PLANT PHENOTYPING AND PHENOMICS FOR PLANT BREEDING

EDITED BY: Gustavo A. Lobos, Anyela V. Camargo, Alejandro del Pozo, Jose L. Araus, Rodomiro Ortiz and John H. Doonan

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

PLANT PHENOTYPING AND PHENOMICS FOR PLANT BREEDING

Topic Editors:
Gustavo A. Lobos, Universidad de Talca, Chile
Anyela V. Camargo, National Institute of Agricultural Botany, United Kingdom
Alejandro del Pozo, Universidad de Talca, Chile
Jose L. Araus, University of Barcelona, Spain
Rodomiro Ortiz, Swedish University of Agricultural Sciences, Sweden
John H. Doonan, Aberystwyth University, United Kingdom

Plant phenotyping under controlled conditions.
Image: National Plant Phenomics Centre (NPPC, Aberystwyth University, United Kingdom).

As a consequence of the global climate change, both the reduction on yield potential and the available surface area of cultivated species will compromise the production of food needed for a constant growing population. There is consensus about the significant gap between world food consumption projected for the coming decades and the expected crop yield-improvements, which are estimated to be insufficient to meet the demand.

The complexity of this scenario will challenge breeders to develop cultivars that are better adapted to adverse environmental conditions, therefore incorporating a new set of morpho-physiological and physico-chemical traits; a large number of these traits have been found to be linked to heat and drought tolerance.

Currently, the only reasonable way to satisfy all these demands is through acquisition of high-dimensional phenotypic data (high-throughput phenotyping), allowing researchers with a holistic comprehension of plant responses, or 'Phenomics'.

Phenomics is still under development. This Research Topic aims to be a contribution to the progress of methodologies and analysis to help understand the performance of a genotype in a given environment.

Citation: Lobos, G. A., Camargo, A. V., del Pozo, A., Araus, J. L., Ortiz, R., Doonan, J. H., eds. (2018). Plant Phenotyping and Phenomics for Plant Breeding. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-532-4

First Latin-American Conference on Plant Phenotyping and Phenomics for Plant Breeding
November 30, December 1 & 2, 2015

Hosted by

Universidad de Talca, Talca, Chile

Conveners

Dr. Gustavo A. Lobos
Plant Breeding and Phenomic Center
Facultad de Ciencia Agrarias
Universidad de Talca
Talca
Chile

Dr. Anyela Camargo
National Plant Phenomics Centre
IBERS, Plas Gogerddan
Aberystwyth University
Aberystwyth
United Kingdom

Supporters

Sponsors

Table of Contents

3. NOVEL METHODOLOGICAL APPROACHES AND SOFTWARE DEVELOPMENT

frontiers
in Plant Science

Editorial: Plant Phenotyping and Phenomics for Plant Breeding

Gustavo A. Lobos [1]*, Anyela V. Camargo [2]*, Alejandro del Pozo [1], Jose L. Araus [3], Rodomiro Ortiz [4] and John H. Doonan [5]

[1] *PIEI Adaptación de la Agricultura al Cambio Climático, Facultad de Ciencias Agrarias, Plant Breeding and Phenomic Center, Universidad de Talca, Talca, Chile, [2] The John Bingham Laboratory, Genetics and Breeding, National Institute of Agricultural Botany, Cambridge, United Kingdom, [3] Plant Physiology Section, University of Barcelona, Barcelona, Spain, [4] Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden, [5] National Plant Phenomics Centre, Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, United Kingdom*

Editorial on the Research Topic

Plant Phenotyping and Phenomics for Plant Breeding

INTRODUCTION

A major challenge for food production in the coming decades is to meet the food demands of a growing population (Beddington, 2010). The difficulty of expanding agricultural land, along with the effect of climate change and the increase in world population are the current societal changes that make necessary to accelerate research to improve yield-potential and adaptation to stressful environments (Lobos et al., 2014; Camargo and Lobos). Increasing yields will require implementing novel approaches in gene discovery and plant breeding that will significantly increase both production per unit of land area and resource use efficiency (Parry and Hawkesford, 2010; Tanger et al., 2017). A critical component for accelerating the development of new and improved cultivars is the rapid and precise phenotypic assessment of thousands of breeding lines, clones or populations over time (Fu, 2015) and under diverse environments. The only reasonable way to satisfy all these demands is through acquisition of high-dimensional phenotypic data (high-throughput phenotyping) or "phenomics" (Houle et al., 2010). This approach may predict complex characters that are relevant for plant selection (forward phenomics), and will also provide explanations as to why given genotypes stands out in a specific environment (reverse phenomics) (Camargo and Lobos).

Phenotype can be defined as the characteristics of an organism resulting from the interaction between genotype, environment and crop management. Phenomics involves the gathering of appropriate phenotypic data at multiple levels of organization, to progress toward a more complete characterization of phenotypic space generated by a particular genome or set of genomes (Dhondt et al., 2013). Thus, plant phenotyping can operate at different levels of resolution and dimensionality, from the molecular to the whole plant, and in different environments, from controlled to field conditions. Although, each level focuses on particular traits, the ultimate goal is to integrate knowledge from the bottom up to produce cultivars with higher performance. In that regard, the use of plant phenotyping methods as part of breeding programs has become a powerful research tool to help breeders generate cultivars more adaptable to diverse challenging environmental scenarios (Camargo and Lobos).

This Research Topic (RT) issue is based on contributions from invited speakers to the First Latin-American Conference on Plant Phenotyping and Phenomics for Plant Breeding carried out during 2015 at Universidad de Talca (Talca, Chile), and from other scientists who are currently researching on phenomics and plant breeding.

Thus, the categories and scope are diverse (review, perspective, and original research) and address different objectives through various levels of resolution and dimensionality. Interestingly, even though most of the phenotyping and phenomics for plant breeding research have been developed for model plants and cereals (Lobos and Hancock; Camargo and Lobos), this RT highlights the feasibility of implement these approaches on breeding programs targeted to other crop species.

CHARACTERIZATION OF THE PLANT: FROM THE GENE TO POPULATION RESPONSES BY REMOTE SENSING

Many of the articles comment on knowledge gained from model plants that was applied to crops. For example, Liu et al. proposed a pathway model for trichome development in cucumber and compared it with the model from *Arabidopsis thaliana*. Yu et al. identified *VaERD15* as a transcription factor gene associated with cold-tolerance in Chinese wild *Vitis amurensis*, and the expression levels increased after low-temperature treatment, enhancing cold tolerance. Cao et al. analyzed gene expression on transgenic material to unveil the functional roles of four expressed *FT*-like genes in tomato. These authors also demonstrated the functional differentiation between *FT*-like genes in controlling flowering through overexpression in *A. thaliana* and VIGS-mediated knocking down in tomato. Awlia et al. were able to show further that phenotyping multiple quantitative traits in one experimental setup can provide new insights into the dynamics of plant responses to stress, suggesting the use of forward genetics studies to identify genes underlying early responses to stress.

The small grain cereals are very well represented in this RT issue. For example, Wu et al. found that overexpressing *OsDof12* in rice could lead to reduced plant height, erected leaf, shortened leaf blade, and smaller panicle resulted from decreased number of primary and secondary branches. Barber et al. performed 1-day transfers of pot-grown wheat to replicated controlled environments, providing strong evidence that the key phases susceptible to heat stress at booting and anthesis in wheat are discrete and that genotypes vary with regards to the most susceptible growth stage; at anthesis, the north European allele *Rht-D1b* was related with higher tolerance to heat stress. Adriani et al. analyzed the effect of different quantitative trait loci (QTL) and their interaction with growing conditions on panicle size and number in rice. Their results showed that grain production was enhanced by *qTSN* only under shading conditions, where panicle number was not affected while photosynthesis and starch storage in internodes were enhanced. Similarly, Camargo et al. established the value of systematically phenotyping genetically unstructured populations to reveal the genetic architecture underlying morphological variation in commercial wheat, QTL with phenological characters such as heading, and the onset of flag leaf senescence, as well as morphological traits such as stem height. In order to have consistency between stress conditions and seasons, del Pozo et al. highlighted the importance of defining and deeply characterizing the target environment before determining the set of phenotyping traits for selection.

Integration of molecular analyses with whole plant phenotyping deepens our understanding of responses to environmental variables as demonstrated by Sanchez-Bragado et al. who proposed the combination of δ^{15}N and N content as an affordable tool to phenotype the relative contribution of different plant parts to the grain N in wheat under contrasting water and nitrogen conditions. Similarly, as selection criteria in breeding N-efficient rice cultivars, Wu et al. suggested that promoting pre-heading growth could increase total nitrogen uptake at maturity, while high biomass accumulation during the grain filling period and large panicles are important for higher nitrogen use efficiency for grain production. Nigro et al. also suggested the use of glutamine synthetase activity and expression as a candidate proxy to select genotypes having high grain protein content in wheat. Fisher et al. found that at early stages of drought stress, metabolic profiling could be used as an efficient tool to discriminate among tolerant or susceptible genotypes of the model plant *Brachypodium distachyon*. Medina et al. concluded for wheat that the combination of phenotyping and gene expression analysis is a useful approach to identify phenotype-genotype relationships and their behavior in response to different environments, which mostly follows from the combination of water regimes and CO_2 levels during vegetative stages.

Kumar et al. used molecular, phenotypic, and geographical diversity to develop a compact composite core collection in the oilseed crop safflower (*Carthamus tinctorius* L.) that will facilitate the identification of genetic determinants of trait variability. Mora et al. emphasized that large number of spurious QTL could be detected when the genetic covariance matrix is ignored in the mixed model analysis, increasing the rate of false-positives. In strawberry breeding programs, Hancock et al. proved that much of the cost associated with DNA marker discovery for marker-aided breeding (MAB) can be eliminated if a diverse, segregating population is generated, genotyped, and made available to the global breeding community. Tan et al. proposed the use of Digital Gen Expression (DGE) as an efficient tool to find differences in transcriptional responses of different tissues/organs of castor plants (*Ricinus communis* L.) subjected to stress, but also to understand molecular mechanisms associated to sex variation. Among approaches to improve yield potential, delayed leaf senescence or stay-green attributes were also addressed in this RT (Balakrishnan et al., 2016; Camargo et al.; Fisher et al.; Wu et al.; Yang et al.; del Pozo et al.).

NOVEL METHODOLOGICAL APPROACHES AND SOFTWARE DEVELOPMENT

This issue also considers methodological approaches to characterize and classify cells and to quantify fluorescence at sub-cellular level (Hall et al.). Likewise, using a wheat panel of elite cultivars and non-adapted genetic resources growing under different adverse environments, Tattaris et al. compared on-ground proximal assessments of canopy temperature (CT) and NDVI with data collected from unmanned aerial vehicles (UAV) and satellite; considering statistical analyses, cost and the feasibility of performing measurements of a high number of

genotypes at any moment, the authors recommend the use of UAV for plant breeding purposes. Vergara-Díaz et al. highlighted the advantages of using spectral reflectance indices (SRI) derived from Red-Green-Blue (RGB) digital images as a low-cost tool for prediction of several traits (e.g., grain yield, leaf N concentration and the ratio of carbon to nitrogen) that are highly valuable for maize breeders. Likewise, Garriga et al. proposed the use of classification methods (PLS-DA) as an efficient tool to directly identify the elite genotypes group instead to focus on the estimation of predicted trait-values.

Software developments were also highlighted in this RT. For example, Deery et al. showed the efficiency of a custom-developed airborne thermography image processing software (ChopIt), used for plot segmentation and extraction of canopy temperature from each individual plot by a non-technical user. Kaushik et al. performed a deep characterization of eggplant using the Tomato Analyzer (Rodríguez et al., 2010). Using Spectral Knowledge (SK-UTALCA; Lobos and Poblete-Echeverría, 2017), Garriga et al. performed an exploratory analysis of high-resolution spectral reflectance data, testing 255 SRI at the same time. RGB images were filtered by a retina filter (Benoit et al., 2010), which enhanced the contrast between plant and background, improving color consistency and providing spatial noise removal and luminance correction (Fisher et al.). Awlia et al. used the PlantScreen™Compact System (PSI, Czech Republic) equipped with chlorophyll fluorescence and RGB imaging, as well as with automatically weighing and watering of plants to test the response of *A. thaliana* to salinity. The open source Breedpix 0.2 software (Casadesus et al., 2007) was used by Vergara-Díaz et al. for the calculation of several RGB vegetation indices based on the different properties of color inherent in RGB images.

CONCLUSIONS

This RT highlights the importance of a holistic characterization of the crop, from the cell to the plant level using for that current available tools for plant phenotyping. Gathering, integrating, and making inferences on these data will help breeding programs to speed up the release of cultivars tolerant to stress environments. Alongside more data, training programs should be established to guarantee the use and adoption of new technologies.

In addition to aspects associated to phenomics and breeding, the RT also discussed future challenges such as the need for multi-disciplinary research and the better and deeper characterization of the environment. The progress of forward and reverse phenomics has great potential to accelerate the improvement of yield potential and we expect to see that rapid developments will continue in this subject area, especially in Latin America.

AUTHOR CONTRIBUTIONS

GL, AC, AdP, JA, RO, and JD co-wrote this editorial based on the various contributions to this Research Topic.

REFERENCES

Balakrishnan, D., Subrahmanyam, D., Badri, J., Raju, A. K., Rao, Y. V., Beerelli, K., et al. (2016). Genotype × environment interactions of yield traits in backcross introgression lines derived from Oryza sativa cv. Swarna/*Oryza nivara*. *Front. Plant Sci.* 7:1530. doi: 10.3389/fpls.2016.01530

Beddington, J. (2010). Food security: contributions from science to a new and greener revolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 61–71. doi: 10.1098/rstb.2009.0201

Benoit, A., Caplier, A., Durette, B., and Herault, J. (2010). Using human visual system modeling for bio-inspired low level image processing. *Comput. Vis. Image Underst.* 114, 758–773. doi: 10.1016/j.cviu.2010.01.011

Casadesus, J., Kaya, Y., Bort, J., Nachit, M. M., Araus, J. L., Amor, S., et al. (2007). Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water–limited environments. *Ann. Appl. Biol.* 150, 227–236. doi: 10.1111/j.1744-7348.2007.00116.x

Dhondt, S., Wuyts, N., and Inzé, D. (2013). Cell to whole-plant phenotyping: the best is yet to come. *Trends Plant Sci.* 18, 428–439. doi: 10.1016/j.tplants.2013.04.008

Fu, Y.-B. (2015). Understanding crop genetic diversity under modern plant breeding. *Theor. Appl. Genet.* 128, 2131–2142. doi: 10.1007/s00122-015-2585-y

Houle, D., Govindaraju, D. R., and Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet.* 11, 855–866. doi: 10.1038/nrg2897

Lobos, G. A., Matus, I., Rodriguez, A., Romero-Bravo, S., Araus, J. L., and del Pozo, A. (2014). Wheat genotypic variability in grain yield and carbon isotope discrimination under Mediterranean conditions assessed by spectral reflectance. *J. Integr. Plant Biol.* 56, 470–479. doi: 10.1111/jipb.12114

Lobos, G. A., and Poblete-Echeverría, C. (2017). Spectral Knowledge (SK-UTALCA): software for exploratory analysis of high-resolution spectral reflectance data on plant breeding. *Front. Plant Sci.* 7:1996. doi: 10.3389/fpls.2016.01996

Parry, M. A. J., and Hawkesford, M. J. (2010). Food security: increasing yield and improving resource use efficiency. *Proc. Nutr. Soc.* 69, 592–600. doi: 10.1017/S0029665110003836

Rodríguez, G. R., Moyseenko, J. B., Robbins, M. D., Morejón, N. H., Francis, D. M., and van der Knaap, E. (2010). Tomato analyzer: a useful software application to collect accurate and detailed morphological and colorimetric data from two-dimensional objects. *J. Vis. Exp.* 37:e1856. doi: 10.3791/1856

Tanger, P., Klassen, S., Mojica, J. P., Lovell, J. T., Moyers, B. T., Baraoidan, M., et al. (2017). Field-based high throughput phenotyping rapidly identifies genomic regions controlling yield components in rice. *Sci. Rep.* 7:42839. doi: 10.1038/srep42839

Latin America: A Development Pole for Phenomics

Anyela V. Camargo[1]* and Gustavo A. Lobos[2]*

[1] The National Plant Phenomics Centre, Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, UK, [2] Facultad de Ciencias Agrarias, Plant Breeding and Phenomic Center, PIEI Adaptación de la Agricultura al Cambio Climático (A2C2), Universidad de Talca, Talca, Chile

Latin America and the Caribbean (LAC) has long been associated with the production and export of a diverse range of agricultural commodities. Due to its strategic geographic location, which encompasses a wide range of climates, it is possible to produce almost any crop. The climate diversity in LAC is a major factor in its agricultural potential but this also means climate change represents a real threat to the region. Therefore, LAC farming must prepare and quickly adapt to an environment that is likely to feature long periods of drought, excessive rainfall and extreme temperatures. With the aim of moving toward a more resilient agriculture, LAC scientists have created the Latin American Plant Phenomics Network (LatPPN) which focuses on LAC's economically important crops. LatPPN's key strategies to achieve its main goal are: (1) training of LAC members on plant phenomics and phenotyping, (2) establish international and multidisciplinary collaborations, (3) develop standards for data exchange and research protocols, (4) share equipment and infrastructure, (5) disseminate data and research results, (6) identify funding opportunities and (7) develop strategies to guarantee LatPPN's relevance and sustainability across time. Despite the challenges ahead, LatPPN represents a big step forward toward the consolidation of a common mind-set in the field of plant phenotyping and phenomics in LAC.

Keywords: LAC, climate change, genomic, phenotyping, plant breeding, LatPPN

WHY PHENOMICS IS KEY TO FACE CLIMATE CHANGE AND FOOD SECURITY?

In the past decades, climatic variations related to El Niño or La Niña phenomena have brought serious challenges to the agricultural sector in LAC. While drought is the main threat to food production associated to La Niña, El Niño can cause heavy rains, flooding or extremely hot or cold weather (Allen and Ingram, 2002). In the last 150 years, earth's temperature increased at a rate of $0.045°C$ per decade, with almost four-fold ($0.177°C$) in the last 25 years (IPCC, 2007), and will continue to raise by another $1.1–6.4°C$ over the next century (Jin et al., 2011). This increase in temperature can lead to several agricultural associated problems such as yield reduction as a results of droughts, and the emergence and spreading of plant diseases and pests (FAO, 2016). Therefore, a better use of plant genetic resources and plant breeding (Borrás and Slafer, 2008), are key to tackling the imminent impact of climate change in food security. Further, a multidisciplinary approach that includes disciplines such as omics technologies (e.g., genomics, phenomics, proteomics, and metabolomics), plant physiology, eco-physiology, plant pathology and entomology, and soil science will be critical to increase crop resilience to climate change (Reynolds et al., 2016).

Undoubtedly, public and private breeding programs have the challenge of producing stress tolerant cultivars whose yield potential and quality are also high. In order to increase the chances of producing desirable cultivars, breeders make a high number of crosses (e.g., Chilean wheat breeding programs generate ∼800 crosses per year) and screen them under a limited number of environmental conditions (Araus and Cairns, 2014). Line crossing is a common experimental design for mapping quantitative trait loci (QTLs) in plant breeding. Crosses are initiated from at least two inbred lines, such as backcrosses, F2, and more derived generations (Xie et al., 1998). To increase the statistical inference space of the estimated QTL variance and ensure that polymorphic alleles are present in the parental gene pool, a sufficient number of parents must be sampled (Muranty, 1996). The number of traits measured per plot is normally limited to the size of the population. Increasing the number of traits to be measured requires additional time, resources and the use of skilled labor (Kipp et al., 2014). This represents a limitation toward the understanding of the interaction genotype × environment (G × E) (Furbank and Tester, 2011; Yang et al., 2014; Großkinsky et al., 2015; Rahaman et al., 2015).

Although, genome sequencing has become relatively fast, cheap, and easy to produce, plant phenomics still lags behind. This unbalance has become a bottleneck in the understanding of G × E and it also limits the possibility of carrying out tests under field conditions (Lobos and Hancock, 2015). Therefore, there is a need to incorporate the evaluation of multiple morpho-physiological and physico-chemical traits at the high-throughput level to be able to understand for example pleiotropy or genomic variants that gave rise to a particular phenotype (Houle et al., 2010; Fahlgren et al., 2015).

Due to the cost of high-throughput plant phenotyping, several international phenotyping networks have been established with the idea of joining efforts and produce research with impact. Some of the most prominent networks are: the European Plant Phenotyping Network (EPPN), Food and Agriculture COST Action FA1306, the International Plant Phenotyping Network (IPPN), the Australian Phenomics Network (APN), the German Plant Phenotyping Network (DPPN) and the U.K. Plant Phenomics Network (UKPPN). In Asia, the 1st Asia-Pacific Plant Phenotyping will be held in Beijing, China in October 2016 and the 3rd International Plant Phenotyping Symposium was held in Chennai, India in 2014. More recently in North America, the United States of America recently launched the North American Plant Phenotyping Network (NAPPN).

DOES LATIN AMERICA AND THE CARIBBEAN NEED TO WORRY ABOUT PHENOMIC DEVELOPMENT?

Latin America is a region that includes Mexico, the Spanish/Portuguese speaking countries in Central America and the whole of South America, as well as the Caribbean (Latin America and the Caribbean—LAC). The region is highly heterogeneous in terms of climate, ecosystems, human population distribution, politics, economy and incomes, and cultural traditions. Out of a total of 17 megadiverse countries identified by the World Conservation Monitoring Centre (http://www.unep-wcmc.org), six are in Latin American, namely Brazil, Colombia, Ecuador, Mexico, Peru, and Venezuela. Furthermore, from the eight primary centers of origin and diversity, numbers VII (South Mexican and Central American) and VIII (South America Andes region: Bolivia, Peru, Ecuador; VIIIa The Chilean Center, and VIIIb Brazilian-Paraguayan Center) are based in the region (Vavilov, 1992).

Due to LAC's diverse geography, climate change will impact the region severely. Compared to pre-industrial times, it is estimated that the mean temperature on the region will increase about 4.5°C by the end of the century (Reyer et al., 2015). Temperatures are expected to increase dramatically in the tropics and moderate at the subtropical regions in the north (Mexico) and south (southern Chile, Argentina and Uruguay) (Reyer et al., 2015). Annual precipitations are also likely to increase in Argentina, Uruguay, Brazil, Peru, Ecuador, and Colombia and decrease in the rest of the countries (Reyer et al., 2015). These changes have a direct impact on agricultural crop yields. It's expected that crops such as wheat, soybean and maize will reduce its yield potential, while others such as rice and sugar cane will increase it (Fernandes et al., 2012; Marin et al., 2012).

The economic development of the regions where plant phenotyping and phenomics have been developed in the last 10 years (high-income countries) is completely different to that of LAC. According to the World Bank, around 37% of the LAC population lives under poverty or extreme poverty (World Bank, 2014), and near 60% of the people living in rural areas is under extreme poverty (RIMISP, 2011). Therefore, besides the climate change effects impacting LAC agriculture, there is also a significant knock-on the region economy, affecting particularly the lower socioeconomic strata (Ortiz, 2012).

Although, LAC countries are wealthier, government efforts are mainly focused on priority areas such as education, health, employability, and infrastructure. Research and innovation in areas such as agriculture has been given a low priority. As a result, most Latin American farmers do not have the resources or the support to effectively adapt to a changing climate that is already showing its negative impact in agriculture (Lobos and Hancock, 2015). Therefore, LAC scientists and private sector must work together to develop strategies aiming at moving toward a more resilient agriculture, and one of them is the use of plant phenomics and phenotyping for breeding.

Phenomics has become a powerful research tool to help breeders to generate cultivars adaptable to more challenging environmental scenarios. In the past decade, phenomics has been focused mainly on breeding of grain crops, but their application in other species of relevance for LAC (e.g., fruit, vegetables, forage and others) is almost absent (Lobos and Hancock, 2015).

The potential of recent advances in phenomics encouraged the Plant Breeding and Phenomic Center (Dr. Gustavo A. Lobos, Universidad de Talca, Talca, Chile) and the National Plant Phenomics Centre (Dr. Anyela Camargo, IBERS, Aberystwyth University, U.K.) to organize the First Latin American Conference on Plant Phenotyping and Phenomics for Plant Breeding (November 30st to December 2nd 2015,

Talca, Chile). This event had three main goals: (1) bring to Latin American researchers and students, international keynote speakers and plant breeding companies from around the world, to present their ongoing work on plant phenomics and phenotyping for plant breeding; (2) perform a workshop to train Latin American scientists and postgraduate students in the use of key plant phenotyping tools, the analysis of data and the mapping of traits to the genome; and (3) set up the Latin American Plant Phenomics Network (LatPPN), conceived to facilitate the training on high-throughput phenotyping and pre-breeding methodologies, scientific exchange of young/senior researchers and students, and to improve access to resources and research facilities.

The conference covered a broad range of topics such as pre-breeding and breeding strategies, methods to measure and analyse trait data for plant breeding and the strategies to translate research from the bench to the field. International keynote speakers gave seminal talks and chaired the track of their expertise. Challenges and opportunities were also explored such as the handling of the high amount of data generated through high-throughput phenotyping. Multiple ideas were discussed to deal with every particular challenge. Participants also had the opportunity to attend five workshops that covered aspects such as the use of software and equipment for plant phenotyping (mainly by remote sensing), and data handling and manipulation.

The LatPPN, which is chaired for two years by Chile (Dr. Gustavo A. Lobos) and Colombia (Dr. Anyela Camargo), had it first reunion during the 3rd day of the conference. Representatives from LAC (Argentina, Brazil, Chile, Colombia, Ecuador, Mexico, and Uruguay) and from other countries (Australia, Germany, Saudi Arabia, Spain, U.K., and U.S.A.) got together to discuss what LAC's breeding programs needed to do to become more efficiency in terms of plant phenotyping and phenomics. They also discussed the differences between phenotyping and the more complex concept of phenomics. This discussion helped to define where LAC currently stands (more focused on the phenotyping of few traits and low number of genotypes) and where it needs to be in the future (mostly oriented to the multidimensional approach of phenomics, considering a high number of genotypes assessed). For example, the wheat breeding program of INIA Chile, used to consider a classical approximation where the numbers of traits evaluated increases insofar the number of generation progresses: ∼9 traits at F2–F5: susceptibility to *Puccinia triticina*, *P. graminis*, and *P. striiformis*, plant height, tillering capacity, type of spike, grain color, type of grain, and black point or other grain defects; ∼16 at F6–F8: previous ones plus heading date, grain yield, and some grain characteristics such as test weight, protein and gluten content, sedimentation, and seed hardiness; and ∼19 at F9–F10 where less than 5% of the original crosses are evaluated: previous ones plus some other required by millers such as W flour value, falling number, and some bakery aptitudes. Today, using spectrometry and thermography, this breeding program is aimed to predict some of these traits but also to consider other 30 morpho-physiological and physico-chemical characters (some examples covered in next section), screening ∼800 genotypes per day.

IS LAC ORIENTED TO PHENOMICS OR PLANT PHENOTYPING?

Due to resources' availability such as equipment, skills and infrastructure, LAC has mainly focused on plant phenotyping. Although phenomics in LAC has not yet had a proper expansion, there are some good examples of institutions focusing on it: (i) The International Maize and Wheat Improvement Center (CIMMYT—Mexico) routinely uses remote sensing and high spec sensor technologies to screen for wheat and maize's responses to biotic and abiotic stresses, among them yield and its components, biomass, senescence (stay-green), water stress, and water use efficiency, canopy cover, photosynthetic capacity and activity (Zaman-Allah et al., 2015). Special emphasis is also put on 3D reconstruction for plant height, spike number and biomass determination; (ii) The Plant Breeding and Phenomic Center (University of Talca—Chile) have focused its efforts on the prediction of physiological traits by spectrometry and thermography (e.g., gas exchange, modulated chlorophyll fluorescence, pigments concentration, stem water potential, hydric and osmotic cell potential, cell membrane stability, lipid peroxidation, proline content, C and O isotopic composition) on several breeding programs (wheat, blueberries, alfalfa, strawberries, and quinoa) oriented to abiotic stresses (salt, water deficit and high temperature) (Garriga et al., 2014; Lobos et al., 2014; Estrada et al., 2015; Hernandez et al., 2015), developing also a software for exploratory analysis of high-resolution spectral reflectance data on plant breeding (Lobos and Poblete-Echeverría, in press).

In terms of phenotyping, most research institutes across the region have done some form of low to medium throughput phenotyping, for example: (i) The International Centre for Tropical Agriculture (CIAT—Colombia) is screening root architecture to identify markers associated to drought stress tolerance in beans and grasses (Villordo-Pineda et al., 2015; Rao et al., 2016); (ii) Embrapa (Brazil) uses traditional phenotyping to screen for root morphology in wheat (Richard et al., 2015); (iii) Universidade Federal de Mato Grosso, Brazil, uses traditional phenotyping tools (e.g., gas exchange measurements) to look for photosynthetic responses of tree species to seasonal variations in hydrology in the Brazilian Cerrado and Pantanal (Dalmagro et al., 2016); (iv) Researchers from Argentina uses conventional phenotyping equipment to investigate the response of seed weight and composition to changes in assimilate supply from leaves, to the incident solar radiation reaching the pods and to the combination of both, changes in assimilate supply from the leaves and incident solar radiation on pods of soybean plants (Bianculli et al., 2016), they are also trying to develop low cost tools in order to make that technology accessible to researches from LAC; (v) The International Potato Center (Peru) have improved the screening of potato breeding lines by spectroscopy (Ayvaz et al., 2016); and (vi) INIA (Uruguay) in collaboration with INIA (Chile) and the Plant Breeding and Phenomic Center (University of Talca—Chile), applied genotyping-by-sequencing to identify single-nucleotide polymorphisms, in the genomes of 384 wheat genotypes that were field tested in Chile under three different water regimes (Lado et al., 2013).

HOW WILL LAC BENEFIT FROM LATPPN?

The conference served as a platform to showcase LAC capabilities, investigate strengths, and weaknesses, and thereby identify where the challenges lie and what the knowledge and the technological gaps between the region and the rest of the world are.

Given LAC's high heterogeneity in terms of climate, ecosystems and genetic diversity, as well as the differences of each country vulnerability to climate change, it was agreed how important it is for LAC's agri-food chain to take a more proactive role in the development of strategies leading to the selection of crops capable to withstand the impact of climate change.

With the aim of identifying what LatPPN needed to do to strength LAC's plant phenotyping and phenomics research, the panel of participants identified the following key challenges: (i) develop LatPPN's own tailored identity: there is not a common crop but rather a wide diversity of them, from grasses to forest species. As previously mentioned, plant phenotyping and phenomics has been developed almost exclusively on cereal improvement, however LatPPN needs to focus on other breeding programs that are important for particular countries. For example: blueberries for Chile (Chile is the biggest exporter of fresh blueberries in the world, ~90,000 ton during 2015/16), potato for Peru (production was estimated to be 4.5 million tons for 2015), tangerines for Uruguay (production was ~6000 tons in 2014), pineapple for Costa Rica (since 2000, pineapple production has increased by nearly 300%, however production is very inefficient, each plant only produces two fruit over a period of 18–24 months, and requires significant amount fertilizer to do so) and Coffee for Colombia (exports account for ~810,000 ton in 2015) and Brazil (exports account for ~2.6 million tons in 2014). The production of these cash crops will face serious challenges (e.g., post-harvest life, or the incidence of physiological disorders, pests and diseases) in the coming decades due to the sensitivity of them to water shortages and heat stress. In this meeting, it was also highlighted: (ii) training on plant phenomics and phenotyping using strategies that allow the participation of several countries at the same time. We are aiming at finding resources to implement distance-training courses using currently available technologies such as webinars and teleconferences; (iii) learn from experienced researchers and current plant phenotyping and phenomics initiatives. In order to facilitate the interaction between researchers and institution, senior researchers on plant phenotyping and phenomics were invited to participate in the first meeting; (iv) since high-throughput phenotyping requires a broad range of capabilities (e.g., programmers, bioinformaticians, statisticians, biologists, agronomists, geneticists, physiologists), is important to promote interdisciplinary work between researchers; (v) identify the state of art of plant phenotyping and phenomics in LAC. In order to identify strengths, opportunities and weaknesses and develop targeted strategies, key information such as breeding programs, researchers, equipment and infrastructure, regional and local financial sources, and capabilities should be surveyed. All this information should be included on the future LatPPN webpage; (vi) distribute efforts on common goals (e.g., researchers from different countries working on the same species or problem), it will be necessary to standardize measurements and protocols; (vii) sharing of equipment and infrastructure; and (viii) LatPPN visibility and presence. To avoid early disenchantment, LatPPN needs to carry out activities to promote the network (e.g., events, postgraduate grants or proposal calls).

In relation to weaknesses, the lack of a permanent budget to run network activities is one of LatPPN's main concerns. Currently, the Director and Co-Director, the executive committee (Dr. Paulo Hermann from EMBRAPA—Brazil and Dr. Gustavo Pereyra from INTA-CONICET—Argentina), and the representative members (three per country in charge of meet the local demands, thematic promotion, and economic resources leveraging) devote part of their time and resources to consolidate the network. However, they are looking into sources of support within LAC and worldwide. At the country level, there are a number of countries that have access to grants provided by their own governments. At regional level, there are a number of organizations such as PROCISUR and PROCITROPICOS, which provide regular grant support for agricultural research initiatives. At international level, there are several organizations such as FAO (the Food and Agricultural Organization), EU (the European Union), and IBS (the Inter-American Development Bank) who support agricultural research in LAC.

Another weakness is LAC's low publication rate and the lack of accessibility of LAC institutions to main bibliographic databases. According to the World Bank, the number of publications produced by the most important economies in LAC in 2012 was 48,622 from Brazil, 13,112 from Mexico, 8,053 from Argentina, 5,158 from Chile and 4,456 from Colombia. Brazil is the only country whose output is equivalent to high-income countries where phenomics have been developing in the last 10 years; U.S.A. (412,542), Germany (101,074), U.K. (97,332), France (72,555), Spain (53,342), and Australia (47,806) (World Bank, 2012). In term of access to bibliographic databases, most of the institutions in the region have limited or no access to main bibliographic databases such as Scopus and Web of Knowledge. This is serious limitation to the dissemination of the work developed in LAC, especially if we are aiming at improving plant breeding programs through the use of plant phenotyping and phenomics.

Despite the weaknesses, currently there are several international research institutes who are already formally collaborating with LAC on plant phenotyping and phenomics. Some of them are, Lemnatec (Germany), CSIRO (Australia), IBERS (U.K.), Universidad de Barcelona (Spain), the Julich Plant Phenomics Centre (German), and the James Hutton Institute (U.K.).

The establishment of LatPPN represented a big step forward toward the consolidation of a common mind-set in the field of plant phenotyping and phenomics across LAC. Clearly there are more opportunities than disadvantages, and each weakness needs to be addressed having in mind a regional approach.

CONCLUSIONS AND FUTURE WORK

Phenomics can complement the potential of new molecular/ genotyping technologies, and together with agronomy and plant breeding efforts would be a real contribution to develop new strategies to help mitigate the impact of climate change in agriculture. There are major opportunities for phenomics in LAC, not only because it has been adopted in isolated initiatives, but also as worldwide development has focused mainly on grain breeding programs. LAC researchers have identified the need to collaborate to exploit the opportunities and gathered together to organize the Latin American Plant Phenomics Network (LatPPN). Currently, LatPPN has prioritized the work on several fronts to consolidate the network (e.g., grant application to CYTED and Procisur, LatPPN's second meeting in April 2016 (Balcarce, Argentina) and planning a second regional conference organized by EMBRAPA during 2017, drafting of LatPPN's survey, drafting of LatPPN's white paper, and construction of LatPPN's webpage). What follows next is the development of strategies leading to the sustainability of the network. We are aware of the work ahead of us and know that the collaboration within LatPPN members and with other networks will be crucial to build on the foundations laid.

AUTHOR CONTRIBUTIONS

Both authors contributed equally.

ACKNOWLEDGMENTS

REFERENCES

Allen, M. R., and Ingram, W. J. (2002). Constraints on future changes in climate and the hydrologic cycle. *Nature* 419, 224–232. doi: 10.1038/nature01092

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Ayvaz, H., Bozdogan, A., Giusti, M. M., Mortas, M., Gomez, R., and Rodriguez-Saona, L. E. (2016). Improving the screening of potato breeding lines for specific nutritional traits using portable mid-infrared spectroscopy and multivariate analysis. *Food Chem.* 211, 374–382. doi: 10.1016/j.foodchem.2016.05.083

Bianculli, M. L., Aguirrezábal, L. A. N., Pereyra Irujo, G. A., and Echarte, M. M. (2016). Contribution of incident solar radiation on leaves and pods to soybean seed weight and composition. *Eur. J. Agron.* 77, 1–9. doi: 10.1016/j.eja.2016.03.002

Borrás, L., and Slafer, G. A. (2008). Agronomy and plant breeding are key to combating food crisis. *Nature* 453, 1177. doi: 10.1038/4531177a

Dalmagro, H. J., de A. Lobo, F., Vourlitis, G. L., Dalmolin, C. Â., Antunes, M. Z. Jr., et al. (2016). Photosynthetic response of a wetland- and an upland-adapted tree species to seasonal variations in hydrology in the Brazilian Cerrado and Pantanal. *Acta Physiol. Plant.* 38:107. doi: 10.1007/s11738-016-2125-7

Estrada, F., Escobar, A., Romero-Bravo, S., González-Talice, J., Poblete-Echeverría, C., Caligari, P. D. S., et al. (2015). Fluorescence phenotyping in blueberry breeding for genotype selection under drought conditions, with or without heat stress. *Sci. Hortic.* 181, 147–161. doi: 10.1016/j.scienta.2014.11.004

Fahlgren, N., Gehan, M. A., and Baxter, I. (2015). Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Curr. Opin. Plant Biol.* 24, 93–99. doi: 10.1016/j.pbi.2015.02.006

FAO (Food and Agriculture Organization of the United Nations) (2016). *El Niño*. Available online at: http://www.fao.org/el-nino/en/ [Accessed May 04, 2016].

Fernandes, E. C. M., Soliman, A., Confalonieri, R., Donatelli, M., and Tubiello, F. (2012). *Climate Change and Agriculture in Latin America, 2020–2050*. Washington, DC: World Bank.

Furbank, R. T., and Tester, M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Garriga, M., Retamales, J. B., Romero, S., Caligari, P. D. S., and Lobos, G. A. (2014). Chlorophyll, anthocyanin and gas exchangechanges assessed by spectroradiometry in *Fragaria chiloensis* under salt stress. *J. Integr. Plant Biol.* 56, 505–515. doi: 10.1111/jipb.12193

Großkinsky, D. K., Svensgaard, J., Christensen, S., and Roitsch, T. (2015). Plant phenomics and the need for physiological phenotyping across scales to narrow the genotype-to-phenotype knowledge gap. *J. Exp. Bot.* 66, 5429–5440. doi: 10.1093/jxb/erv345

Hernandez, J., Lobos, G. A., Matus, I., del Pozo, A., Silva, P., and Galleguillos, M. (2015). Using ridge regression models to estimate grain yield from field spectral data in bread wheat (*Triticum aestivum* L.) grown under three water regimes. *Remote Sens.* 7, 2109–2126. doi: 10.3390/rs70202109

Houle, D., Govindaraju, D. R., and Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet.* 11, 855–866. doi: 10.1038/nrg2897

IPCC (2007). *Climate Change 2007: Impacts, Adaptation and Vulnerability*. Contribution of Working Group II to the 4th Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press.

Jin, B., Wang, L., Wang, J., Jiang, K. Z., Wang, Y., Jiang, X. X., et al. (2011). The effect of experimental warming on leaf functional traits, leaf structure and leaf biochemistry in *Arabidopsis thaliana*. *BMC Plant Biol.* 11:35. doi: 10.1186/1471-2229-11-35

Kipp, S., Mistele, B., Baresel, P., and Schmidhalter, U. (2014). High-throughput phenotyping early plant vigour of winter wheat. *Eur. J. Agron.* 52, Part B, 271–278. doi: 10.1016/j.eja.2013.08.009

Lado, B., Matus, I., Rodríguez, A., Inostroza, L., Poland, J., Belzile, F., et al. (2013). Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data. *G3 (Bethesda)* 3, 2105–2114. doi: 10.1534/g3.113.007807

Lobos, G. A., and Hancock, J. F. (2015). Breeding blueberries for a changing global environment: a review. *Front. Plant Sci.* 6:782. doi: 10.3389/fpls.2015.00782

Lobos, G. A., Matus, I., Rodriguez, A., Romero-Bravo, S., Araus, J. L., and del Pozo, A. (2014). Wheat genotypic variability in grain yield and carbon isotope discrimination under Mediterranean conditions assessed by spectral reflectance. *J. Integr. Plant Biol.* 56, 470–479. doi: 10.1111/jipb.12114

Lobos, G. A., and Poblete-Echeverría, C. (in press). Spectral Knowledge (SK-UTALCA): software for exploratory analysis of high-resolution spectral reflectance data on plant breeding. *Front. Plant Sci.*

Marin, F. R., Jones, J. W., Singels, A., Royce, F., Assad, E. D., Pellegrino, G. Q., et al. (2012). Climate change impacts on sugarcane attainable yield in southern Brazil. *Clim. Change* 117, 227–239. doi: 10.1007/s10584-012-0561-y

Muranty, H. (1996). Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Heredity* 76, 156–165. doi: 10.1038/hdy.1996.23

Ortiz, R. (2012). *Climate Change and Agricultural Production.* Technical Notes IDB-TN 383. Inter-American Development Bank, Washington, DC. Available online at: https://issuu.com/idb_publications/docs/technicalnotes_en_64698 [Accessed May 04, 2016].

Rahaman, M. M., Chen, D., Gillani, Z., Klukas, C., and Chen, M. (2015). Advanced phenotyping and phenotype data analysis for the study of plant growth and development. *Front. Plant Sci.* 6:619. doi: 10.3389/fpls.2015.00619

Rao, I. M., Miles, J. W., Beebe, S. E., and Horst, W. J. (2016). Root adaptations to soils with low fertility and aluminium toxicity. *Ann. Bot.* 118, 593–605. doi: 10.1093/aob/mcw073

Reyer, C. P. O., Adams, S., Albrecht, T., Baarsch, F., Boit, A., Canales Trujillo, N., et al. (2015). Climate change impacts in Latin America and the Caribbean and their implications for development. *Reg. Environ. Change* 7, 1–21. doi: 10.1007/s10113-015-0854-6

Reynolds, M. P., Quilligan, E., Aggarwal, P. K., Bansal, K. C., Cavalieri, A. J., Chapman, S. C., et al. (2016). An integrated approach to maintaining cereal productivity under climate change. *Glob. Food Secur.* 8, 9–18. doi: 10.1016/j.gfs.2016.02.002

Richard, C. A., Hickey, L. T., Fletcher, S., Jennings, R., Chenu, K., and Christopher, J. T. (2015). High-throughput phenotyping of seminal root traits in wheat. *Plant Methods* 11:13. doi: 10.1186/s13007-015-0055-9

RIMISP (2011). *Poverty and Inequality: Latin American Report.* Latin American Center for Rural Development (RIMISP), Santiago.

Vavilov, N. I. (1992). *Origin and Geography of Ultivated Plants.* Great Britain: Cambridge University Press.

Villordo-Pineda, E., González-Chavira, M. M., Giraldo-Carbajo, P., Acosta-Gallegos, J. A., and Caballero-Pérez, J. (2015). Identification of novel drought-tolerant-associated SNPs in common bean (*Phaseolus vulgaris*). *Front. Plant Sci.* 6:546. doi: 10.3389/fpls.2015.00546

World Bank (2012). *Scientific and Technical Journal Articles.* Available online at: http://data.worldbank.org/indicator/IP.JRN.ARTC.SC?year_high_desc=true [Accessed June 08, 2016].

World Bank (2014). *Social Gains in the Balance—A Fiscal Policy Challenge for Latin America and the Caribbean.* Washington, DC: LAC Poverty and Labor Brief (February), World Bank.

Xie, C., Gessler, D. D., and Xu, S. (1998). Combining different line crosses for mapping quantitative trait loci using the identical by descent-based variance component method. *Genetics* 149, 1139–1146.

Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., et al. (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* 5:5087. doi: 10.1038/ncomms6087

Zaman-Allah, M., Vergara, O., Araus, J. L., Tarekegne, A., Magorokosho, C., Zarco-Tejada, P. J., et al. (2015). Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* 11:35. doi: 10.1186/s13007-015-0078-2

Breeding blueberries for a changing global environment: a review

Gustavo A. Lobos[1,2]* and James F. Hancock[2]

[1] Faculty of Agricultural Sciences, Plant Breeding and Phenomic Center, Universidad de Talca, Talca, Chile, [2] Department of Horticulture, Michigan State University, East Lansing, MI, USA

Today, blueberries are recognized worldwide as one of the foremost health foods, becoming one of the crops with the highest productive and commercial projections. Over the last 100 years, the geographical area where highbush blueberries are grown has extended dramatically into hotter and drier environments. The expansion of highbush blueberry growing into warmer regions will be challenged in the future by increases in average global temperature and extreme fluctuations in temperature and rainfall patterns. Considerable genetic variability exists within the blueberry gene pool that breeders can use to meet these challenges, but traditional selection techniques can be slow and inefficient and the precise adaptations of genotypes often remain hidden. Marker assisted breeding (MAB) and phenomics could aid greatly in identifying those individuals carrying adventitious traits, increasing selection efficiency and shortening the rate of cultivar release. While phenomics have begun to be used in the breeding of grain crops in the last 10 years, their use in fruit breeding programs it is almost non-existent.

Keywords: Vaccinium, drought, heat, UV, phenotype, highbush, MAB, phenomics

Introduction

Over the last 100 years, the geographical area where highbush blueberries are grown has expanded dramatically (Retamales and Hancock, 2012). The northern highbush blueberry (NHB) is native to the eastern and mid-western portions of the USA, where winters are very cold, summers are moderate and chilling hours are high (**Table 1**). The Industry was first established in New Jersey (1910), but within a few decades had expanded to North Carolina (1920), Michigan (1930), and the Pacific Northwest (1940). From there it leapfrogged to Europe (1970s), New Zealand/Australia (1980s), central Chile (1980s), and most recently China (2000s).

The expansions into the Pacific Northwest, Mexico, and Chile were into climates with much less severe winters, while the introductions into China were into much colder regions. Cultivars developed in Michigan and New Jersey have generally thrived in the milder climates of the Pacific Northwest, but many of them suffer from the high irradiance in Chile and the cold of China.

Southern highbush blueberry (SHB) types were originally developed in the 1980s by incorporating genes from native species from the southern US to reduce the chilling requirement of NHB. SHB were first established in Florida and Georgia (1980s) and then moved to north central Chile (1980), Argentina and Spain (1990), California (2000) and most recently Mexico, Peru and Ecuador (2010s).

The introductions of SHB into California, north Central Chile, and Spain were into hotter and dryer climates than those in Florida and Georgia (**Table 1**), and in southern Chile with much higher UV levels (Huovinen et al., 2006). The expansions into Mexico and Ecuador were from low to

TABLE 1 | Climates of major global highbush and rabbiteye blueberry production regions (adapted from Retamales and Hancock, 2012).

Region	State	City	Rainfall (mm)		Temperature (°C)				Frost-free days	Chilling hours (7°C)
			Annual	Summer	Mid-summer high	Mid-summer low	Mid-winter high	Mid-winter low		
North America (Atlantic)	N. Carolina	Wilmington	1378	478	33.0	21.0	14.0	1.5	246	500–800
North America (Northeastern)	New Jersey	Hammonton	1097	284	30.0	19.0	5.0	-4.5	182	1000+
North America (Midwestern)	Michigan	Holland	1021	294	27.0	14.0	-2.0	-10.0	156	1000+
North America (Southeastern)	Florida, north	Gainesville	1234	495	33.5	22.5	22.0	10.0	285	150–350
	Florida, central	Orlando	1228	528	33.0	21.5	19.0	6.0	315	400–500
	Georgia	Alma	1248	432	33.5	22.0	17.0	5.0	250	450–600
	Mississippi	Poplarville	1606	414	33.5	22.0	15.5	3.5	256	450–600
North America (Northwestern)	British Columbia	Vancouver	1201	134	21.5	13.0	6.0	0.5	170	1000+
	Oregon	Corvallis	1168	92	28.0	13.5	9.0	2.0	190	1000+
	Washington	Vancouver	1267	71	25.0	12.0	7.5	0.0	177	1000+
North America (Southwestern)	California	Bakersfield	163	5	36.0	21.0	13.5	4.0	277	450–550
	Mexico	Guadalajara	927	676	32.4	16.8	26.5	10.2	365	0
Africa	South Africa	Cape Town	515	52	26.1	15.7	17.5	7.0		400–600
Asia	China	Dalian	632	405	26.1	20.7	-0.9	-7.7		1000+
	Japan	Tokyo	1465	481	30.8	24.2	9.8	2.1		1000+
Europe	Poland	Warsaw	520	203	23.6	12.9	0.4	-4.8		1000+
	Germany	Hamburg	773	224	22.1	12.7	3.5	-1.4		1000+
	France	Bordeaux	986	179	22.6	15.2	10.0	2.8		1000+
	Spain	Huelva	490	16	29.6	21.4	16.1	7.0		200–400
	Netherlands	Amsterdam	778	194	21.8	12.5	5.4	0.2		1000+
	Italy	Venice	810	154	27.5	17.8	5.8	-0.9		1000+
Pacific Rim	Australia	Coffs Harbor	1704	570	27.0	19.0	19.1	7.0		400–500
		Melbourne	665	154	24.8	14.5	13.4	5.9		800+
	New Zealand	Auckland	1135	246	25	14	16	7		800+
South America	Chile, north-central	Santiago	311	3	29.7	13.0	14.9	3.9		800+
	Chile, south-central	Osorno	1383	160	23.8	8.6	11.3	3.8		800+
	Argentina	Buenos Aires	1215	348	30.4	20.4	14.9	7.4		300–400

moderate chill conditions to regions with few to no hours under 7°C. In general, the cultivars that have done well in Florida and Georgia have also performed well in the hotter, drier production regions of California, central Chile, and Spain. However, only a few low chill cultivars have performed well in Mexico and Peru, and many of them suffer under the high UV levels of Chile.

During the last couple of decades, a constant stream of successful cultivars has been released from a number of breeding programs. These programs have focused on releasing cultivars with reduced chilling hours in warmer regions, increased cold hardiness in colder regions, and higher performance under high pH, temperature and radiative stress, but there is still much room for improvement. To achieve these goals, blueberry breeders have incorporated genes from many species within the *Vaccinium* genus through inter-specific hybridization (**Table 2**), which should prove to be a rich genetic pool for further improvements.

In this paper, we review the environmental challenges facing blueberry cultivation due to global warming. We describe the state of the art of blueberry breeding and outline how future varietal development can be enhanced by marker assisted breeding (MAB) and phenomics.

Environmental Challenges to Blueberry Cultivation

The expansion of highbush blueberry growing into colder and warmer regions will be challenged by the alterations in global temperature and rainfall patterns, both associated with increases in atmospheric CO_2 concentrations. From the "Industrial Revolution" carbon dioxide has increased in a significant way and will continue to do so. It is estimated that under the most conservative scenario, atmospheric CO_2 concentrations at the end of the century will be at least double to the pre-industrial era, increasing by 35% from year 2005 (IPCC, 2007b). As atmospheric concentrations of greenhouse gases rise due to the human activity, worldwide climatic patterns are being greatly altered (United Nations, 2010). The Intergovernmental Panel on Climate Change (IPCC), reports that during the past 150 years, global mean temperatures raised 0.045°C per decade, but in the last 25 years have increased almost four times (0.177°C) (IPCC, 2007a). Two separate analyses done recently by NASA (National Aeronautics and Space Administration) and NOAA (National Oceanic and Atmospheric Administration) have concluded that 2014 was the warmest year since 1880 (NASA, 2015). It is

TABLE 2 | Genetic composition of some of the cultivated blueberries.

Cultivar	Specie composition (%)							
	VC	VA	VD	Va	VT	Vc	VE	Others
Elliott, Brigitta, Liberty, Aurora, Lateblue, Jersey	100.0							
Duke	96.0	4.0						
Bluecrop	93.6	6.4						
Hannah's Choice	92.2	7.8						
Avonblue	86.7	0.8	5.0	7.5				
Lenoir	85.2	2.3	12.5					
Draper	84.5	6.0	1.6	1.2	0.4			6.3
O'Neal	84.0	10.0	3.0	3.0				
Misty	81.0	1.0	9.0	6.0	3.0			
Ozarkblue	77.3	3.9	11.3	7.5				
Summit	77.3	3.9	11.3	7.5				
Reveille	77.2	4.1	3.1	2.3	0.8			12.5
Sampson	76.6	10.9	12.5					
Magnolia	75.5	5.7	10.0	7.5	1.3			
Legacy	73.4	1.6	25.0					
Star	71.9	7.7	7.2	5.9	1.0			6.3
Camellia	71.8	1.6	19.7	3.8				3.1
Bluetta, Patriot, Sunrise	72.0	28.0						
Carteret	71.5	3.5					25.0	
Millennia	66.5	5.3	1.3	1.9				25.0
Jubilee	56.6	2.7	26.9	7.5				6.3
Emerald	54.4	1.9	13.9	1.5	0.2			28.1
Sierra	50.0	2.0	20.0	15.0		13.0		
Cara's Choice	47.7	2.3	20.0	15.0		15.0		
Sharpblue	43.7	28.8	15.0					12.5
Biloxi	41.8	1.8	32.5	11.3				12.6

VC, V. corymbosum; VA, V. angustifolium; VD, V. darrowii; Va, V. ashei; VT, V. tenellum; Vc, V. constablei; VE, V. elliottii (Hancock and Siefker, 1982; Ehlenfeldt, 1994; Clark et al., 1996; Hancock et al., 1997; Brevis et al., 2008; Lee et al., 2012; Rowland et al., 2013).

expected that during the next century global temperatures will be increased by an additional 1.1–6.4°C (Jin et al., 2011).

The increases in temperature are associated with extreme variations in weather patterns, resulting in severe droughts, unusually heavy rains and atypically hot temperatures (Allen and Ingram, 2002). Since the 1970s, the frequency of warm nights and days is increasing dramatically (IPCC, 2007a). For example, in the main blueberry production area in Chile (approximately between 35 and 38° Latitude S), precipitation diminished around 25% during the 20th Century, and it is estimated that there will be a further reduction of 5–15% over the next 30 years (Meza et al., 2003; Santibañez and Santibañez, 2007, 2008; United Nations, 2010). These dramatic changes led Friend (2010) to suggest that "Quantifying and explaining the current global distribution of plant production, and predicting its future responses to climate change and increasing atmospheric CO_2, are therefore major scientific objectives." High temperatures and drought can significantly reduce the productivity and the quality of the harvested organ (Moretti et al., 2010), restricting the areas (latitudes and soils) where economically important species can be grown.

The activity and development of humanity has not only increased atmospheric CO_2 levels but also levels of chlorofluorocarbons from aerosols, refrigerators, and other equipment that conditions the air. These compounds destroy the ozone layer, which selectively absorbs ultraviolet light. Ozone absorbs 100% of UV-C, prevents the passage of UV-B (near 90%) but does not affect the UV-A transmission (de Gruijl and van der Leun, 2000). In the southern (35–60°) and northern (35–60°) hemisphere, the annual mean ozone quantities during 2006–2009 were lower than between 1964 and 1980 (6 and 3.5%, respectively) (WMO, 2011).

The average UV erythemal irradiance, which indicates potential biological damage to human skin from solar ultraviolet radiation, has steadily risen as the amount of ozone has decreased (WMO, 2011). Compared with the 1970s, surface erythemal UV radiation has increased 7% in winter-spring and 4% in summer-fall in the northern hemisphere mid-latitudes, 6% year-round in the mid-southern hemisphere latitudes, and 22% in the Antarctic and Arctic in the spring (Madronich et al., 1998). In the summertime, erythemal UV irradiance in the southern hemisphere is up to 40% higher than values in the northern hemisphere (Madronich et al., 1998). If the Montreal Protocol is followed, it is possible that UV values will return to 1980 levels by the middle of this century, but this is dependent on multifaceted global cooperation (Kazantzidis et al., 2010; McKenzie et al., 2011).

Implications of Climate Change on Blueberry Breeding

The aspect of global warming that most needs attention from blueberry breeders is the dramatic seasonal fluctuations now occurring in rainfall and temperature patterns. Cultivars well adapted to "average conditions," often do not have sufficient plasticity to perform well under the range of conditions now being faced. For example, an unusually warm spring in Michigan in 2012 lead to very early floral development, and as a result, when temperatures returned to normal later in the spring, a high percentage of flowers were damaged by frost. An unusually hot summer in the Pacific Northwest in 2012, resulted in the fruit of most cultivars being too soft for extended storage. This was followed by an unusually cold winter in 2013–2014, where high percentages of the floral buds were heavily damaged. In Chile, falls and winters are becoming progressively milder in many areas, causing some cultivars to bloom out of season (O'Neal, Snowchaser and Misty, among others).

To maintain and extend the geographic range where blueberries are grown, breeders will need to be much more cognizant of the potential range of environments that the cultivars will face. They will need to take care not to release cultivars that are narrowly adapted to average conditions. Among the environmental challenges faced by blueberry breeders are:

Winter Cold

The range of the highbush blueberry has been limited by extreme winter cold. Cold hardiness is a complex interaction between rate of acclimation (development of freezing tolerance) and deacclimation (loss of developed freezing tolerance), as well as degree of mid-winter tolerance. This is extremely important since unseasonably warm midwinter spells can trigger a premature deacclimation, exposing the bush to freeze damage (Arora and Rowland, 2011).

In general, northern highbush cultivars survive much colder mid-winter temperatures than southern highbush ones, although considerable variability exists within groups and among *Vaccinium* species (Hancock et al., 1997; Ehlenfeldt et al., 2003, 2006, 2007; Dhanaraj et al., 2004; Rowland et al., 2004; Ehlenfeldt and Rowland, 2006; Hanson et al., 2007). In full dormancy, northern highbush genotypes have been found to range in tolerance from −20 to −30°C. Few southern highbush have been evaluated, although "Legacy" tolerates temperatures to −17°C and "Ozarkblue" to −26°C. US 245, an inter-species hybrid of US 75 ("Bluecrop" × *V. darrowii* "Fla 4B") × "Bluecrop," is tolerant to at least −24°C.

To date, the primary approach to developing more cold tolerant blueberries has been to hybridize lowbush with highbush to produce "half-high" types (Trehane, 2004; Hancock et al., 2008a). However, the shorter stature of the half-highs and the fact they become covered and protected with snow may be the primary basis of their increased tolerance (El-Shiekh et al., 1996). Due to the lack of formal comparisons of flower bud tolerance to winter cold in highbush, lowbush and half-highs, it is unknown how much more cold tolerant highbush can be made through introgression. It would be productive to determine if several other wild species carry useful genes for cold hardiness including *V. boreale*, *V. constablaei,* and *V. myrtilloides* (Galletta and Ballington, 1996; Ehlenfeldt and Rowland, 2006). Ehlenfeldt et al. (2007) showed that when *V. ashei* was hybridized with *V. constablaei*, cold hardiness was positively associated with the percentage of *V. constablaei* genes.

Little formal genetic analysis of cold tolerance of tetraploid blueberry has been performed. Arora et al. (2000) found

in diploid populations that the cold hardiness data fit a simple additive-dominance model of gene action, with the additive effects being greater than the dominance ones. During cold acclimation, specific genes are expressed in floral buds that increase cold tolerance (Naik et al., 2007). Arora et al. (1997a), working with "Bluecrop," "Tifblue," and "Gulfcoast," found a close relationship between floral bud dehydrin concentration and the level of cold hardiness. Similar results were found by Rowland et al. (2004) and Dhanaraj et al. (2005). This suggests that dehydrin concentration might be a way to predict the cold hardiness of selections in a breeding program.

It is important to note that studies of cold hardiness under field or artificial conditions can lead to different conclusions. When "Bluecrop" (NHB) and "Tifblue" (Rabbiteye blueberry—RE) flower buds were assessed in the field, LT_{50} (maximum level of cold-hardiness) were close to −27 and -25°C (respectively), whereas the same cultivars in cold room conditions (4°C) reached maximums around −24 and −17°C, respectively (Arora et al., 1997b; Arora and Rowland, 2011). There were almost twice as many "Bluecrop" genes expressed in the cold room than in the field, suggesting that many of the genes induced in the cold room were responding to low temperature (specifically 4°C) and were not contributing to freezing tolerance *per se*. In contrast, more "Tifblue" genes were expressed in the field than under the controlled conditions. This suggests that there is a strong genotype × environment interaction associated with cold tolerance and any screen designed to select cold hardy genotypes, must be conducted under field conditions or under realistic controlled protocols (Arora and Rowland, 2011).

It may be possible to determine when a plant is approaching full dormancy by measuring the expression of the β-amylase gene. Lee et al. (2012) showed in the NHB "Jersey" and the SHB "Sharpblue" that there was an abrupt reduction in starch in shoots in the middle of cold acclimation, which was associated with an increase in the expression of the β-amylase gene. This change was positively correlated with the total amount of soluble solids in the wood, which likely served as osmoprotectants able to reduce the freezing point. Inter-species differences in the level of expression of β-amylase genes in northern and southern highbush were described by Rowland et al. (2008).

Spring and Fall Frost

Freezing damage to developing flowers in the spring is a major problem in most blueberry production regions, with both NHB and SHB. It is a rare year when at least a fraction of the flower buds is not damaged. Rate of deacclimation likely plays a role in early spring flower bud tolerance. Ehlenfeldt (2003) found the northern highbush "Duke" deacclimated the fastest in a mixed group of 12 cultivars, while the southern highbush "Magnolia," the northern highbush × rabbiteye pentaploid hybrid "Pearl River," the rabbiteye × *V. constablaei* "Little Giant" and the half-highs "Northcountry" and "Northsky" were the slowest. Northern highbush "Bluecrop" and "Weymouth," southern highbush "Legacy" and "Ozarkblue" were intermediate. While there is evidence of considerable variability, no formal genetic studies have been done on deacclimation rates.

Identifying late bloom or slower deacclimating genotypes will be useful for breeding spring-frost tolerant cultivars (Rowland et al., 2005). Because of the chances of frosts and the direct relation between the stage of floral development and the relative bud hardiness, those cultivars with late bloom dates tend to suffer less frost damage than those flowering earlier (Spiers, 1976; Hancock et al., 1987; Patten et al., 1991; Lin and Pliszka, 2003). When Hancock et al. (1987) assessed flower bud injury in 18 highbush blueberry cultivars after two spring frosts in Michigan, they found significant differences in proportion of brown ovaries among cultivars, ranging from 25 to 94%. Most of the variation was associated with stage of bud development.

Bloom date is strongly correlated with ripening date, but early ripening cultivars have been developed that have later than average flowering dates such as the NHB "Duke," "Huron" and "Spartan," and the SHB "Santa Fe" and "Star." Bloom date, ripening interval and harvest dates are highly heritable in blueberry populations (Lyrene, 1985; Hancock et al., 1991) with strong genotype by environmental interactions (Finn et al., 2003). Finn and Luby (1986) found additive genetic variation was more important than non-additive effects for date of 50% bloom, 50% ripe fruit and for length of fruit development interval in populations from hybrids between *V. angustifolium* and *V. corymbosum*. Where spring frosts are a problem, breeders can focus on developing cultivars with late bloom dates and where earliness is premium, selection will need to be made on ripening interval as well.

Flower buds also can be damaged by rapid freezes in the fall. The flower buds of SHB cultivars are generally considered to acclimate more slowly in the fall than those of NHB ones, and as a result are more subject to late fall freezes; however, few formal screens of germplasm have been conducted on this characteristic (Rowland et al., 2005, 2013; Hanson et al., 2007). Leaf retention in the fall does not appear to be a good predictor of rate of DA, as Hanson et al. (2007) found that "Ozarkblue" and US 245 retain their leaves until the very late fall, but they are just as hardy as the mid-season standard "Bluecrop." Bittenbender and Howell (1975) also found no correlation between flower bud hardiness and fall leaf retention.

There are not many studies that have evaluated the effect of the spring frost on open flowers, and most of them have been done on RE (Spiers, 1976; Gupton, 1983; NeSmith et al., 1999). Among RE cultivars, "Southland" proved to be more frost-tolerant than "Delite," "Woodard," "Climax," and "Tifblue" (Gupton, 1983). Nevertheless, interspecific crosses with RE would not be recommended to increase bud tolerance to frost since, at similar stages of floral bud development, RE tend to be more sensitive than NHB and SHB (Patten et al., 1991). Rowland et al. (2013), studying the sensitiveness of five northern highbush blueberries cultivars ("Bluecrop," "Elliott," "Hannah's Choice," "Murphy," and "Weymouth") to frost damage of open flowers, concluded that "Hannah's Choice" and "Murphy" were the most tolerant whereas "Bluecrop" was the most susceptible. Among the cultivars analyzed by Rowland et al. (2013), female parts from "Elliott" (styles), "Hannah's Choice" (styles and exterior ovaries) and "Murphy" (styles) were more frost-tolerant than those structures of "Bluecrop," and the male organs from

"Murphy" (filaments and anthers) were more frost-tolerant than "Bluecrop." These differences need to be studied in other cultivars and exploited for breeding.

Chilling Requirement

Expanding the range of adaptation of the NHB by reducing its chilling requirement has been a major breeding goal over the last 50 years (Hancock et al., 2008a). This was largely accomplished by incorporating genes from the southern diploid species *V. darrowii* into *V. corymbosum* via unreduced gametes, although hybridizations with native southern *V. corymbosum* and *V. ashei* also played a role. Cultivars with an almost a continuous range of chilling requirements (hours below 7°C) are now available from 0 to 1000 h.

Most SHB are grown in areas with 250–600 chilling hours each winter. SHB cultivars vary widely in their performance without any chilling hours. Surprisingly, one of the best adapted cultivars to this system is "Biloxi," which requires 500 chilling hours (<7°C) in Mississippi, where it was developed. The response to chilling is clearly a complex interaction and many factors play a role including sensitivity to temperature shifts, floral development time, response to photoperiodic change and temperature thresholds.

The genetics of the chilling requirement has not been formally determined, although segregation patterns suggest that it is largely quantitatively inherited with the low chilling requirement showing some dominance. The precise temperature necessary to break dormancy has not been determined, but Mainland et al. (1977) and Spiers (1976) have proposed that the chilling requirement of highbush blueberries is at least partially satisfied by temperatures below 1.4 and above 12.4°C. It is possible that blueberry genotypes vary in the threshold temperatures that are required to break dormancy, although this has not been documented. Southern highbush cultivars with complex ancestry may be particularly variable in their temperature thresholds.

Heat and Drought Tolerance

When heat stress is present in blueberries, the quick response needed to supply the atmospheric demand, puts the plant and its fruit at permanent threat (Chen et al., 2012). High summer temperatures, such as in the subtropical southeast China or the dry Mediterranean north of Chile, impact on the productivity of highbush blueberries across much of their range (Darnell, 2000; Chen et al., 2012). It is thought that SHB are more tolerant of high temperatures than NHB, but both types commonly experience summer temperatures in the field that have negative impacts on CO_2 assimilation rates and fruit quality. In general, optimal temperatures have been shown to vary between 20 and 25°C (Davies and Flore, 1986).

No formal studies have been conducted on the genetics of photosynthetic heat tolerance in blueberry, but genetic variation has been documented. Moon et al. (1987b) evaluated the optimum temperature for photosynthesis in different highbush cultivars, determining ranges of 18–6°C for Jersey and 14–22°C for Bluecrop. A temperature of 30°C has been shown to reduce photosynthesis in NHB cultivars by 22–51% (Hancock et al., 1992); the authors reported that "Jersey," "Elliott," and "Rubel"

showed a decrease in photosynthesis between 22 and 27% whereas for "Spartan," "Bluejay," and "Patriot" it was between 41 and 51%. Trehane (2004) describes "Ozarkblue" and "Jubilee" as varieties that perform well in hot summers. Chen et al. (2012) found that at high temperatures, up to 40–45°C, a number of photosynthetic parameters were damaged in "Brigitta," but they stayed largely intact in "Sharpblue," and "Duke"; "Misty" performed in the middle. In this study, at high temperatures, there were increases in hydrogen peroxide, super oxide radical and F_0 (minimum fluorescence in the dark-adapted state), while Fv/Fm (maximum photochemical quantum yield of photosystem II) and ÔPS II (quantum efficiency of PSII photochemistry) decreased.

Southern highbush cultivars may have obtained higher photosynthetic heat tolerance from *Vaccinium darrowii* (Lyrene, 2002). Moon et al. (1987a) found that CO_2 assimilation (A) in Fla 4B of *V. darrowii* was similar at 20 and 30°C, while A in the pure northern highbush "Bluecrop" dropped by almost 30% across this same range. Transpiration rates were also much lower in Fla 4B than "Bluecrop." This difference was found to be heritable, with a tetraploid F_1 hybrid actually having higher A than the two parents (Moon et al., 1987b; Hancock et al., 1992). The selection, Fla 4B has been used to generate many of the important southern highbush cultivars including "Biloxi," "Emerald," "Legacy," and "Star" (Lyrene and Sherman, 2000; Draper and Hancock, 2003). There may also be additional sources of heat tolerance in the native southern species *V. tenellum*, *V. myrsinites*, *V. pallidum*, *V. ashei*, *V. elliottii*, *V. stamineum*, *Vaccinium arboreum*, southern diploids and tetraploids of *V. corymbosum* (Luby et al., 1991).

It would seem likely that the photosynthetic heat tolerance of both NHB and SHB types can be increased by crossing the most heat tolerant genotypes, since there is considerable genetic variability for this trait both within and among blueberry species.

High temperatures also negatively impact fruit quality and storage life of highbush blueberries. Temperatures higher than 32°C during the maturation of the fruit can give rise to smaller, soft fruits and with waxes that have greater susceptibility of being lost by means of the rubbing (by leaves or during the harvest) (Mainland, 1989).

Blueberries have a relatively inefficient water conducting systems, characterized by the lack of root hairs (Gough, 1994). Root anatomy and architecture should be a key trait, but unfortunately it is almost unexplored, e.g., *V. arboreum* is drought tolerant specie because it has deep tap roots in contrast to the spreading, shallow root systems of highbush blueberry. Hence, drought tolerance of highbush blueberries might also be enhanced by using species material in breeding.

In his screens of wild species material, Erb et al. (1988a,b) found *V. elliottii*, *V. darrowii*, and *V. ashei* to be the most drought tolerant species and this characteristic was transmitted to hybrid progeny. Moon et al. (1987a) found transpiration rates and leaf conductance (gs) to water vapor to be much lower in *V. darrowii* than "Bluecrop" at high temperature. This suggests that *V. darrowii* may have higher drought tolerance through decreased stomatal opening and subsequent restriction of water loss.

Other sources of drought tolerance likely include the native species *Vaccinium stamineum* and *V. arboretum* (Hancock et al., 2008a). *V. stamineum* is the most drought tolerant species in the southeastern U.S.A., but hybrids derived with species in section Cyanococcus have not been vigorous (Ballington, 1980; Lyrene, 2006). The use of *V. arboreum* appears to be more promising, as this species can be crossed with *V. darrowii* to produce vigorous hybrids, and these hybrids can be used as a bridge to tetraploid SHB types (Lyrene, 1991; Brooks and Lyrene, 1998; Olmstead et al., 2013).

In response to water deficit, plants stop shoot growth affecting their final height and diameter (Mingeau et al., 2001). Bluecrop is one of the cultivars which has been studied most, proving to be highly sensitive to water deficit, showing a rapid stomatal closure, and reduced gas exchange (Cameron et al., 1989; Rho et al., 2012), berry size and yields (Améglio et al., 2000). When "Bluecrop" was subjected to severe hydric restriction, a reduction in the yield (31–49%) was observed (Perrier et al., 2000). Similar results were found for other highbush cultivars like Rancocas (Lee et al., 2006) and Jersey (Cameron et al., 1989). Rho et al. (2012) also found that along with the reduction in gas exchange found in "Bluecrop" under water deficit (-1.9 MPa), an increment in the electron transport rate (ETR) occurs, indicating photorespiration is also affected.

When Estrada et al. (2015) studied how SHB, RE, and NHB responded to drought conditions, with or without heat stress, they found that under each stress SHB and RE had a better photoprotection capacity, while the NHB showed increments in its photochemical capacity. When both stresses were present, just NHB "Liberty" and "Elliott" had increased ETR_{max} (maximum electron transport rate), the coefficient of photochemical fluorescence quenching (qP and qL) and the effective photochemical quantum yield of photosystem II [Y(II)]. This suggests these two cultivars should be considered as parents if reduction of photo-oxidative damage is required. This tool could be used to screen genotypes much faster than other classic measurements, such as gas exchange rate, chlorophyll content, stem water potential, etc., making the monitoring of stressed plants more efficient (Ralph and Gademann, 2005; Estrada et al., 2015).

The priority of this trait for the breeder varies depending on location and irrigation availability. It should be mentioned that currently most blueberries are cultivated under irrigation or with irrigation supplementation.

High UV Light
Ozone depletion itself, is not a major contributor to global warming, but increases in UV irradiance have large, direct impacts on plant productivity (Boesgaard et al., 2012). In some latitudes, plants will not only have to deal with extreme photosynthetic active radiation and heat, but also with high UV radiation. Is not unusual to observe damaged plants in commercial fields (Yáñez et al., 2009), as well as among breeding families, in central Chile with reddened and curled leaves and what appear to be localized burns on fruit. Shading nets have been shown to enhance productivity of blueberries in central Chile, but the direct influence of UV light has not be

investigated (Retamales et al., 2008; Lobos et al., 2009, 2012, 2013, 2015).

Kakani et al. (2003), reviewing 129 studies of the effect of UV-B on 35 crops, reported that higher levels of UV-B (most affected by the ozone depletion) were associated with vegetative and reproductive morphology alterations, decreases in chlorophyll content and photosynthesis, and chlorotic or necrotic patches on leaves or fruit. Little formal work has been done on the effect of UV light on northern highbush blueberries, and most of what has been done has focused on postharvest improvement through short treatments of UV-B on harvested fruit (Perkins-Veazie et al., 2008; Eichholz et al., 2011).

The damaging effects of high UV light have been documented in other *Vaccinium* species. Albert et al. (2008) and Boesgaard et al. (2012) found a reduction of net photosynthesis in *V. uliginosum* throughout the season and damage to photosystem II (PS II) through the diminution of the fluorescence measured as Fv/Fm. Kossuth and Biggs (1978) tested the effects of 15, 24, and 44 units of UV-B on the rabbiteye blueberry "Woodard" and found that the higher doses reduced fruit growth and surface bloom, and under high doses of UV-B the fruit skin actually appeared burned.

Among the mechanisms that might be selected to improve UV tolerance in blueberries is the ability of the leaf surface to reflect part of the incident radiation (Semerdjieva et al., 2003a,b). The thickness of the epidermis of the leaf and the concentration of absorbent compounds could also be improved to counteract the damaging effects of UV radiation (Batschauer, 1998; Boesgaard et al., 2012). Increases in levels of phenols (flavonoids and hidroxamic acid), could also help counter the degradative effects of high UV-B on DNA (Rozema et al., 1997; Ruhland et al., 2005). Defense responses to UV-B have not been evaluated in highbush blueberry; however, Semerdjieva et al. (2003a,b) noted in three other species (*V. myrtillus* L., *V. vitis-idaea* L., and *V. uliginosum* L.) in the north of Sweden, that there were noticeable differences in quantity of phenolic compounds. In all species, high UV-B led to an increase in phenolic compounds, but some genotypes responded more than others, and the plants with the highest flavonoid content had the least UV-B damage.

New Frontiers for the Breeding of Blueberries

The breeding of highbush blueberries can be a long and tedious process. Traditional approaches take from 10 to 20 years from the original cross to cultivar release and often the precise adaptive range of a cultivar is not known until farmers have grown it for a number of years. Two relatively new techniques called "MAB" and "Phenomics" could greatly facilitate blueberry breeding. MAB would aid in the selection of those individuals most likely to carry adventitious traits and Phenomics would allow for much more easy, fast and precise characterization of the superior types. It is possible that individuals could be selected for their adaptability to variable environmental conditions with MAB even though they were not exposed to those conditions in the field.

Marker Assisted Breeding (MAB)

All breeding programs revolve around identifying the optimal traits for a cultivar. Most blueberry breeding programs utilize traditional approaches to identify desirable types, such as walking along rows of crosses in the field or doing simple laboratory assays on fruit quality and disease resistance. However, in many other crops, MAB is used to facilitate and speed up the release of new cultivars (Cabrera-Bosquet et al., 2012; Araus and Cairns, 2014).

MAB is based on DNA diagnostic tests that can identify potential parents and progeny carrying desirable traits. This process allows selection to be moved all the way back to conception in the breeder's minds, helping them to only make crosses that create desirable trait combinations in offspring, increasing the efficiency of the entire process. It also permits selection to be moved from the field to the greenhouse, so that only seedlings predicted to be superior are planted in the field for further evaluation. In addition, MAB allows for the assessment of traits that are difficult to predict in the field such as chilling hour requirement or heat tolerance.

To use MAB to broaden the environmental range where highbush blueberries can be grown, it will be necessary to find genetic variability associated with expanded adaptations. The rich germplasm diversity currently being used by blueberry breeders (**Table 2**) is likely to contain useful genes. The major stumbling block to using MAB will be the collection of precise data on the adaptations of potential breeding parents. Evaluations of genotypes in the field will require that the extreme conditions occur when the plants are in the field (Arora and Rowland, 2011). Care will need to be taken to evaluate genotypes in appropriate environments and in many cases controlled experiments will need to be undertaken. Most likely a genotype will need to be evaluated in multiple environments to obtain an accurate representation of its adaptations. This will often be tedious and time consuming, but once markers are found that are tightly linked to the genes regulating adaptations of interest, future screening will be greatly facilitated through MAB. As we discuss below, field screening could be greatly streamlined using phenomic techniques involving spectrometry and thermography.

The first genetic maps of blueberries are beginning to emerge that will set the groundwork for MAB. Rowland's group at the USDA-ARS (Genetics of Fruit and Vegetable Improvement Laboratory, Beltsville, MD, USA) developed the first blueberry map using a diploid population segregating for chilling requirement (Rowland and Levi, 1994). Their population was a cross between an inter-specific hybrid (*V. darrowii* × *V. corymbosum*) and another clone of *V. corymbosum*. They have continued to periodically add markers to this map and at last report had 265 markers mapped to 12 linkage groups. They have used this map to identify quantitative trait loci for cold tolerance and chilling requirement (Rowland et al., 2014). Allan Brown (North Carolina State University, NC, USA) and Eric Jackson (General Mills, Crop Bioscience, MN, USA) have led teams that sequenced the genome of a *V. darrowii* × *V. corymbosum* hybrid and used this information to generate a more dense chromosome map with 1200 markers.

A major research initiative has been undertaken by Lisa J. Rowland, Nahla Bassil (USDA-ARS, National Clonal Repository, Corvallis, OR, USA), Julie Graham and Susan McCallum (The James Hutton Institute, Dundee, UK) and Jim Olmstead (University of Florida, Gainesville, FL, USA) to develop a linkage map of the tetraploid cross "Jewel" (SHB) × "Draper" (NHB) (Rowland et al., 2012). Replicated progeny of that cross were planted at five locations across the USA and data was collected on a wide array of traits including fruit quality, developmental rates, chilling hour requirements and growth patterns. A QTL analysis is currently being conducted to search for markers for these traits.

The first few thousand expressed sequence tags (ESTs) have been generated and made publicly available for the Ericaceae family, about 5000 from blueberry and about 1200 from rhododendron (Rowland et al., 2010, 2014; Die and Rowland, 2013) (http://bioinformatics.towson.edu/BBGD/). These ESTs from blueberry and rhododendron were generated as parts of projects focused on cold acclimation research, and the ESTs are from non-acclimated and cold acclimated flower bud libraries, in the case of blueberry (Dhanaraj et al., 2004, 2007), and from non-acclimated and cold acclimated leaf libraries, in the case of rhododendron (Wei et al., 2005). Another ~16,000 ESTs have been generated from blueberry fruit by the New Zealand Institute for Plant and Food Research Ltd. (formerly HortResearch), but they are not publicly available.

Phenomics

The success of breeding programs is reflected in the number of individuals released at the end of the selection process (Hancock et al., 2008b). In order to be successful, breeders must generate thousands of hybrids annually and evaluate them for a number of years. What ultimately is selected is dependent on local environmental conditions (Araus and Cairns, 2014).

Because of the large number of genotypes that need to be evaluated, deep phenotypic characterizations of the material often becomes impractical due to the time and costs involved (White et al., 2012; Kipp et al., 2014). For this reason, conventional breeding generally focuses on different visual characteristics (e.g., fruit color, cluster tightness, disease resistance, growth habit, flowering, and ripeness dates) and a few that require measurements of average complexity (e.g., yield, soluble solids, firmness).

To effectively develop cultivars well adapted to fluctuations in environmental stresses, blueberry breeders will have to evaluate a number of morpho-physiological and physico-chemical traits that they are not used to considering (**Figure 1**). The only reasonable way to fulfill all these needs is through acquisition of high-dimensional phenotypic data (high-throughput field phenotyping) or "Phenomics" (Houle et al., 2010). Nowadays, there are a number of remote sensing devices, techniques and analysis, mostly non-destructive, which have proven very helpful in the characterization of the phenotype (Furbank and Tester, 2011; White et al., 2012; Araus and Cairns, 2014).

Among the remote sensing technologies with the greatest potential for use in phenomics are spectrometry and thermography. Spectroradiometers are widely used to measure plant reflectance (R), whose spectral signature (graphic
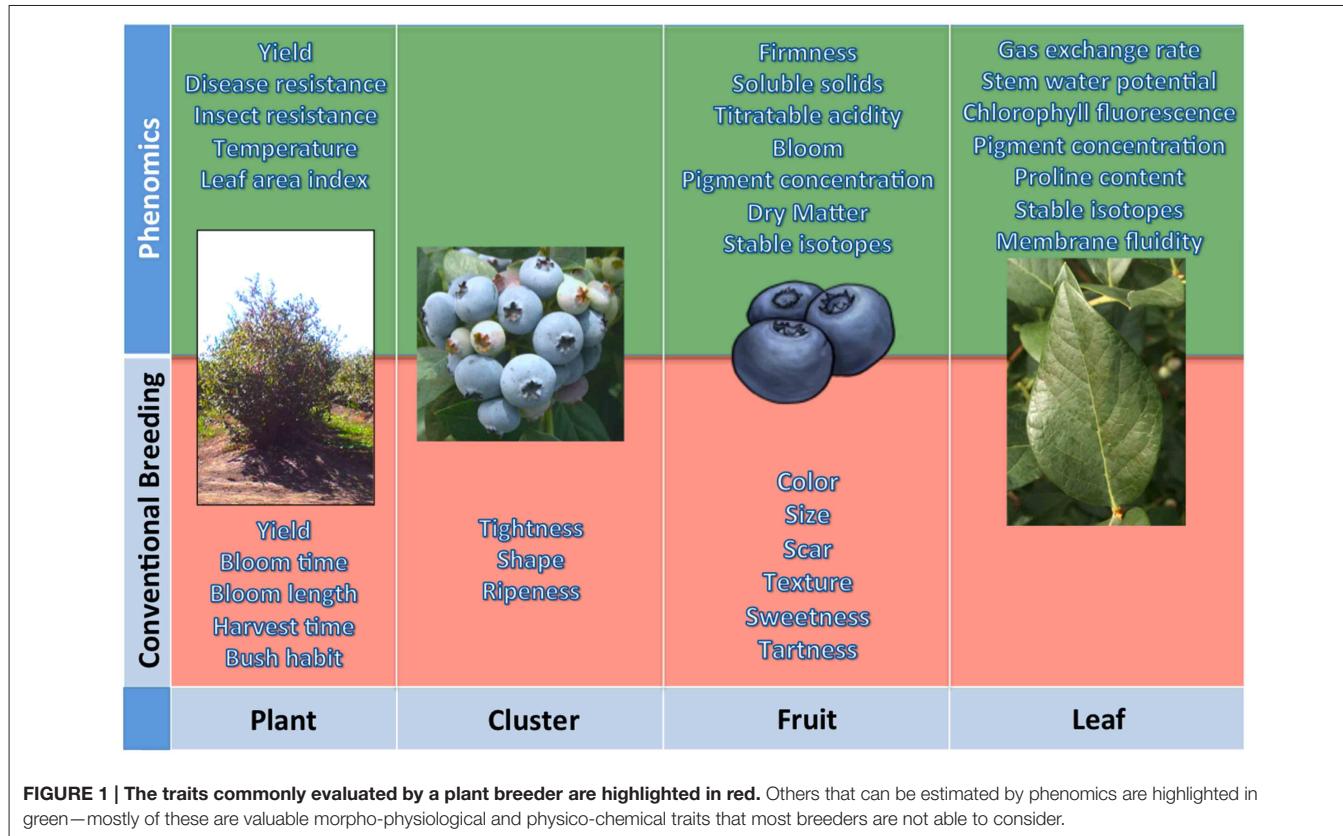
FIGURE 1 | The traits commonly evaluated by a plant breeder are highlighted in red. Others that can be estimated by phenomics are highlighted in green—mostly of these are valuable morpho-physiological and physico-chemical traits that most breeders are not able to consider.

characterization of reflected segment of wavelengths) is closely associated with the absorption at certain wavelengths that are linked to specific characters or plant conditions (Araus and Cairns, 2014). Thermography uses plant temperature as an efficient tool for the study of the spatial and temporal heterogeneity of plant water status and how it responds to the environment.

While these techniques are not new, their use was expanded tremendously during mid 1980s and is now being widely used in plant ecophysiology and postharvest studies (estimation of yield, nutritional content in leaves, gas exchange rate, fruit quality, biotic and abiotic stress, etc.) (Garriga et al., 2014; Lobos et al., 2014). Measurements that before usually took months, weeks, or days can now be accomplished in hours or even minutes for a large number of genotypes (**Figure 2**).

In spectrometry, reflectance data is used to generate "Spectral Reflectance indices" (SRI). Initially SRIs were simple relationships between wavelengths or spectral bands. The first SRI was the "Simple Ratio," calculated as the ratio of the near infrared (NIR) to the visible (VIS) (SR = R_{NIR}/R_{VIS}), and the "Normalized Difference Vegetation Index" [NDVI = (R_{NIR}-R_{VIS})/(R_{NIR}+R_{VIS})]. Since then, and incorporating specific wavelengths, SRIs have been used in different species to estimate green biomass and leaf area index (Tucker and Sellers, 1986), plant water status (Peñuelas et al., 1993), radiation use efficiency (Peñuelas et al., 1995), water content in leaves (Sims and Gamon, 2003), photosynthetic capacity and efficiency (Inoue et al., 2008),

micro and macro nutrients in leaves (Basayigit and Senol, 2009), yield and carbon isotope discrimination (Lobos et al., 2014) among many others.

Since the 1960s, plant temperature has been widely used as an indicator of water status (Tanner, 1963). Initially, temperature measurements were performed by thermocouples in contact with the leaves. Later, the development of infrared sensors allowed faster measurements of leaves or canopies. With thermal imaging it is possible to detect biotic or abiotic pre-symptomatic responses, providing a powerful tool to evaluate a high number of samples in only a few minutes (Costa et al., 2013; Araus and Cairns, 2014). The development of cheaper devices has made this approach available to farmers and breeders. Thermometry analysis has been fine-tuned over time and the different parts of the image (soil, air, leaves, stems, branches, etc.) can now be isolated, allowing for the evaluation of specific tissues, organs or individuals (Costa et al., 2013; Araus and Cairns, 2014).

To date, spectrometry and thermography have been used to only a limited extent on blueberries. There are a few studies where the antioxidant content in NHB was evaluated using spectrometry (total phenols, total flavonoids, total anthocyanins, and ascorbic acid) (Sinelli et al., 2008; Bai et al., 2014). The ideal "Brigitta" harvest date was determined in the field using reflectance data (Beghi et al., 2009), and a blueberry ripeness index (BRI) has been developed (Beghi et al., 2013). Guidetti et al. (2009) used a portable spectroradiometer Vis-NIR to accurately estimate soluble solids, firmness and functional compounds

FIGURE 2 | When drought or heat stress needs to be assessed, the evaluation window per day is reduced to 4 h. A basic physiological plant evaluation (gas exchange rate, stem water potential, chlorophyll fluorescence, and pigment concentration) **(A)** takes about 16 min per plant, representing a characterization of a maximum 15 genotypes per day. When thermography **(B)** and spectrometry **(C)** is considered, linear and non-linear modeling **(D)** can streamline the evaluation of plant status, further increasing the number of genotypes that can be evaluated in a breeding population per day **(E)**.

(anthocyanins, flavonoids, polyphenols, and ascorbic acid) in fresh and homogenized fruit samples of "Brigitta" and "Duke." Other work on blueberries focused on monitoring osmo and air dehydration processes (Sinelli et al., 2011), SHB cultivar identification (Yang and Lee, 2011; Yang et al., 2012), and the recognition of foreign materials (leaves and stems) among frozen blueberries (Tsuta et al., 2006; Sugiyama et al., 2010). In the wild lowbush blueberry (*V. angustifolium*) reflectance data has also been used for detection of internal larvae fruit infestation (Peshlov et al., 2009), *in situ* levels of foliar nitrogen (Bourguignon, 2006; Maqbool et al., 2012) and to evaluate vegetative (leaf area index) and reproductive (flower number, fruit set, and berry yield) parameters (Percival et al., 2012). Most recently, hyperspectral imaging has been used to predict soluble solids content and firmness in NHB fruit (Leiva-Valenzuela et al., 2013, 2014), to identify damaged fruit (Leiva-Valenzuela et al., 2012; Leiva-Valenzuela and Aguilera, 2013), to classify blueberry fruit growth stages (Yang et al., 2012, 2014), and as a tool for early detection of leaf rust in blueberries (Ahlawat et al., 2011). Escobar-Opazo (2015) found that in blueberry some of the physiological parameters were significantly correlated with reflectance data (e.g., ETR_{max} and chlorophyll a/b > 0.90; A and gs > 0.65).

During the last decade, spectrometry and thermography have begun to be used in the breeding of grain crops, but their use in fruit breeding programs is almost non-existent. Even in the grain crops, their use has been limited to the evaluation of small numbers of genotypes, in general <20. SRI will need to be replaced by more complex bio-mathematical models, to fully provide breeders with the solid and reliable information

they need for plant selection. When this happens, the efficiency of selection should dramatically improve, and well adapted genotypes will be released at a faster rate. The future challenge will be to develop techniques that can screen a large number of genotypes simultaneously (hundreds or thousands) (**Figure 2**).

Concluding Remarks and Future Perspectives

Considerable genetic variability exists in the highbush blueberry germplasm base that can be used by breeders to meet the environmental challenges associated with climate change. There has already been extensive range expansion of blueberries into hotter and drier environments. There is no reason to believe that there is not additional genetic variability that can be deployed to further enhance cold acclimation, heat, high UV, and drought tolerance of blueberries. Perhaps the greatest challenge associated with climate change and blueberry range expansion will be the development of blueberry cultivars that can resist extremes in environmental variability. Ongoing research to develop DNA diagnostic markers for key physiological tolerances will aid greatly in the breeding of these stress resistant types. To generate robust markers for MAB, it will be necessary to have precise phenotypic characterizations, making phenomics a powerful tool that could aid greatly in identifying the superior types.

Acknowledgments

References

Ahlawat, V., Jhorar, O., Kumar, L., and Backhouse, D. (2011). "Using hyperspectral remote sensing as a tool for early detection of leaf rust in blueberries," in *34th International Symposium on Remote Sensing of Environment - The GEOSS Era: Towards Operational Environmental Monitoring* (Sydney).

Albert, K. R., Mikkelsen, T. N., and Ro-Poulsen, H. (2008). Ambient UV-B radiation decreases photosynthesis in high arctic *Vaccinium uliginosum*. *Physiol. Plant* 133, 199–210. doi: 10.1111/j.1399-3054.2008.01065.x

Allen, M. R., and Ingram, W. J. (2002). Constraints on future changes in climate and the hydrologic cycle. *Nature* 419, 224–232. doi: 10.1038/nature01092

Améglio, T., Roux, X. L., Mingeau, M., and Perrier, C. (2000). Water relations of highbush blueberry under drought conditions. *Acta Hortic.* 537, 273–278. doi: 10.17660/actahortic.2000.537.30

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Arora, R., and Rowland, L. J. (2011). Physiological research on winter-hardiness: deacclimation resistance, reacclimation ability, photoprotection strategies, and a cold acclimation protocol design. *HortScience* 46, 1070–1078.

Arora, R., Rowland, L. J., and Panta, G. R. (1997a). Chill-responsive dehydrins in blueberry: are they associated with cold hardiness or dormancy transitions? *Physiol. Plant.* 101, 8–16. doi: 10.1111/j.1399-3054.1997.tb01813.x

Arora, R., Rowland, L. J., Lehman, J. S., Lim, C. C., Panta, G. R., and Vorsa, N. (2000). Genetic analysis of freezing tolerance in blueberry (Vaccinium section Cyanococcus). *Theor. Appl. Genet.* 100, 690–696. doi: 10.1007/s001220051341

Arora, R., Wisniewski, M. E., and Rowland, L. J. (1997b). Low temperature-induced expression of dehydrins in deciduous fruit crops and their relation to cold acclimation and/or dormancy. *Acta Hortic.* 441, 175–182. doi: 10.17660/actahortic.1997.441.23

Bai, W., Yoshimura, N., and Takayanagi, M. (2014). Quantitative analysis of ingredients of blueberry fruits by near infrared spectroscopy. *J. Near Infrared Spectrosc.* 22, 357–365. doi: 10.1255/jnirs.1129

Ballington, J. R. (1980). Crossability between subgenus Cyanococcus (Gray) Klotzsch and subgenus Polycodium (Raf.) Sleumer in Vaccinium. *HortScience* 15, 419.

Basayigit, L., and Senol, H. (2009). Prediction of plant nutrient contents in deciduous orchards fruits using spectroradiometer. *Int. J. ChemTech. Res.* 1, 212–224.

Batschauer, A. (1998). Photoreceptors in higher plants. *Planta* 206, 479–492. doi: 10.1007/s004250050425

Beghi, R., Giovenzana, V., Spinardi, A., Guidetti, R., Bodria, L., and Oberti, R. (2013). Derivation of a blueberry ripeness index with a view to a low-cost, handheld optical sensing device for supporting harvest decisions. *Trans. ASABE* 56, 1551–1559. doi: 10.13031/trans.56.10169

Beghi, R., Guidetti, R., Oberti, R., and Bodria, L. (2009). "Feasibility of a simplified handheld optical system for blueberries ripeness field evaluation," in *5th International Technical Symposium on Food Processing, Monitoring Technology in Bioprocesses and Food Quality Management,* eds M. Zude and O. Schlüter (Potsdam).

Bittenbender, H., and Howell, G. (1975). Interactions of temperature and moisture content on spring de-acclimation of flower buds of highbush blueberry. *Can. J. Plant Sci.* 55, 447–452. doi: 10.4141/cjps75-070

Boesgaard, K. S., Albert, K. R., Ro-Poulsen, H., Michelsen, A., Mikkelsen, T. N., and Schmidt, N. M. (2012). Long-term structural canopy changes sustain net photosynthesis per ground area in high arctic *Vaccinium uliginosum* exposed to changes in near-ambient UV-B levels. *Physiol. Plant.* 145, 540–550. doi: 10.1111/j.1399-3054.2011.01564.x

Bourguignon, C. (2006). *The Feasibility of Determining Foliar Nitrogen, Phosphorus and Potassium Contents using Existing Hyperspectral Technologies in Wild Blueberry*. Halifax, NS: Dalhousie University.

Brevis, P. A., Bassil, N. V., Ballington, J. R., and Hancock, J. F. (2008). Impact of wide hybridization on highbush blueberry breeding. *J. Am. Soc. Hortic. Sci.* 133, 427–437.

Brooks, S. J., and Lyrene, P. M. (1998). Level of self- and cross-fertility of derivatives of Vaccinium arboreum x Vaccinium section Cyanococcus hybrids. *HortScience* 33, 1066–1068.

Cabrera-Bosquet, L., Crossa, J., von Zitzewitz, J., Serret, M. D., and Luis Araus, J. (2012). High-throughput Phenotyping and genomic selection: the frontiers of crop breeding converge. *J. Integr. Plant Biol.* 54, 312–320. doi: 10.1111/j.1744-7909.2012.01116.x

Cameron, J. S., Brun, C. A., and Hartley, C. A. (1989). The influence of soil moisture stress on the growth and gas exchange characteristics of young highbush blueberry plants (*Vaccinium corymbosum* L.). *Acta Hortic.* 241, 254–259. doi: 10.17660/actahortic.1989.241.41

Chen, W., Cen, W., Chen, L., Di, L., Li, Y., and Guo, W. (2012). Differential sensitivity of four highbush blueberry (*Vaccinium corymbosum* L.) cultivars to heat stress. *Pak. J. Bot.* 44, 853–860.

Clark, J. R., Moore, J. N., and Draper, A. D. (1996). 'Ozarkblue' southern highbush blueberry. *HortScience* 31, 1043–1045.

Costa, J. M., Grant, O. M., and Chaves, M. M. (2013). Thermography to explore plant–environment interactions. *J. Exp. Bot.* 64, 3937–3949. doi: 10.1093/jxb/ert029

Darnell, R. L. (2000). "Blueberries," in *Temperate Fruit Crops in Warm Climates,* ed A. Erez (Dordrecht; London: Kluwer Academic Publishers), 429–444.

Davies, F. S., and Flore, J. A. (1986). Gas exchange and flooding stress of highbush and rabbiteye blueberries. *J. Am. Soc. Hortic. Sci.* 111, 565–571.

de Gruijl, F. R., and van der Leun, J. (2000). Environmental and health: 3. Ozone depletion and ultraviolet radiation. *Can. Med. Assoc. J.* 163, 851–855.

Dhanaraj, A. L., Alkharouf, N. W., Beard, H. S., Chouikha, I. B., Matthews, B. F., Wei, H., et al. (2007). Major differences observed in transcript profiles of blueberry during cold acclimation under field and cold room conditions. *Planta* 225, 735–751. doi: 10.1007/s00425-006-0382-1

Dhanaraj, A. L., Slovin, J. P., and Rowland, L. J. (2004). Analysis of gene expression associated with cold acclimation in blueberry floral buds using expressed sequence tags. *Plant Sci.* 166, 863–872. doi: 10.1016/j.plantsci.2003.11.013

Dhanaraj, A. L., Slovin, J. P., and Rowland, L. J. (2005). Isolation of a cDNA clone and characterization of expression of the highly abundant, cold acclimation-associated 14 kDa dehydrin of blueberry. *Plant Sci.* 168, 949–957. doi: 10.1016/j.plantsci.2004.11.007

Die, J. V., and Rowland, L. J. (2013). Advent of genomics in blueberry. *Mol. Breed.* 32, 493–504. doi: 10.1007/s11032-013-9893-1

Draper, A., and Hancock, J. (2003). Florida 4B: native blueberry with exceptional breeding value. *J. Am. Pomol. Soc.* 57, 138–141.

Ehlenfeldt, M. K. (1994). The genetic composition and tetrasomic inbreeding coefficients of highbush blueberry cultivars. *HortScience* 29, 1342–1345.

Ehlenfeldt, M. K. (2003). Investigations of metaxenia in northern highbush blueberry (*Vaccinium corymbosum* L.) cultivars. *J. Am. Pomol. Soc.* 57, 26–31.

Ehlenfeldt, M. K., and Rowland, L. J. (2006). Cold-hardiness of *Vaccinium ashei* and V. constablaei germplasm and the potential for northern-adapted rabbiteye cultivars. *Acta Hortic.* 715, 77–80. doi: 10.17660/actahortic.2006.715.8

Ehlenfeldt, M. K., Ogden, E. L., Rowland, L. J., and Vinyard, B. (2006). Evaluation of midwinter cold hardiness among 25 rabbiteye blueberry cultivars. *HortScience* 41, 579–581.

Ehlenfeldt, M. K., Rowland, L. J., and Arora, R. (2003). Bud hardiness and deacclimation in blueberry cultivars with varying species ancestry: Flowering time may not be a good indicator of deacclimation. *Acta Hortic.* 626, 39–44. doi: 10.17660/actahortic.2003.626.4

Ehlenfeldt, M. K., Rowland, L. J., Ogden, E. L., and Vinyard, B. T. (2007). Floral bud cold hardiness of *Vaccinium ashei*, V. constablaei, and hybrid derivatives and the potential for producing northern-adapted rabbiteye cultivars. *HortScience* 42, 1131–1134.

Eichholz, I., Huyskens-Keil, S., Keller, A., Ulrich, D., Kroh, L. W., and Rohn, S. (2011). UV-B-induced changes of volatile metabolites and phenolic compounds in blueberries (*Vaccinium corymbosum* L.). *Food Chem.* 126, 60–64. doi: 10.1016/j.foodchem.2010.10.071

El-Shiekh, A., Wildung, D. K., Luby, J. J., Sargent, K. L., and Read, P. E. (1996). Long-term effects of propagation by tissue culture or softwood single-node cuttings on growth habit, yield, and berry weight of 'Northblue' blueberry. *J. Am. Soc. Hortic. Sci.* 121, 339–342.

Erb, W. A., Draper, A. D., and Swartz, H. J. (1988a). Methods of screening blueberry populations for drought resistance. *HortScience* 25, 312–314.

Erb, W. A., Draper, A. D., and Swartz, H. J. (1988b). Screening interspecific blueberry seedling populations for drought resistance. *J. Am. Soc. Hortic. Sci.* 113, 599–604.

Escobar-Opazo, A. (2015). *Spectroradiometry Estimation Models for Blueberries Genotype selection under Drought Conditions, with or without Heat Stress.* Master thesis. Universidad de Talca, Talca.

Estrada, F., Escobar, A., Romero-Bravo, S., González-Talice, J., Poblete-Echeverría, C., Caligari, P. D. S., et al. (2015). Fluorescence phenotyping in blueberry breeding for genotype selection under drought conditions, with or without heat stress. *Sci. Hortic.* 181, 147–161. doi: 10.1016/j.scienta.2014.11.004

Finn, C. E., Hancock, J. F., Mackey, T., and Serçe, S. (2003). Genotype x environment interactions in highbush blueberry (Vaccinium sp. L.) families grown in Michigan and Oregon. *J. Am. Soc. Hortic. Sci.* 128, 196–200.

Finn, C., and Luby, J. J. (1986). Inheritance of fruit development interval and fruit size in blueberry progenies. *J. Am. Soc. Hortic. Sci.* 111, 784–788.

Friend, A. D. (2010). Terrestrial plant production and climate change. *J. Exp. Bot.* 61, 1293–1309. doi: 10.1093/jxb/erq019

Furbank, R. T., and Tester, M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Galletta, G. J., and Ballington, J. R. (1996). "Blueberries, cranberries and lingonberries," in *Fruit Breeding, Vine and Small Fruit Crops*, Vol. 2, eds J. Janick and J. N. Moore (New York, NY: John Wiley and Sons, Inc), 1–107.

Garriga, M., Retamales, J. B., Romero-Bravo, S., Caligari, P. D. S., and Lobos, G. A. (2014). Chlorophyll, anthocyanin and gas exchangechanges assessed by spectroradiometry in Fragaria chiloensis under salt stress. *J. Integr. Plant Biol.* 56, 505–515. doi: 10.1111/jipb.12193

Gough, R. E. (1994). *The Highbush Blueberry and its Management.* Binghamton, NY: Food Product Press an Imprint of the Haworth Press, Inc.

Guidetti, R., Berghi, R., Bodria, L., Spinardi, A., Mignani, I., and Folini, L. (2009). Prediction of blueberry (*Vaccinium corymbosum*) ripeness by a portable Vis-NIR device. *Acta Hortic.* 810, 877–885. doi: 10.17660/actahortic.2009.810.117

Gupton, C. L. (1983). Variability among rabbiteye blueberry cultivars for tolerance of flowers to frost. *HortScience* 18, 713–714.

Hancock, J. F., and Siefker, J. H. (1982). Levels of inbreeding in highbush blueberry cultivars. *HortScience* 17, 363–366.

Hancock, J. F., Erb, W. A., Goulart, B. L., and Scheerens, J. C. (1997). Blueberry hybrids with complex genetic backgrounds evaluated on mineral soils: cold hardiness as influenced by parental species and location. *Acta Hortic.* 446, 389–396. doi: 10.17660/actahortic.1997.446.57

Hancock, J. F., Haghighi, K., Krebs, S. L., Flore, J. A., and Draper, A. D. (1992). Photosynthetic heat stability in highbush blueberries and the possibility of genetic improvement. *HortScience* 27, 1111–1112.

Hancock, J. F., Lyrene, P., Finn, C. E., N., V., and Lobos, G. A. (2008a). "Blueberry and cranberry," in T*emperate Fruit Crop Breeding: Germplasm to Genomics*, ed J. F. Hancock (Dordrecht: Kluwer Academic Publishers), 115–149.

Hancock, J. F., Sjulin, T. M., and Lobos, G. A. (2008b). "Strawberries," in *Temperate Fruit Crop Breeding: Germplasm to Genomics* ed J. F. Hancock (Dordrecht: Kluwer Academic Publishers), 393–437.

Hancock, J., Nelson, J., Bittenbender, H., Callow, P., Cameron, J., Krebs, S., et al. (1987). Variation among highbush blueberry cultivars in susceptibility to spring frost. *J. Am. Soc. Hortic. Sci.* 112, 702–706.

Hancock, J., Sakin, M., and Callow, P. (1991). Heritability of flowering and harvest dates in *Vaccinium corymbosum*. *Fruit Varieties J.* 45, 173–176.

Hanson, E. J., Berkheimer, S. F., and Hancock, J. F. (2007). Seasonal changes in the cold hardiness of the flower buds of highbush blueberry with varying species ancestry. *J. Am. Pomol. Soc.* 61, 14–18.

Houle, D., Govindaraju, D. R., and Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet.* 11, 855–866. doi: 10.1038/nrg2897

Huovinen, P., Gómez, I., and Lovengreen, C. (2006). A five-year study of solar ultraviolet radiation in southern Chile (39 degrees S): potential impact on physiology of coastal marine algae? *Photochem. Photobiol.* 82, 515–522. doi: 10.1562/2005-07-05-RA-601

Inoue, Y., Peñuelas, J., Miyata, A., and Mano, M. (2008). Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO2 flux measurements in rice. *Remote Sens. Environ.* 112, 156–172. doi: 10.1016/j.rse.2007.04.011

IPCC. (2007a). "Climate change 2007: impacts, adaptation and vulnerability," in *Contribution of Working Group II to the 4th Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge, UK).

IPCC. (2007b). "Climate Change 2007: The Physical Basis," in *Contribution of Working Group I to the 4th Assessment of the Intergovernmental Panel on Climate Change* (Cambridge, UK).

Jin, B., Wang, L., Wang, J., Jiang, K. Z., Wang, Y., Jiang, X. X., et al. (2011). The effect of experimental warming on leaf functional traits, leaf structure and leaf biochemistry in Arabidopsis thaliana. *BMC Plant Biol.* 11:35. doi: 10.1186/1471-2229-11-35

Kakani, V. G., Reddy, K. R., Zhao, D., and Sailaja, K. (2003). Field crop responses to ultraviolet-B radiation: a review. *Agric. For. Meteorol.* 120, 191–218. doi: 10.1016/j.agrformet.2003.08.015

Kazantzidis, A., Tourpali, K., and Bais, A. F. (2010). Variability of cloud-free ultraviolet dose rates on global scale due to modeled scenarios of future ozone recovery. *Photochem. Photobiol.* 86, 117–122. doi: 10.1111/j.1751-1097.2009.00645.x

Kipp, S., Mistele, B., Baresel, P., and Schmidhalter, U. (2014). High-throughput phenotyping early plant vigour of winter wheat. *Eur. J. Agron.* 52, Part B, 271–278. doi: 10.1016/j.eja.2013.08.009

Kossuth, S. V., and Biggs, R. H. (1978). "Sunburned blueberries," in *Proceedings of the Florida State Horticultural Society* (Gainesville, FL).

Lee, H. J., Kim, S. J., Yu, D. J., Lee, B. Y., and Kim, T. C. (2006). Changes in photosynthetic characteristics in water-stressed 'Rancocas' blueberry leaves. *Acta Hortic.* 715, 111–117. doi: 10.17660/actahortic.2006.715.15

Lee, J. H., Yu, D. J., Kim, S. J., Choi, D., and Lee, H. J. (2012). Intraspecies differences in cold hardiness, carbohydrate content and β-amylase gene expression of *Vaccinium corymbosum* during cold acclimation and deacclimation. *Tree Physiol.* 32, 1533–1540. doi: 10.1093/treephys/tps102

Leiva-Valenzuela, G. A., and Aguilera, J. M. (2013). Automatic detection of orientation and diseases in blueberries using image analysis to improve their postharvest storage quality. *Food Control* 33, 166–173. doi: 10.1016/j.foodcont.2013.02.025

Leiva-Valenzuela, G. A., Lu, R., and Aguilera, J. M. (2012). "Assessment of internal quality of blueberry using hyperspectral imaging," in *American Society of Agricultural and Biological Engineers Annual International Meeting* (Dallas, TX), 2333–2347.

Leiva-Valenzuela, G. A., Lu, R., and Aguilera, J. M. (2013). Prediction of firmness and soluble solids content of blueberries using hyperspectral reflectance imaging. *J. Food Eng.* 115, 91–98. doi: 10.1016/j.jfoodeng.2012.10.001

Leiva-Valenzuela, G. A., Lu, R., and Aguilera, J. M. (2014). Assessment of internal quality of blueberries using hyperspectral transmittance and reflectance images with whole spectra or selected wavelengths. *Innov. Food Sci. Emerg. Technol.* 24, 2–13. doi: 10.1016/j.ifset.2014.02.006

Lin, W., and Pliszka, K. (2003). Comparison of spring frost tolerance among different highbush blueberry (*Vaccinium corymbosum* L.) cultivars. *Acta Hortic.* 626, 337–341. doi: 10.17660/actahortic.2003.626.45

Lobos, G. A., Acevedo-Opazo, C., Guajardo-Moreno, A., Valdés-Gómez, H., Taylor, J., and Laurie, V. F. (2015). Effects of kaolin-based particle film and fruit zone netting on Cabernet Sauvignon physiology and fruit quality. *J. Int. Sci. Vigne Vin* 49, 137–144.

Lobos, G. A., Matus, I., Rodriguez, A., Romero-Bravo, S., Araus, J. L., and del Pozo, A. (2014). Wheat genotypic variability in grain yield and carbon isotope discrimination under Mediterranean conditions assessed by spectral reflectance. *J. Integr. Plant Biol.* 56, 470–479. doi: 10.1111/jipb.12114

Lobos, G. A., Retamales, J. B., Del Pozo, A., Hancock, J. F., and Flore, J. A. (2009). Physiological response of *Vaccinium corymbosum* 'Elliott' to shading nets in Michigan. *Acta Hortic.* 810, 465–470. doi: 10.17660/actahortic.2009.810.60

Lobos, G. A., Retamales, J. B., Hancock, J. F., Flore, J. A., Cobo, N. G., and del Pozo, A. (2012). Spectral irradiance, gas exchange characteristics and leaf traits of *Vaccinium corymbosum* L. 'elliott' grown under photo-selective nets. *Environ. Exp. Bot.* 75, 142–149. doi: 10.1016/j.envexpbot.2011.09.006

Lobos, G. A., Retamales, J. B., Hancock, J. F., Flore, J. A., Romero-Bravo, S., and del Pozo, A. (2013). Productivity and fruit quality of *Vaccinium corymbosum* cv. Elliott under photo-selective shading nets. *Sci. Hortic.* 153, 143–149. doi: 10.1016/j.scienta.2013.02.012

Luby, J. J., Ballington, J. R., Draper, A. D., Pliszka, K., and Austin, M. E. (1991). Blueberries and cranberries (Vaccinium). *Acta Hortic.* 290, 393–458. doi: 10.17660/actahortic.1991.290.9

Lyrene, P. (1985). Effects of year and genotype on flowering and ripening dates in rabbiteye blueberry. *HortScience* 20, 407–409.

Lyrene, P. (1991). Fertile derivatives from sparkleberry × blueberry crosses. *J. Am. Soc. Hortic. Sci.* 116, 899–902.

Lyrene, P. (2006). Breeding southern highbush and rabbiteye blueberries. *Acta Hortic.* 715, 29–36. doi: 10.17660/actahortic.2006.715.1

Lyrene, P. M. (2002). Development of highbush blueberry cultivars adapted to Florida. *J. Am. Pomol. Soc.* 56, 79–85.

Lyrene, P. M., and Sherman, W. B. (2000). 'Star' southern highbush blueberry. *HortScience* 35, 956–957.

Madronich, S., McKenzie, R. L., Björn, L. O., and Caldwell, M. M. (1998). Changes in biologically active ultraviolet radiation reaching the Earth's surface. *J. Photochem. Photobiol B Biol.* 46, 5–19. doi: 10.1016/S1011-1344(98)00182-1

Mainland, C. M. (1989). "Harvesting, sorting and packing quality blueberries," in *23rd Annual Open House, Southeastern Blueberry Council*, North Carolina Univ.

Mainland, C. M., Buchanan, D. W., and Bartholic, J. F. (1977). The effects of five chilling regimes on bud break of highbush (*Vaccinium corymbosum* L.) and rabbiteye (*V. ashei* Reade) blueberry hardwood cuttings. *HortScience* 12, 411.

Maqbool, R., Percival, D. C., Adl, M. S., Zaman, Q. U., and Buszard, D. (2012). In situ estimation of foliar nitrogen in wild blueberry using reflectance spectra. *Can. J. Plant Sci.* 92, 1155–1161. doi: 10.4141/cjps2011-203

McKenzie, R. L., Aucamp, P. J., Bais, A. F., Björn, L. O., Ilyas, M., and Madronich, S. (2011). Ozone depletion and climate change: impacts on UV radiation. *Photochem. Photobiol. Sci.* 10, 182–198. doi: 10.1039/c0pp90034f

Meza, F. J., Wilks, D. S., Riha, S. J., and Stedinger, J. R. (2003). Value of perfect forecasts of sea surface temperature anomalies for selected rain-fed agricultural locations of Chile. *Agric. Forest Meteorol.* 116, 117–135. doi: 10.1016/S0168-1923(03)00004-2

Mingeau, M., Perrier, C., and Améglio, T. (2001). Evidence of drought-sensitive periods from flowering to maturity on highbush blueberry. *Sci. Hortic.* 89, 23–40. doi: 10.1016/S0304-4238(00)00217-X

Moon, J. W. Jr., Flore, J. A., and Hancock, J. F. Jr. (1987a). A comparison of carbon and water vapor gas exchange characteristics between a diploid and highbush blueberry. *J. Am. Soc. Hortic. Sci.* 112, 134–138.

Moon, J. W. Jr., Hancock, J. F. Jr., Draper, A. D., and Flore, J. A. (1987b). Genotypic differences in the effect of temperature on CO_2 assimilation and water use efficiency in blueberry. *J. Am. Soc. Hortic. Sci.* 112, 170–173.

Moretti, C. L., Mattos, L. M., Calbo, A. G., and Sargent, S. A. (2010). Climate changes and potential impacts on postharvest quality of fruit and vegetable crops: a review. *Food Res. Int.* 43, 1824–1832. doi: 10.1016/j.foodres.2009.10.013

Naik, D., Dhanaraj, A. L., Arora, R., and Rowland, L. J. (2007). Identification of genes associated with cold acclimation in blueberry (*Vaccinium corymbosum* L.) using a subtractive hybridization approach. *Plant Sci.* 173, 213–222. doi: 10.1016/j.plantsci.2007.05.003

NASA. (2015). *NASA, NOAA Find 2014 Warmest Year in Modern Record* (New York, NY).

NeSmith, D. S., Krewer, G., and Lindstrom, O. M. (1999). Fruit set of rabbiteye blueberry (*Vaccinium ashei*) after subfreezing temperatures. *J. Am. Soc. Hortic. Sci.* 124, 337–340.

Olmstead, J. W., Armenta, H. P. R., and Lyrene, P. M. (2013). Using sparkleberry as a genetic source for machine harvest traits for southern highbush blueberry. *HortTechnology* 23, 419–424.

Patten, K., Neuendorff, E. W., Nimr, G., Clark, J. R., and Fernandez, G. (1991). Cold injury of southern blueberries as a function of germplasm and season of flower bud development. *HortScience* 26, 18–20.

Peñuelas, J., Filella, I., and Gamon, J. A. (1995). Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytol.* 131, 291–296.

Peñuelas, J., Filella, I., Biel, C., Serrano, L., and Savé, R. (1993). The reflectance at the 950–970 nm region as an indicator of plant water status. *Int. J. Remote Sens.* 14, 1887–1905.

Percival, D. C., Sharpe, S., Maqbool, R., and Zaman, Q. (2012). Narrow band reflectance measurements can be used to estimate leaf area index, flower number, fruit set and berry yield of the wild blueberry (*Vaccinium angustifolium* Ait.). *Acta Hortic.* 926, 363–370. doi: 10.17660/actahortic.2012.926.51

Perkins-Veazie, P., Collins, J. K., and Howard, L. (2008). Blueberry fruit response to postharvest application of ultraviolet radiation. *Postharvest Biol. Technol.* 47, 280–285. doi: 10.1016/j.postharvbio.2007.08.002

Perrier, C., Mingeau, M., and Améglio, T. (2000). Effects of water stress on transpiration, radial growth and yield in highbush blueberry. *Acta Hortic.* 537, 923–928. doi: 10.17660/actahortic.2000.537.112

Peshlov, B. N., Dowelt, F. E., Drummond, F. A., and Donahue, D. W. (2009). Comparison of three near infrared spectrophotometers for infestation detection in wild blueberries using multivariate calibration models. *J. Near Infrared. Spec.* 17, 203–212. doi: 10.1255/jnirs.842

Ralph, P. J., and Gademann, R. (2005). Rapid light curves: a powerful tool to assess photosynthetic activity. *Aquat. Bot.* 82, 222–237. doi: 10.1016/j.aquabot.2005.02.006

Retamales, J. B., and Hancock, J. F. (2012). *Blueberries*. Wallingford, UK: CABI Publishing.

Retamales, J. B., Montecino, J. M., Lobos, G. A., and Rojas, L. A. (2008). Colored shading nets increase yields and profitability of highbush blueberries. *Acta Hortic.* 770, 193–197. doi: 10.17660/actahortic.2008.770.22

Rho, H., Yu, D. J., Kim, S. J., and Lee, H. J. (2012). Limitation factors for photosynthesis in 'Bluecrop' highbush blueberry (*Vaccinium corymbosum*) leaves in response to moderate water stress. *J. Plant Biol.* 55, 450–457. doi: 10.1007/s12374-012-0261-1

Rowland, L. J., and Levi, A. (1994). RAPD-based genetic linkage map of blueberry derived from a cross between diploid species (*Vaccinium darrowi* and *V. elliottii*). *Theor. Appl. Genet.* 87, 863–868. doi: 10.1007/BF00221139

Rowland, L. J., Bell, D. J., Alkharouf, N., Bassil, N. V., Drummond, F. A., Beers, L., et al. (2012). Generating genomic tools for blueberry improvement. *IJFS* 12, 276–287. doi: 10.1080/15538362.2011.619452

Rowland, L. J., Dhanaraj, A. L., Naik, D., Alkharouf, N., Matthews, B., and Arora, R. (2008). Study of cold tolerance in blueberry using EST libraries, cDNA microarrays, and subtractive hybridization. *HortScience* 43, 1975–1981.

Rowland, L. J., Ogden, E. L., and Ehlenfeldt, M. K. (2010). EST-PCR markers developed for highbush blueberry are also useful for genetic fingerprinting and relationship studies in rabbiteye blueberry. *Sci. Hortic.* 125, 779–784. doi: 10.1016/j.scienta.2010.05.008

Rowland, L. J., Ogden, E. L., Bassil, N., Buck, E. J., McCallum, S., Graham, J., et al. (2014). Construction of a genetic linkage map of an interspecific diploid blueberry population and identification of QTL for chilling requirement and cold hardiness. *Mol. Breed.* 34, 2033–2048. doi: 10.1007/s11032-014-0161-9

Rowland, L. J., Ogden, E. L., Ehlenfeldt, M. K., and Vinyard, B. (2005). Cold hardiness, deacclimation kinetics, and bud development among 12 diverse blueberry genotypes under field conditions. *J. Am. Soc. Hortic. Sci.* 130, 508–514.

Rowland, L. J., Ogden, E. L., Takeda, F., Glenn, D. M., Ehlenfeldt, M. K., and Vinyard, B. T. (2013). Variation among highbush blueberry cultivars for frost tolerance of open flowers. *HortScience* 48, 692–695.

Rowland, L. J., Panta, G. R., Mehra, S., and Parmentier-Line, C. (2004). Molecular genetic and physiological analysis of the cold-responsive dehydrins of blueberry. *J. Crop Improv.* 10, 53–76. doi: 10.1300/J411v10n01_05

Rozema, J., van de Staaij, J., Björn, L. O., and Caldwell, M. M. (1997). UV-B as an environmental factor in plant life: stress and regulation. *Trends Ecol. Evol.* 12, 22–28. doi: 10.1016/S0169-5347(96)10062-8

Ruhland, C. T., Xiong, F. S., Clark, W. D., and Day, T. A. (2005). The influence of ultraviolet-B radiation on growth, hydroxycinnamic acids and flavonoids of Deschampsia antarctica during Springtime ozone depletion in Antarctica. *Photochem. Photobiol.* 81, 1086–1093. doi: 10.1562/2004-09-18-RA-321

Santibañez, F., and Santibañez, P. (2007). Cambio climático y degradación de tierras en Lationoamerica y Chile. *Ambiente Desarrollo* 3, 54–63.

Santibañez, F., and Santibañez, P. (2008). "Efectos del cambio climático en la agricultura," in *IV Jornadas de Actualización en Riego y Fertirriego*, Facultad de Ciencias Agrarias – U. N. Cuyo.

Semerdjieva, S. I., Phoenix, G. K., Hares, D., Gwynn-Jones, D., Callaghan, T. V., and Sheffield, E. (2003a). Surface morphology, leaf and cuticle thickness of four dwarf shrubs from a sub-Arctic heath following long-term exposure to enhanced levels of UV-B. *Physiol. Plant.* 117, 289–294. doi: 10.1034/j.1399-3054.2003.00006.x

Semerdjieva, S. I., Sheffield, E., Phoenix, G. K., Gwynn-Jones, D., Callaghan, T. V., and Johnson, G. N. (2003b). Contrasting strategies for UV-B screening in sub-Arctic dwarf shrubs. *Plant Cell Environ.* 26, 957–964. doi: 10.1046/j.1365-3040.2003.01029.x

Sims, D. A., and Gamon, J. A. (2003). Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sens. Environ.* 84, 526–537. doi: 10.1016/S0034-4257(02)00151-7

Sinelli, N., Casiraghi, E., Barzaghi, S., Brambilla, A., and Giovanelli, G. (2011). Near infrared (NIR) spectroscopy as a tool for monitoring blueberry osmo-air dehydration process. *Food Res. Int.* 44, 1427–1433. doi: 10.1016/j.foodres.2011.02.046

Sinelli, N., Spinardi, A., Di Egidio, V., Mignani, I., and Casiraghi, E. (2008). Evaluation of quality and nutraceutical content of blueberries (*Vaccinium corymbosum* L.) by near and mid-infrared spectroscopy. *Postharvest Biol. Technol.* 50, 31–36. doi: 10.1016/j.postharvbio.2008.03.013

Spiers, J. M. (1976). Chilling regimens affect bud break in 'Tifblue' rabbiteye blueberry. *J. Am. Soc. Hortic. Sci.* 101, 84–86.

Sugiyama, T., Sugiyama, J., Tsuta, M., Fujita, K., Shibata, M., Kokawa, M., et al. (2010). NIR spectral imaging with discriminant analysis for detecting foreign materials among blueberries. *J. Food Eng.* 101, 244–252. doi: 10.1016/j.jfoodeng.2010.06.026

Tanner, C. B. (1963). Plant temperatures. *Agron. J.* 55, 210–211. doi: 10.2134/agronj1963.00021962005500020043x

Trehane, J. (2004). *Blueberries, Blueberries and Other Vacciniums*. Portland, OR: Timber Press.

Tsuta, M., Takao, T., Sugiyama, J., Wada, Y., and Sagara, Y. (2006). Foreign substance detection in blueberry fruits by spectral imaging. *Food Sci. Technol. Res.* 12, 96–100. doi: 10.3136/fstr.12.96

Tucker, C. J., and Sellers, P. J. (1986). Satellite remote sensing of primary production. *Int. J. Remote Sens.* 7, 1395–1416. doi: 10.1080/0143116860848944

United Nations. (2010). *Economics of Climate Change in Latin America and the Caribbean. Summary 2010*. Santiago: United Nations.

Wei, H., Dhanaraj, A. L., Rowland, L. J., Fu, Y., Krebs, S. L., and Arora, R. (2005). Comparative analysis of expressed sequence tags from cold-acclimated and non-acclimated leaves of *Rhododendron catawbiense* Michx. *Planta* 221, 406–416. doi: 10.1007/s00425-004-1440-1

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., and Conley, M. M., Wang, G. (2012). Field-based phenomics for plant genetics research. *Field Crops Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003

WMO. (2011). *Scientific Assessment of Ozone Depletion: 2010, Global Ozone Research and Monitoring Project–Report No. 52. World Metereological Organization*. Geneva.

Yáñez, P., Retamales, J. B., Lobos, G. A., and del Pozo, A. (2009). Light environment within mature rabbiteye blueberry canopies influences flower bud formation. *Acta Hortic.* 810, 417–473. doi: 10.17660/ActaHortic.2009.810.61

Yang, C., and Lee, W. S. (2011). "Spectral signatures of blueberry fruits and leaves," in *American Society of Agricultural and Biological Engineers Annual International Meeting 2011, ASABE 2011* (Louisville, KY).

Yang, C., Lee, W. S., and Gader, P. (2014). Hyperspectral band selection for detecting different blueberry fruit maturity stages. *Comput. Electron. Agric.* 109, 23–31. doi: 10.1016/j.compag.2014.08.009

Yang, C., Lee, W. S., and Williamson, J. G. (2012) Classification of blueberry fruit and leaves based on spectral signatures. *Biosys. Eng.* 113, 351–362.

The *qTSN* Positive Effect on Panicle and Flag Leaf Size of Rice is Associated with an Early Down-Regulation of Tillering

Dewi E. Adriani[1,2], Tanguy Lafarge[1], Audrey Dardou[1], Aubrey Fabro[3], Anne Clément-Vidal[1], Sudirman Yahya[4], Michael Dingkuhn[1,3] and Delphine Luquet[1]*

[1] CIRAD, UMR AGAP, F-34398 Montpellier, France, [2] Faculty of Agriculture, University of Lambung Mangkurat, Banjarbaru, Indonesia, [3] Crop and Environment Science Division, International Rice Research Institute, Los Baños, Philippines, [4] Department of Agronomy and Horticulture, Bogor Agricultural University, Bogor, Indonesia

The *qTSN4* was identified as rice QTL (Quantitative Traits Locus) increasing total spikelet number per panicle and flag leaf area but potentially reducing panicle number depending on the environment. So far, this trade-off was mainly observed at grain maturity and not specifically studied in details, limiting the apprehension of the agronomic interest of *qTSN4*. This study aimed to understand the effect of *qTSN4* and of the environment on panicle sizing, its trade-off with panicle number, and finally plant grain production. It compared two high yielding genotypes to their Near Isogenic Lines (NIL) carrying either QTL *qTSN4* or *qTSN12, two* distinct QTLs contributing to the enlarged panicle size, thereafter designated as *qTSN*. Traits describing C sink (organ appearance rate, size, biomass) and source (leaf area, photosynthesis, sugar availability) were dynamically characterized along plant and/or panicle development within two trials (greenhouse, field), each comparing two treatments contrasting for plant access to light (with or without shading, high or low planting densities). The positive effect of *qTSN* on panicle size and flag leaf area of the main tiller was confirmed. More precisely, it could be shown that *qTSN* increased leaf area and internode cross-section, and in some cases of the photosynthetic rate and starch reserves, of the top 3–4 phytomers of the main tiller. This was accompanied by an earlier tillering cessation, that coincided with the initiation of these phytomers, and an enhanced panicle size on the main tiller. Plant leaf area at flowering was not affected by *qTSN* but fertile tiller number was reduced to an extent that depended on the environment. Accordingly, plant grain production was enhanced by *qTSN* only under shading in the greenhouse experiment, where panicle number was not affected and photosynthesis and starch storage in internodes was enhanced. The effect of *qTSN* on rice phenotype was thus expressed before panicle initiation (PI). Whether early tillering reduction or organ oversizing at meristem level is affected first cannot be entirely unraveled. Further studies are needed to better understand any signal involved in this early regulation and the *qTSN* × Environment interactions underlying its agronomic interest.

Keywords: rice, *qTSN4*, down-regulation of tillering, panicle size, top leaf and internode size, main stem growth rate

INTRODUCTION

Grain yield elaboration in cereals depends on the establishment of panicle number per square meter and panicle size, grain filling rate, and individual grain size (Chen et al., 2008; Gaju et al., 2014). All these traits are known to compensate per unit area i.e., among plants (Zhang and Yamagishi, 2010) and within the plant where they compete for the same pool of C and N resources (Okawa et al., 2003; Hashida et al., 2013). Amongst these traits in rice, panicle number, and panicle size are those characterized with the highest plasticity under favorable growing conditions thus with the highest impact on yield elaboration. These two traits are determined along plant cycle, particularly based on tillering and green leaf area dynamics, internode reserve remobilization, and reproductive sink size and number. All these processes are depending on C assimilated and N availability, as reported by Lafarge and Bueno (2009) and Dingkuhn et al. (2015) in rice, and by Dreccer et al. (2008) in wheat. Regarding panicle size, many studies reported that panicle development, and the determination of spikelet number per panicle, is closely correlated to early plant vigor and underlying traits such as leaf appearance rate (Dong et al., 2004; Streck et al., 2009; Itoh and Shimizu, 2012; Rebolledo et al., 2012), tillering (Lafarge et al., 2002, 2010; Borràs-Gelonch et al., 2012), culm (Fujita and Yoshida, 1984; Wu et al., 2011), and peduncle (Liu et al., 2008) size, plant height (Wang et al., 2014; Chen et al., 2015), or even neck internode diameter (Zhang and Yamagishi, 2010).

Several genes were reported to control morphogenetic processes positively during both the vegetative and the reproductive phases. Amongst them, the genes *MOC1* (Monoculm 1, Li et al., 2003) and LAX (Komatsu et al., 2003) were shown to promote axillary meristem initiation and accordingly to regulate both shoot and panicle branching. The *Gn1* gene is involved in the control of plant height and grain number per panicle (Ashikari et al., 2005). The *DEP1* enhances shoot apical meristem activity and grain number per panicle (Huang et al., 2009). The *OsSPL14* promotes early tillering and grain number per panicle (Miura et al., 2010). The *APO1* increases grain number per plant and harvest index (Ikeda et al., 2007; Terao et al., 2010). The *OsEBS* enhances plant biomass and spikelet number per panicle (Dong et al., 2013). By contrast, other genes were shown to have trade-off effects on some morphogenetic traits expressed during the vegetative and the reproductive phases. This is the case of the rice mutant *this1* characterized by higher tiller number but lower plant height and fertile spikelet number compared to the wild type (Liu et al., 2013). The *Ghd7* increases spikelet number per panicle through panicle branching but decreases tillering in a density-dependent manner (Weng et al., 2014); the OsSPS1 gene inhibits plant height but enhances spikelet number per panicle (Hashida et al., 2013). The *Ltn* (low tiller number) and similarly *SPIKE* (*qTSN4*) induces higher total spikelet number per panicle but decreases final panicle number (Fujita et al., 2010, 2013).

These studies point out that a gene has unlikely a positive effect on both organ size and number and accordingly that molecular breeding efforts are frequently confronted to the issue of physiological trade-offs among traits, compromising the direct use of a candidate gene or QTL in crop improvement. The trade-off between tiller number and organ size is of major concern for rice, not only with respect to vegetative growth (Rebolledo et al., 2012) but also reproductive growth as it comes as a parameter regulating panicle size vs. number (Lafarge et al., 2010). The *qTSN4* was recently identified as a QTL enhancing flag leaf width and total spikelet number per panicle in IR64 background (Fujita et al., 2009, 2012, 2013). This QTL is known to co-locate with *Nal1* gene involved in the determination of leaf structure, veining pattern, and carboxylation (Qi et al., 2008), as well as other early traits like stem length and vascular bundle number (Fujita et al., 2012, 2013). It was also shown that its positive effect on panicle size was potentially depressive on panicle number (Fujita et al., 2012, 2013). Recently, Okami et al. (2015) confirmed that an isogenic line carrying *qTSN4* produced fewer but larger tillers than its parent (IR64).

No studies, however, were conducted to further understand at plant level the physiological and environmental determinisms regulating (i) the trade-off between panicle size and number in genotypes introgressed with *qTSN4* and (ii) accordingly, the positive effect of *qTSN4* on plant grain production. This is the aim of the present study that, analyzes the effect of *qTSN* on plant growth and development and thus on yield formation processes along plant cycle in contrasting situations of plant access to light. Two high yielding genotypes (IR64 and IRRI146) were compared to their Near Isogenic Lines (NIL) carrying QTL *qTSN4* (Fujita et al., 2013) and QTL *qTSN12*, a different QTL reported for having a similar function as *qTSN4*, i.e., enhanced panicle size (Ishimaru, personal communication). This was conducted within three trials under contrasted situations of plant access to light and thus to C assimilates in order to modulate the level of competition among sinks within the plant. Plant growth was then characterized in terms of morphogenesis and biomass accumulation per organ type along plant development.

MATERIALS AND METHODS

Plant Materials

Two pairs of parents/near-isogenic lines (P/NIL): IR64 vs. IR64 NIL and IRRI146 vs. IRRI146 NIL were studied in two experiments. In one of these experiments, an additional pair of P/NIL was considered asIR64 and NIL1 carrying *qTSN12*. This QTL has been observed to increase panicle size and reduce panicle number—as observed with *qTSN4*—and was detected in chromosome 12 from donor parent YP4. The NILs were developed by self-pollination of a plant selected from BC4F2 population (in IR64 background) and BC3F1 population (in IRRI146 background; Fujita et al., 2013). The IR64 NIL was identified by composite interval mapping carrying the high total spikelet number (TSN) QTL between Simple Sequence Repeat (SSR) markers RM3423 and RM17492 on the long arm of chromosome 4 (Fujita et al., 2012), whereas IRRI146 NIL was developed from recurrent backcrossing to IRRI146 and marker-assisted selection (MAS; Fujita et al., 2013). The details (cross combination, donor and category) of plant materials are available in **Table 1**. The recipient lines (parents) were chosen because

TABLE 1 | Details of plant materials.

Designation	Species	Cross combination	Donor	Category
IR64	Indica	IR64		Recurrent parent recipient
IR64 NIL	Isoline of IR64	(IR64/IR68522-10-2-2//3*IR64)/IR64	IR68522-10-2-2	IR64-qTSN4.4
IR64 NIL1	Isoline of IR64	(IR64/IR65564-2-2-3//3*IR64)/IR65	IR65564-2-2-3	IR64-qTSN12.2
IRRI146	Indica	NSIC Rc158		Recurrent parent recipient
IRRI146 NIL	Isoline of NSIC Rc158	NSIC Rc158/IR65564-2-2-3//3*NSIC Rc158	IR65564-2-2-3	NSIC Rc158-qTSN4.1

of the wide adaptability as of IR64 as a mega variety grown in many parts of the world (Khush, 1987) and IRRI146, a 2nd-generation New Plant Type (NPT) variety developed at IRRI and released in the Philippines in 2007 under the name NSIC Rc158, a high-yielding *indica* cultivar as well (Brennan and Malabayabas, 2011).

Experiments

GH-CNRS

Experiment I was conducted in the greenhouse from May to August 2013 at the National Center for Scientific Research (CNRS), Montpellier, France. This experiment tested *qTSN4* in IRRI146 and IR64 genetic backgrounds comparing one control treatment (C) and a shading (S) treatment with a 58% light attenuation from panicle initiation (PI) up to heading (H) by using a gray net all around the table bringing the plants.

The seeds were grown in a germination chamber at 29°C, then transplanted 4 days after germination in 3 l pots (three seeds per pot) when seedlings were about 3 cm tall. The thinning of plant population to one plant per pot (downsizing to 45.65 plants m^{-2}) was conducted at four-leaf stage. Pots contained about $3/4$ of their volume with EGOT 140 media (17N-10P-14K, pH of 5) and were placed side by side (corresponding to 14.8- cm spacing) on tables filled with 5 cm water depth. Basal fertilizer was applied using Basacot 6M+ at 2g l^{-1}, 11N-9P-19K +2Mg incorporated before transplanting. Pots were arranged in four aluminum tables and each table was containing 104 pots, including border plants, at the beginning of the experiment. The tables were moved every week from 2 weeks after transplanting until maturity to avoid any bias due to the green house structure.

Weather data were collected from the AWS (*Automatic Weather Station*) that was installed in the center of the tables measuring Photosynthetic Active Radiation (PAR), global radiation (Rg), air temperature (T), and relative humidity (RH). The average daily air temperature throughout the crop cycle was 27.3 ± 0.6°C. The average daily PAR for the whole crop cycle was 24.7 ± 7.1 mol m^{-2} d^{-1} under full light and 10.3 ± 3.4 mol m^{-2} d^{-1} under shading, and the average RH was 66.8 ± 7.7%.

Field-IRRI

Experiment II was performed in the field at International Rice Research Institute (IRRI) experiment station in Los Baños, Philippines (14°11′N, 121°15′E, 21 m altitude), from December 2013 to April 2014. This experiment tested *qTSN4* and *qTSN12* effect in IRRI146 and IR64 genetic backgrounds and at two

planting densities, low (LD, 25 plants m^{-2}), and high density (HD, 100 plants m^{-2}).

The seeds were soaked for 24 h, drained and incubated for another 24 h, then sown in the seeding trays in the greenhouse on December 5, 2013. The 2-week old seedlings were transplanted in the field at one plant per hill in a 2 × 2.4 m^2 plots. The field was initially flooded to hold two puddlings and two harrowings, standing water level, 3–5 cm, was maintained as the IRRI guide field standard. Phosphorus (30 kg P ha^{-1}), potassium (40 kg K ha^{-1}), and zinc (5 kg Zn ha^{-1}) were applied and incorporated into all the plots 2 days before transplanting. 60 kg N ha^{-1} was applied 1 day before transplanting, then 40 kg N ha^{-1} and 60 kg N ha^{-1} were applied at mid-tillering and PI stage, respectively.

Weather data were collected from the IRRI meteorological station measuring radiation (MJ m^{-2}), daylight (h), rainfall (mm), evaporation (mm), average temperature (°C), vapor pressure (kPa), RH (%), and wind speed (m s^{-1}). Average daily air temperature throughout crop cycle was 25.6 ± 1.5°C. The average daily PAR for the whole crop cycle was 31.0 ± 11.3 mol m^{-2} d^{-1}, and the average RH was 84.2 ± 4.8%.

GH-IRRI

Experiment III was performed in an open-top green house during the wet season (August to November 2014) supporting the two experiments described above. This experiment adopted split plot design with three replications and was conducted at IRRI, Los Baños, Philippines (14°11′N, 121°15′E, 21 m altitude). The main factor was plant spacing: crowded density, 20 × 20 cm (Cr, 25 plants m^{-2}) and isolated density, 60 × 60 cm (Is, 2.78 plants m^{-2}) from PI up to flowering (FLO). The subsidiary factor was a pair of rice genotype: IRRI146 (NSIC Rc158) recipient line and its NIL (*qTSN4.1*–YP4).

The seeds were soaked for 24 h, drained and incubated for another 24 h, then sown in the 6 l pots (four seeds per pot). Pots contain about $3/4$ of its volume of Andaqueptic Haplaquoll with a topsoil of 39% clay, 46% silt, 14% sand, pH of 6.38. The thinning of plant population to one plant per pot was conducted 2 weeks after sowing. Water level in the pots was always maintained about 8 cm height. 4 g of Ammonium Sulfate (NH$_4$)$_2$(SO$_4$), 2 g of Single Super Phosphate (SSP) and 2 g of muriate of Potash (KCl) were applied and incorporated into all the pots 2 days before sowing, and 2 g of Ammonium Sulfate was applied at PI stage.

Weather data were collected from IRRI meteorological station measuring radiation (MJ m^{-2}), daylight (h), rainfall (mm), evaporation (mm), average temperature (°C), vapor pressure (kPa), RH (%) and wind speed (m s^{-1}). Average daily air

temperature throughout crop cycle was $27.7 \pm 0.9°C$. The average daily PAR for whole crop cycle was 29.8 ± 11.5 $molm^{-2}d^{-1}$, and RH was $85.7 \pm 5.3\%$.

Plant Measurements

Plant Development and Biomass Accumulation
In both GH-CNRS and field-IRRI, leaf appearance on the main tiller and green tiller number were measured every week from 2 weeks after sowing (in GH-CNRS) and 2 weeks after transplanting (in the field) up to flag leaf stage on 3 and 4 sampled plants in GH-CNRS and field, respectively, in every replication. Thermal time was calculated by daily integration of air temperature minus a base temperature of 12°C (Rebolledo et al., 2012).

Dry weight (DW) of plant shoot organs (leaves, stems of the main tiller and the rest of the whole plant, and panicles) was measured after drying for 72 h in an oven at 70°C, during panicle development (PI + 3 weeks in GH-CNRS and GH IRRI; PI + 2 weeks in the field) and at heading (H; in GH-CNRS) or flowering (FLO; in the field). For each of this sampling main stem dry weight at H or FLO and MAT (MS DW H/FLO, MS DW MAT), shoot dry weight at FLO and MAT (Shoot DW H/FLO, Shoot DW MAT), main stem panicle dry weight at MAT (MS PDW MAT) and plant filled grain dry weight at MAT (Plant FGDW MAT) were measured.

At physiological maturity, the DW of all vegetative organs excluding root biomass was measured as well as MS PDW (after drying under the sun). The five plants harvested at maturity in GH were separated into panicles (after taken pictures for P-TRAP analysis), green leaf blades, senescent leaves, and productive stems (culms + sheaths). In the field, all the plants within a soil base area of $0.12 m^2$ per plot were harvested, that is three plants under LD and 12 plants under HD. They were then separated into panicles, green leaf blades, dead tissues, and productive stems. The panicles were hand-threshed and then the filled spikelets were separated from the unfilled by a densitometric column (in GH) or submerging the spikelets in the water (in the field).

PI was determined by dissecting and observing the main tiller of randomized collected plants (border plants for field experiment) from each unit treatment every second day when PI was close. The occurrence of PI was considered when the first row of floral primordial was visible on the shoot apex. Flowering (FLO) was determined within each unit treatment when an average of 75% spikelets per panicle of the main tiller exerted their anthers. Plants were considered at physiological maturity when 75% of the grains of the panicles had turned yellow and the texture was in dough stage.

Leaf Area
In GH-CNRS, individual leaf area on the main stem was measured by using LI–3100C Area Meter (Lincoln, NE, USA). In the field and GH-IRRI, the length and maximum width of individual green leaf blades on the main stem was measured manually with a ruler, then leaf area was estimated as length × maximum width × 0.725 (Tivet et al., 2001). The measurement was done at PI + 3 weeks (four sampled plants in GH IRRI), and at FLO (four sampled plants in GH-CNRS and two sampled

plants in the field) on the plants used for biomass measurement. The total area of all green leaves per plant was then measured by LI–3100C Area Meter in GH-CNRS and in the field.

Internode Profile and Anatomy
The length of each internode of the main tiller was measured at maturity in GH-CNRS and GH IRRI of sampled plants used for biomass measurement. In GH-CNRS, for anatomical observation purpose a middle part of 2 cm-long of peduncle and of internode-3 was sampled at FLO stage from the same plants used for biomass measurements. The samples were fixed in paraformaldehyde fixative solution and kept in desiccator overnight followed by dehydration with ethanol 70% for at least 24 h then conserved in the freezer prior to observation. The internodes were sliced into pieces of 60–80 μm with a Thermo Scientific Microm HM 650 V Vibration microtome. The pedunles were sliced into 75–90 μm after embedded in agarose 7% for 2 h with a Thermo Scientific Microm HM 650 V Vibration microtome.

In the field, internode profile was measured at FLO stage by sampling plants different from the plants used for biomass measurement. A 2 cm-long section of the peduncle and of the top internode was fixed in Formalin-Acetic-Alcohol (FAA) fixative solution until the date of sectioning. Prior to sectioning, the samples were dehydrate with ethanol 70% then sliced manually by razor blade.

The samples were then observed based on a high resolution imagery system (microscope Leica S8 APO equiped with camera QImaging MicroPublisher 3.3 in GH-CNRS and stereo microscope Olympus SZX7 equiped with camera Olympus DP71 in the field) to analyze the total and inner diameter as well as peripheric bundles number. In this study, peduncle is the uppermost internode between panicle neck node and node I, and top internode is the internode just below the peduncle, which is the same phytomer as flag leaf.

Carbon Assimilation Measurements
Actual assimilation rate (at homogenous level of light in the measurement chamber, i.e., 1500 or 1800 micromole $m^{-2} s^{-1}$), was measured by using Portable Photosynthesis System (GFS–3000 WALZ) during panicle development (2–3 weeks after PI), the same time as biomass measurement for all experiments. The last ligulated leaf of the main tiller from three tagged plants (in GH-CNRS) and two tagged plants (in the field and GH-IRRI) in every replication was chosen for this measurement.

Non Structural Carbohydrates (NSC) Analyses
NSC was analyzed during panicle development (2–3 weeks after PI) at the same time as C assimilation and biomass measurement in all experiments. In GH-CNRS, three plants per treatment (as three replications) were chosen homogenously for NSC analyses. In the field and GH-IRRI, two of sampled plants dedicated for biomass measurement were chosen for NSC analyses. In GH CNRS and GH IRRI trials, plants were dissected to sample the leaf blade of last ligulated leaf from the main stem. One base (only in GH CNRS) and one top internode (internode just below the peduncle) were sampled and immediately frozen in

liquid nitrogen and store at −80°C until fine ground with a ball grinder (Mixer mill MM301, Retsch, Germany). In the field, all green leaves and internodes of the main tiller were sampled and dried for 72 h in an oven at 70°C before fine ground with a grinder (Thomas-Wiley Laboratory Mill Model 4, Thomas Scientific USA).

The method used for sugar content analysis in GH CNRS and GH IRRI was based on high performance liquid chromatography (HPLC). The sugars were extracted three times from 20 mg ground samples with 1 ml of 80% ethanol for 30 min at 75°C, and then centrifuged for 10 min at 10,000 rpm. Soluble sugars (sucrose, glucose and fructose) were contained in the supernatant and starch in the sediment. The supernatant was filtered in the presence of polyvinyl polypyrrolidone and activated carbon to eliminate pigments and polyphenols. After evaporation of solute with Speedvac (RC 1022 and RCT 90, Jouan SA, Saint Herblain, France), soluble sugars were quantified by high performance ionic chromatography (HPIC, standard Dionex) with pulsated amperometric detection (HPAE-PAD). The sediment was solubilized with 0.02 N sodas at 90°C for 1 h 30 min and then hydrolyzed with α-amyloglucosidase at 50°C, pH 4.2 for 1 h 30 min. Starch was quantified as described by Boehringer (1984) with 5 μL of hexokinase (glucose-6-phosphate dehydrogenase), followed by spectro-photometry of NADPH at 340 nm (spectrophotometer UV/VIS V-530, Jasco Corporation, Tokyo, Japan).

In the field IRRI, sugars were extracted two times from 200 mg ground samples with 7 ml of 80% ethanol for 10 min at 80°C, and then centrifuged for 5 min at 3000 rpm. The residue was washed with 5 ml of hot 80% ethanol three times and combined all washings with the supernatant. The residue was dried at 70°C for 24 h prior to starch assay. Total soluble sugars was determined through colorimetric by adding 5 ml anthrone to 0.5 ml aliquot (sugar extraction was diluted with 80% ethanol) then boiled for 10 min at 100°C. After vortex mix and cooling on ice bath for about 5 min, followed by spectro-photometry at 620 nm (DU 800 UV/Vis spectrophotometer, Beckman Coulter). Dried residue was dropped with absolute ethanol and added with 2 ml of acetate buffer then boiled for 3 h at 100°C. The tubes were cooled to 55°C and proceed to hydrolysis step by adding 1 ml acetate buffer and 1 ml amyloglucosidase, then vortex mix. After incubated for 24 h at 37°C, the hydrolysate (supernatant layer) was decanted and save combined with the residue that had been washed with 3 ml of distilled water. Starch was determined through colorimetric assay by adding enzyme Peroxidase Glucose Oxidase (PGO) to 0.6 ml aliquot (starch hydrolysis was diluted with distilled water), then followed by spectro-photometry at 450 nm (DU 800 UV/Vis spectrophotometer, Beckman Coulter) after incubated in the dark room for 30 min.

Physiological Maturity and Yield Component
In GH IRRI four plants were harvested per replication and separated into panicles, green leaf blades, dead tissues and productive tillers then processed as two other experiments described in Adriani et al. (unpublished data). We determined yield components as panicle number per plant and plant FGDW.

Response Rate
The rate of trait response to either *qTSN* introgression or to a reduced plant access to light was quantified as: (ref_value–mod_value)/ref_value; where ref_value is the trait value for the NIL (in the case of response to *qTSN* introgression) or for low plant access to light treatment (shading in GH-CNRS and HD in field-IRRI, in the case of response to low light quantification) and mod_value is the trait value for the parent or for full light treatment, i.e., control in GH-CNRS and LD in field-IRRI). Response rates are synthesized in **Table 3**.

Data Analyses
The graphs describing plant morphogenesis, individual leaf area, grain production, relative NIL-P, and tillering rate relation to growth were represented with mean values and standard error (standard deviation divided by square root of the number of samples). Data of **Tables 2**, **3** and **Figures 3**, **4** were analyzed by an ANOVA procedure and mean comparisons between parent vs. NIL and between treatments for each pair of genotype were analyzed by Duncan's multiple range test using Microsoft® Excel 2010/XLSTAT-PRO statistical software (version 2014, Addinsoft, Inc., Brooklyn, NY, USA). SigmaPlot® Version 11.2 software (for Windows XP and below, copyright 2009–2010), Systat Software Inc. (Chicago, IL, USA) was used for plotting data and nonlinear regressions.

RESULTS

QTL Effects on Plant Morphogenesis
Under Full Light Conditions
Tiller dynamic is delayed in the presence of *qTSN*. This effect was observed at early tillering stage, i.e., starting at 400°C days (accumulation of thermal time from sowing) when eight leaves had appeared on the main stem. This was true for both genetic backgrounds and experiments (**Figures 1A,D**, **2A,D**; **Figures S1B**, **S2B**) and resulted in a reduced tiller number at PI that was however only significant in the field (*P* < 0.001, see **Table 2** for ANOVA). In IR64 background, this reduction rate ranged from 8.5 to 16.7% in GH-CNRS and the highest reduction rate was observed for NIL1 in the field (between 26.3 and 32.7%; **Table 3A**). In IRRI146 background, the reduction rate ranged from 16 to 25.9% (GH CNRS) and up to 29.5% in the field (**Table 3A**). This reduction was associated with a smaller rate of tiller abortion until MAT for the NILs compared to the parents, resulting in a progressive convergence of tiller number of parents and NILs at MAT (**Figures 1A,D**, **2A,D**; **Figure S1B**; **Tables 2**, **3**). Nevertheless, fertile tiller number in IRRI146 background in the field kept smaller at MAT in the NIL compared to the parent (22.3 and 19% of reduction under LD and HD, respectively) as well as for the IR64 parent in HD treatment (21.3 and 24.4% less tillers in NIL and NIL1, respectively, compared to the parent).

The effect of *qTSN* on leaf appearance rate and final leaf number on the main stem was in general weak and not homogenous across treatments. The *qTSN4* had a depressive effect on final leaf number on the main tiller in the field, although less pronounced in IRRI146 background (**Figures 2B,E**; **Figure S2A**). This could be related to a slightly lower rate of

TABLE 2 | ANOVA of flag leaf area, tiller number at PI (Panicle Initiation), FLO (flowering; heading in GH-CNRS), MAT (grain physiological maturity), biomass related traits (DW, Dry Weight) at plant and main stem level at FLO and MAT for vegetative DW and at MAT only for panicle DW and FGDW (FG, Filled Grain).

Source	Genetic background (G)	QTL	Treatment (T)	Replication	G × QTL	G × T	QTL × T
FLAG LEAF AREA							
GH-CNRS	<0.0001	0.0003	0.2902	0.9291	0.8243	0.5827	0.1783
Field	0.652	0.001	0.059	0.245	0.626	0.937	0.212
TILLER NUMBER AT PI							
GH-CNRS	0.026	0.055	0.266	0.235	0.721	0.922	0.627
Field	0.097	0.0005	<0.0001	0.066	0.455	0.084	0.054
TILLER NUMBER AT FLO							
GH-CNRS	0.937	0.187	0.812	0.973	0.479	0.812	0.581
Field	0.099	0.079	<0.0001	0.489	0.980	0.056	0.367
TILLER NUMBER AT MAT							
GH-CNRS	0.959	0.720	0.574	0.924	0.878	0.878	0.959
Field	0.001	0.021	<0.0001	0.038	0.186	0.008	0.584
MS DW FLO							
GH-CNRS	0.005	0.003	<0.0001	0.154	0.224	0.035	0.502
Field	0.002	<0.0001	<0.0001	0.257	0.192	0.870	0.022
MS DW MAT							
GH-CNRS	0.0002	0.002	0.327	0.807	0.014	0.002	0.303
Field	0.126	0.165	0.003	0.949	0.657	0.452	0.475
SHOOT DW FLO							
GH-CNRS	0.492	0.224	<0.0001	0.332	0.197	0.821	0.444
Field	0.0002	0.126	<0.0001	0.868	0.729	0.013	0.112
SHOOT DW MAT							
GH-CNRS	0.080	0.180	0.122	0.565	0.478	0.095	0.502
Field	0.002	0.096	<0.0001	0.113	0.302	0.010	0.769
MS PDW MAT							
GH-CNRS	0.772	0.0001	<0.0001	0.619	0.249	0.678	0.209
Field	0.041	0.004	0.001	0.410	0.258	0.166	0.539
PLANT FGDW MAT							
GH-CNRS	0.089	<0.0001	<0.0001	0.946	0.073	0.863	0.922
Field	0.048	0.083	<0.0001	0.140	0.155	0.078	0.443

leaf appearance considering the similar duration of the vegetative phase. In contrast, in GH-CNRS, higher leaf number (one more in average) in the presence of *qTSN4* was observed in IR64 background only, which was appreciable at the end of panicle development (about 1300°C days, at time of appearance of leaf 14; **Figure 1B**). This can be related to the fact that the vegetative phase was slightly longer in the NIL (later PI), of approximately one phyllochron, i.e., duration between the appearance of two consecutive leaves (**Figures 1B,E**; **Figure S1A**).

Similarly to that observed with leaf number, *qTSN* effect on stem length (considered here as the successive internodes excluding the peduncle) also differed from GH-CNRS to field-IRRI trials as well as between genetic backgrounds. An appreciable decrease in final stem length was observed in IR64 background in the presence of *qTSN* in GH-CNRS (**Figure 1C**), whereas in the field no *qTSN* effect on stem length was observed (**Figure 2C**; **Figure S2C**). No clear difference was observed in IRRI146 background (**Figures 1F, 2F**; **Figure S1C**). The effect of *qTSN* on peduncle length differed with respect to the genetic

background and the experiment. In GH-CNRS, the peduncle (the internode bearing the panicle) was significantly longer in the presence of *qTSN4* in IR64 background (not presented) but it was the opposite in IRRI146 background (**Figure S1E**), which was confirmed in GH-IRRI (data not presented). In the field, the peduncle was shorter in the presence of *qTSN12* in IR64 background (**Figure S2E**), whereas no effect was observed in IRRI146 background (not presented). In IRRI146 background in GH-CNRS, no significant difference in stem length was observed between parent and NIL until heading (**Figure S1C**), but thereafter *qTSN4* positively affected the length of the top three internodes located just below the peduncle. The peduncle was, however, shorter in the NIL compared to the parent (**Figure S1E**).

More stable effect of *qTSN* could be observed on peduncle anatomy and thickness for both genetic backgrounds and trials. The increase of peduncle thickness was 44% in IRRI146 background in GH-CNRS (**Figure S1F**), and 14% in IR64 background (not presented). In the field, the increase of peduncle

TABLE 3 | Response rate of traits to QTLs introgression in each pair of isoline in a given treatment (in the field-IRRI, LD is for Low Density, HD is for high density) and genetic background (IR64 and IRRI146) (A), to access to light in each trial (field-IRRI, GH-CNRS) (B), for each genotype (parent, NIL, NIL1).

A

Traits	IR64-QTL effect		IRRI146-QTL effect		IR64-QTL effect (NIL)		IR64-QTL effect (NIL1)		IRRI146-QTL effect	
	Control	Shading	Control	Shading	LD	HD	LD	HD	LD	HD
Flag leaf area	0.530	0.181	1.091	0.478	0.265	0.156	0.339	0.046	0.291	0.231
Tiller number at PI	−0.167	−0.085	−0.259	−0.160	−0.196	−0.183	−0.263	−0.327	−0.295	−0.161
Tiller number at FLO	−0.014	0.076	−0.147	−0.137	−0.124	−0.069	−0.292	−0.240	−0.072	−0.192
Tiller number at MAT	0.000	0.026	0.041	0.030	−0.052	−0.213	−0.069	−0.244	−0.223	−0.190
MS DW FLO	0.168	0.255	0.126	0.060	0.410	0.579	0.608	0.454	0.358	0.118
MS DW MAT	0.177	0.600	0.082	0.030	0.204	0.018	0.121	0.051	0.290	0.121
Shoot DW FLO	0.030	0.281	0.010	−0.020	0.037	0.068	0.128	0.103	0.133	−0.147
Shoot DW MAT	0.068	0.155	−0.004	0.080	−0.023	−0.179	0.025	−0.040	−0.141	−0.143
MS PDW MAT	0.346	1.158	0.275	0.326	0.058	−0.022	0.271	0.507	0.386	0.129
Plant FGDW MAT	0.310	0.948	0.195	0.198	−0.022	−0.097	0.131	0.075	−0.210	−0.232

B

Traits	IR64-light effect		IRRI146-light effect		IR64-density effect			IRRI146-density effect	
	Parent	NIL	Parent	NIL	Parent	NIL	NIL1	Parent	NIL
Flag leaf area	0.300	0.004	0.328	−0.061	0.024	−0.064	−0.200	−0.057	−0.101
Tiller number at PI					−0.635	−0.629	−0.667	−0.717	−0.664
Tiller number at FLO	−0.096	−0.014	0.176	0.190	−0.659	−0.637	−0.634	−0.540	−0.600
Tiller number at MAT	0.013	0.040	0.050	0.039	−0.588	−0.656	−0.665	−0.694	−0.681
MS DW FLO	−0.429	−0.386	−0.300	−0.338	−0.281	−0.195	−0.349	−0.157	−0.305
MS DW MAT	−0.326	−0.084	0.190	0.138	−0.091	−0.232	−0.148	−0.210	−0.313
Shoot DW FLO	−0.331	−0.168	−0.230	−0.250	−0.708	−0.699	−0.714	−0.674	−0.754
Shoot DW MAT	−0.184	−0.118	−0.030	0.046	−0.678	−0.729	−0.698	−0.735	−0.735
MS PDW MAT	−0.607	−0.369	−0.499	−0.479	−0.319	−0.370	−0.192	−0.048	−0.224
Plant FGDW MAT	−0.589	−0.388	−0.454	−0.453	−0.733	−0.754	−0.746	−0.772	−0.778

Gray columns indicate no treatment effect as shading in GH-CNRS just imposed at PI.

thickness in the presence of *qTSN* was 20% in IR64 background (**Figure S2F**) and 17% in IRRI146 background (not presented). The characteristics of the peduncle were associated with thicker top internode (in the third internode below the peduncle in GH-CNRS, **Figure S1F**; in top internode in the field, **Figure S2F**) and higher number of vascular bundles in the peduncle (data not presented) in the NILs compared to the parents.

A positive effect of *qTSN* on leaf area was observed for the flag leaf (FL) (significant in IR64 background for both trials) and the two to three leaves below the flag leaf (FL-3 or FL-2), but it was more pronounced for FL (**Figure 3**). In GH-CNRS, the increase was 53% for IR64 background (**Figure 3A**; **Table 3A**) and 109% in IRRI146 background (**Figure 3B**; **Figure S1D**; **Table 3A**), which was mainly explained by an increase of leaf length, whereas the width was not affected (data not presented). In the field, *qTSN* effect on the leaf area of the top leaves was already expressed from FL-3 upward for both genetic backgrounds (and significant for all these leaves only in NIL1) (**Figures 3C,D**), with 34% of increase in IR64 background (**Figure S2D**; **Table 3A** for QTL effect in NIL1) and 29% of increase in IRRI146 (**Table 3A**). In GH IRRI it was expressed and significant from FL-2 upward (**Figure 3E**). In the field and GH IRRI, the positive effect of *qTSN* on individual

leaf size was mainly supported by the width (data not presented) rather than the length.

Under Low Light Conditions

Traits related to plant growth and development were more affected by the treatment in the field. This can be easily explained by the fact that treatments were established by planting density from sowing onwards in the field, while shading treatment in GH-CNRS was imposed only from PI time to heading. Accordingly, early tillering (**Figures 1A,D**, **2A,D**; **Tables 2**, **3B** for the rate of trait plasticity in response to light treatment) and leaf appearance rates (**Figures 1B,E**, **2B,E**) were poorly affected by shading in GH-CNRS. By contrast stem length was more affected by shading in GH-CNRS, i.e., decreased, as observed in field-IRRI but only in HD treatment in IRRI146 isolines.

Peduncle and internode thickness were also reduced under low access to light in both experiments but only in the NILs, whereas they were not modified in the parents (not presented). Individual leaf size was almost not affected by the reduction of incoming light. In GH-CNRS, main tiller FL size of that parents of both backgrounds increased by about 30% under shading, whereas under the same conditions it was maintained or reduced
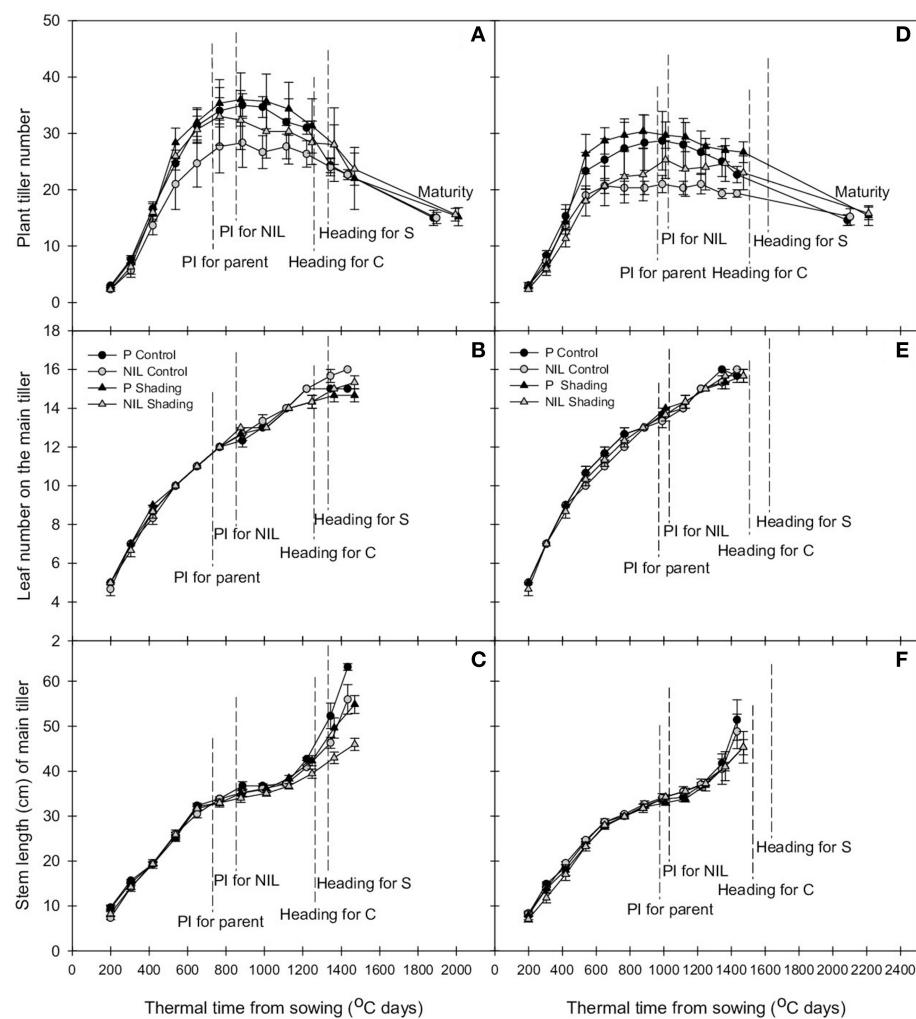
FIGURE 1 | Change with thermal time of plant tiller number (A,D), leaf number on the main tiller (B,E), and stem length of the main tiller (C,E) of parent (black) and NIL (gray) in IR64 (A–C) and IRRI146 (D–F) background, under control (C) and shading (S) in GH-CNRS trial. The values are mean ± SE. *n* = 3.

with the NILs. In the field, slight increase in FL area due to high density was observed only for the parent of IR64 background (2%) (**Figures 3A,D**; **Table 3B**).

As observed under full light conditions, tiller number was reduced with *qTSN* under low plant access to light (shading in GH-CNRS, HD in the field), however, the difference between NIL and parents was not as strong as observed under higher plant access to light (**Figures 1A,D**, **2A,D**; **Table 3B**), and final tiller number was similar between NILs and parents at maturity. The *qTSN* effect on leaf appearance (positive for IR64 background in GH-CNRS; negative in the field; unchanged for IRRI146 background in both trials) and stem elongation (negative or unchanged) was similar than that observed in full light conditions (**Figures 1**, **2**). The *qTSN* positive effect on individual leaf size under low plant access to light was appreciable but less pronounced than that observed under high access to light, as the low access to light increased top leaf size of the parents but not of the NILs (**Figure 3**; **Table 3B**). No QTL effect was observed on

the traits related to peduncle and internode anatomy under low light condition (not presented).

Biomass, Leaf Area and Grain Productions

At plant level, in GH-CNRS, *qTSN4* increased plant final grain production (FGDW) which was significant in all cases (**Table 2**) except for IRRI146 background under shading (**Figure 4B**; **Table 3A**). Meanwhile, plant shoot biomass was not affected by *qTSN4* neither at PI and FLO (heading in GH-CNRS) (not presented) nor at physiological maturity (**Figure 4C**; **Tables 2**, **3A**). In the field, *qTSN* poorly affected grain and straw biomass and not systematically in a positive way. An increase could be observed only in IR64 background in plant grain production but this was not significant. The *qTSN* effect was even significantly negative on plant grain production in IRRI146 under LD (**Figure 4E**; **Table 3A**). These contradictory results between GH-CNRS and field conditions regarding *qTSN4* effect on grain production at plant level were also observed in GH-IRRI where
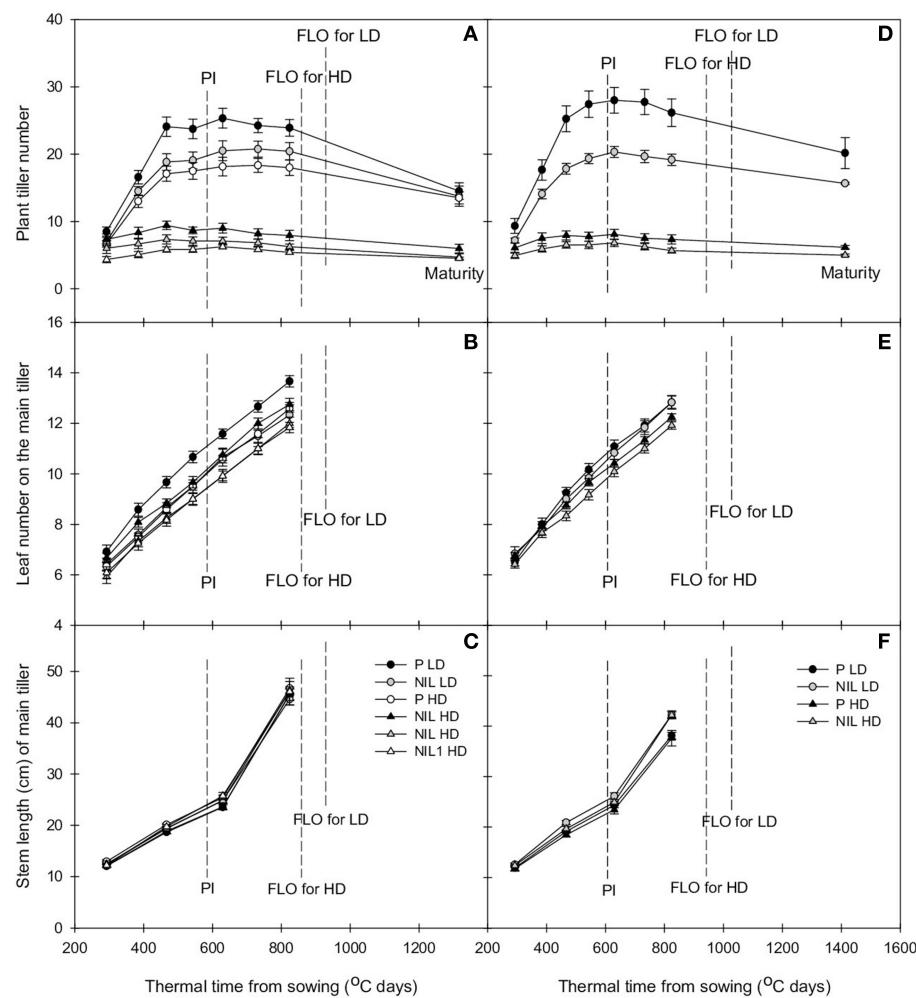
FIGURE 2 | Change with thermal time of plant tiller number (A,D), leaf number on the main tiller (B,E), and stem length of the main tiller (C,E) of parent (black) and NIL (gray) in IR64 (A–C) and IRRI146 (D–F) background, under low density (LD) and high density (HD) in field trial. The values are mean ± SE. *n* = 4.

no significant QTL effect on plant grain production was observed (results not presented). The leaf area per plant at flowering time was unaffected by *qTSN* but the distribution of leaf area was modified in a way that individual leaves were larger but fewer in the presence of *qTSN*. This was true for both genetic backgrounds in both trials (**Figures 4D,H**). However, plant leaf area was affected by light conditions in the field, where it was higher in full light (LD) compared to low light (HD) condition.

At main stem level, panicle dry weight at MAT was increased by *qTSN* in both genetic backgrounds and trials (**Figures 4B,F**; **Table 2**; *P* < 0.01). This increase was more pronounced and systematic in GH-CNRS. It was generally higher under low light conditions for IR64 background (115% under shading and 51% under HD in NIL1, but no effect observed in the NIL of IR64 in field-IRRI, **Table 3A**), whereas it was more homogenous among treatments for IRRI146 (**Table 3A**). The main stem panicle DW was systematically reduced by low access to light and no *qTSN* × treatment interactions were observed (**Tables 2**, **3B**).

The main stem DW at FLO was systematically increased by *qTSN* (*P* < 0.01, **Table 2**), and this was generally stronger in IR64 background compared to IRRI146 and more particularly in field-IRRI. This *qTSN* effect was not maintained until MAT as no significant difference for main stem DW between parents and NIL could be observed at that stage (**Tables 2**, **3**). Low plant access to light systematically reduced main stem DW at FLO, more particularly in GH-CNRS (**Table 3**). This was maintained until maturity only in the field-IRRI as no more significant treatment effect was observed in the GH-CNRS at this time (**Tables 2**, **3**). No *qTSN* × treatment interaction for main stem DW was observed, neither at MAT or FLO, except in field-IRRI at FLO (**Table 2**).

Relationship Between Tillering Dynamics and Main Stem Growth Rate

Overall, above-mentioned results pointed out two key nodes of regulation of plant phenotypes due to *qTSN* introgression,
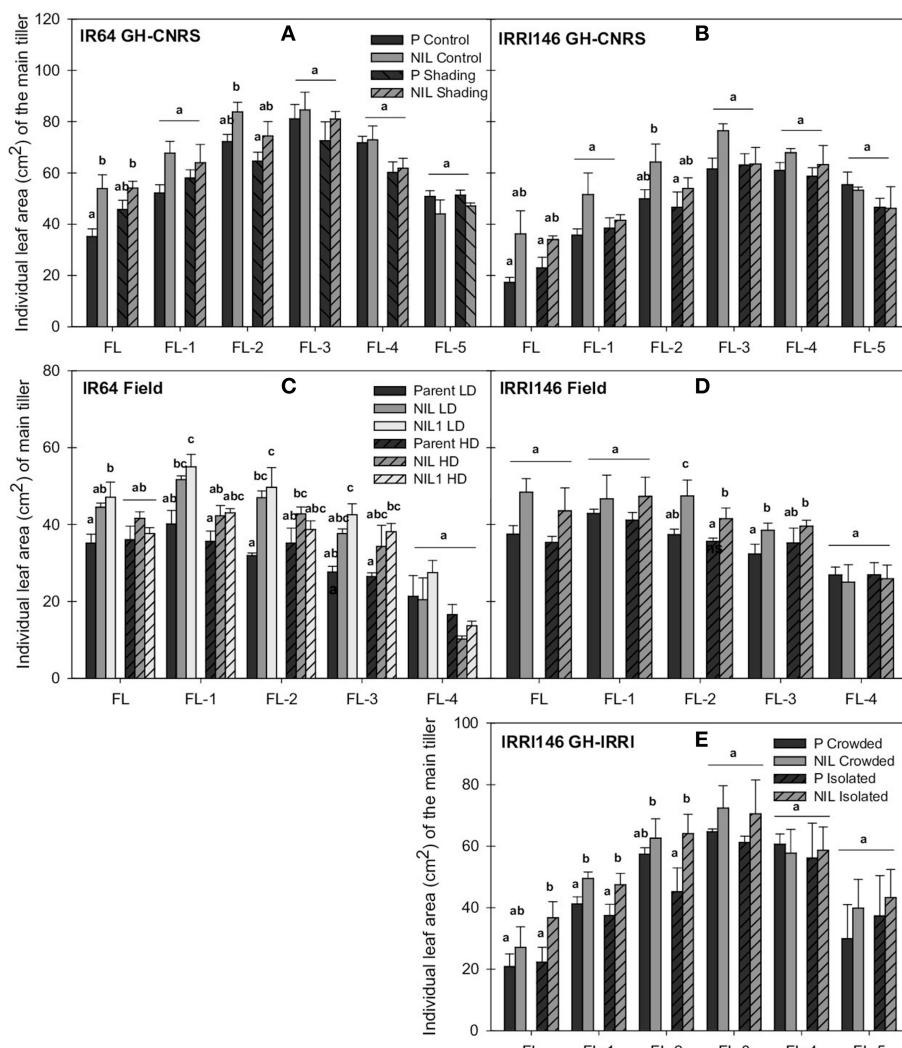
FIGURE 3 | Individual leaf area of main tiller (FL, flag leaf; FL-1–FL-5, one to five leaves below the flag leaf) at flowering (FLO) of parent (black), NIL (gray), and NIL1 (white) in IR64 background (A) and IRRI146 background (B) in GH-CNRS under control and shading, in IR64 background (C) and IRRI146 background (D) in the field under low density (LD) and high density (HD), in IRRI146 background (E) in GH-IRRI under crowded and isolated plants. The values are mean ± SE. Results of Duncan test for multiple comparisons of each genotype per treatment at 5% level are shown in the letters above the bars. $n = 4$ in GH-CNRS and filed-IRRI, $n = 3$ in GH-IRRI.

namely: (i) tillering and tiller number, generally reduced by *qTSN* and (ii) main stem biomass (either before or after FLO), generally increased by *qTSN* (**Table 3A**). In association with the opposite effect of *qTSN4* on these two traits, no additional difference was observed between parents and NILs regarding the resulting plant shoot DW at FLO and MAT. Results of plant FGDW were, however, dependent on cropping conditions (**Figures 4A,D**; **Table 3A**). In order to further explore the relationship between tillering and main stem DW, the change of main stem growth rate during panicle development (PI–FLO period) was plotted against tillering rate before PI. A negative correlation could be observed between these variables, stronger in IRRI146 ($R^2 = 0.5$) than in IR64 ($R^2 = 0.28$) backgrounds (**Figures 5A,B**). Interestingly, the average value for

the NIL was systematically at a higher position on the y-axis (with reference to main stem growth rate from PI to FLO) compared to that of the parents. However, the correlation disappeared when analyzing the same relationship in each experiment separately (**Figures 5C,D**).

In order to evaluate whether this early trade-off between tillering and main stem growth rate (from PI to FLO) could impact grain production, main stem growth rate from PI to FLO was plotted against main stem panicle DW at maturity. This is presented in **Figures 6A,B**, showing a slightly positive correlation between these variables ($R^2 = 0.12$) when analyzing data from parents and NILs together, whereas there was no correlation in IRRI146 ($R^2 = 0.02$) background. This positive correlation was getting even stronger when considering trials separately
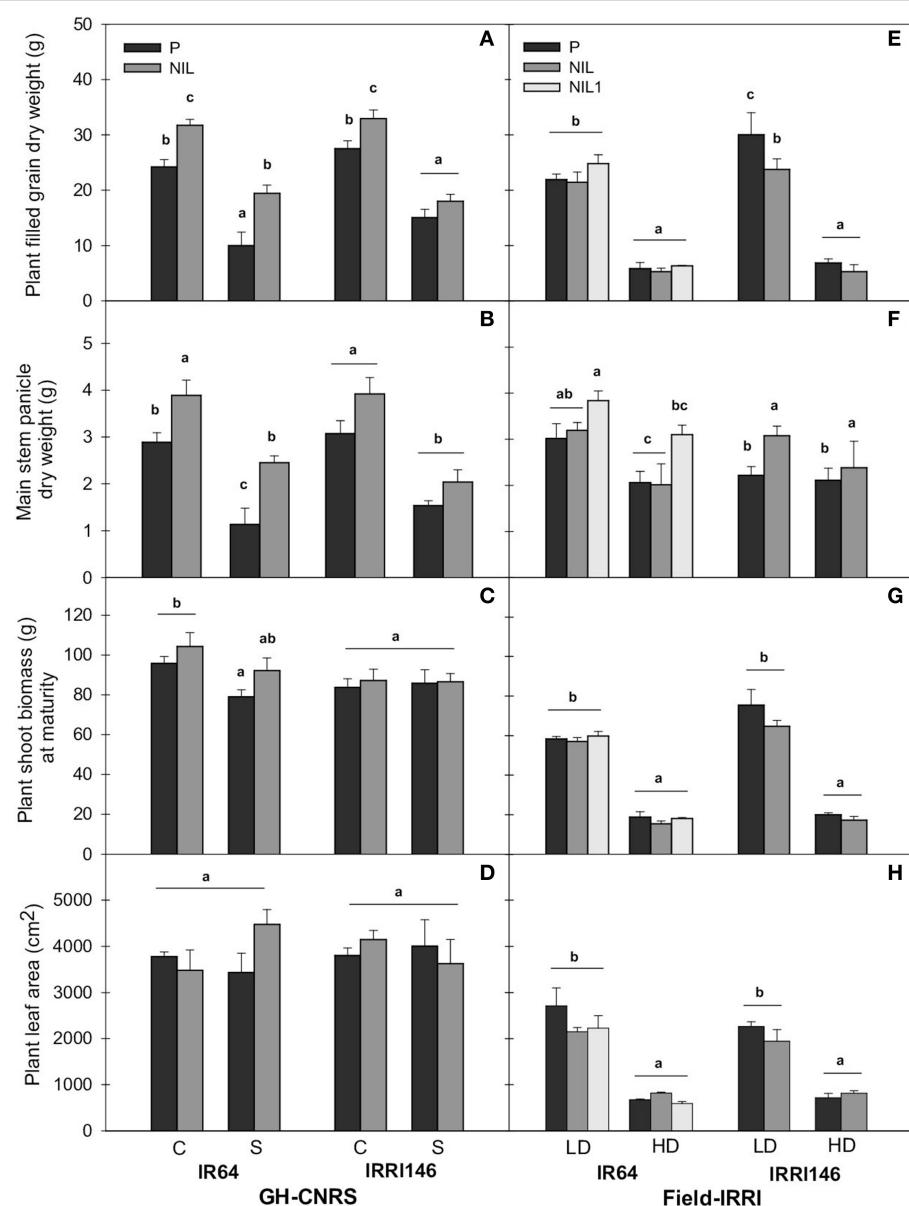
FIGURE 4 | Plant filled grain dry weight (A,E), main stem panicle dry weight at maturity (B,F), plant shoot biomass at maturity (C,G), plant green leaf area at flowering (D,H) under control (C), and shading (S) in GH-CNRS (A–D), under low density (LD), and high density (HD) in field-IRRI (E–J). The values are mean ± SE. Results of Duncan test for multiple comparisons of each genotype per treatment at 5% level are shown by the letters above error bars. $n = 5$ at maturity and $n = 4$ at flowering in GH-CNRS, $n = 4$ in field-IRRI.

(**Figures 6C,D**), in particular at GH-CNRS. In all situations, the average of NIL values showed a higher main stem growth rate from PI to FLO related to a bigger panicle DW on the main stem at maturity.

The relationship between main stem and plant shoot growth rates from PI to FLO was thereafter explored, and no significant correlation was observed (not presented). Meanwhile, plant grain production was positively correlated to the whole plant shoot growth rate from PI to FLO (**Figure 7**). This correlation was higher in IR64 ($R^2 = 0.49$) than in IRRI146

($R^2 = 0.37$) background and in field-IRRI ($R^2 = 0.50$) than in GH-CNRS ($R^2 = 0.28$) trial. Interestingly, with respect to this correlation, the NILs performed better than parents only in GH-CNRS whatever the genetic background (**Figure 7C**), i.e., in the cropping situation where the trade-off between tillering and main stem growth rates was the lowest for the NIL and the nearest from that of parents (**Figure 5C**). It can be mentioned that plant grain production was not correlated to main stem growth rate from PI to FLO (**Figure S3**).
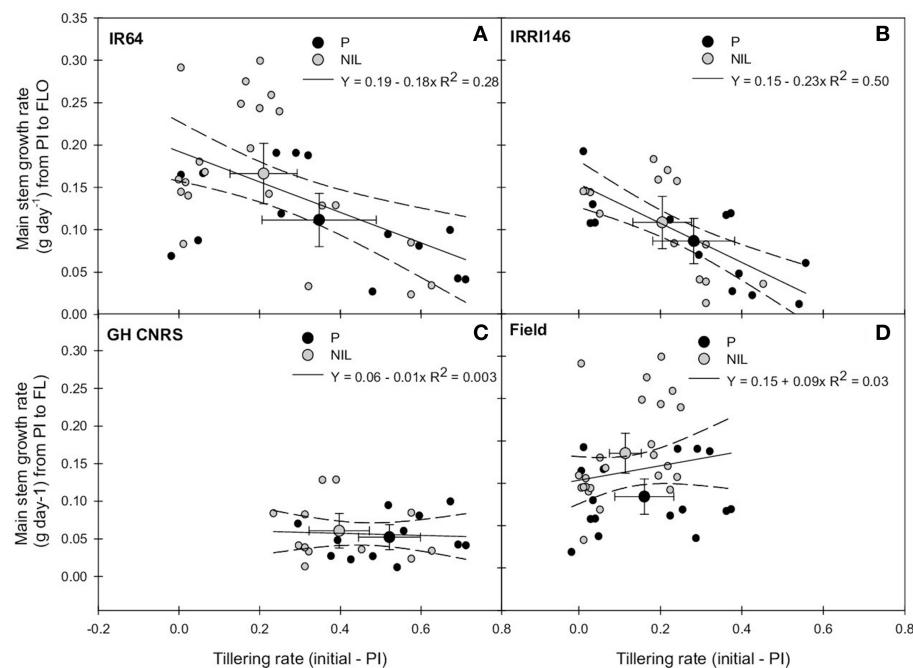
FIGURE 5 | Relationship between tillering rate from initial measurement to panicle initiation (PI) and main stem growth rate from PI to flowering (FLO) of the parent (black symbol) and the NIL (gray symbol), in IR64 background in GH-CNRS and field trials (A) in IRRI146 background in GH-CNRS and field trials (B) in GH-CNRS in IR64 and IRRI146 backgrounds (C) in field-IRRI experiment in IR64 and IRRI146 backgrounds (D). The values are mean ± SE. Regression curves are associated with confidence interval at $P = 0.05$. $n = 40$ for IR64 background and field-IRRI trial, $n = 32$ for IRRI146 background and GH-CNRS trial for regression curve.

Carbon Assimilation and Sugar Related Traits

In order to identify whether the difference in main stem growth rate was associated with a particular metabolic pattern, starch and net assimilation rate at ambient CO_2 concentration of 400 ppm were quantified during panicle development. In GH-CNRS, *qTSN4* enhanced assimilation only in shading treatment by 33% for IR64 and 24% for IRRI146 background (**Figure 8**). Shading significantly reduced assimilation in parent lines, 29 and 16% for IR64 and IRRI146 background, respectively, whereas in the NILs, assimilation was maintained under shading (**Figure 8A**). In the field, no significant *qTSN* and treatment effect was observed on assimilation, even if it was slightly decreased with *qTSN* in both backgrounds (**Figure 8B**). In GH-IRRI, *qTSN4* increased assimilation under low light (crowded plants) conditions but it was the opposite under full light (isolated plants) conditions (**Figure S4A**).

Internode starch content increased in the presence of *qTSN* for both backgrounds across the trials (**Figures 8C,D**; **Figure S4B**). The *qTSN4* effect was stronger under full light conditions (control in GH-CNRS, LD in the field, isolated in GH-IRRI), and significant in IRRI146 background in GH-CNRS (18%) (**Figure 8C**) and IR64 background in the field (11%) (**Figure 8D**). Similar trend was observed in leaf blade, with no significant effect of *qTSN* across the trials (not presented). However, in most cases, leaf starch was strongly reduced under low plant access to light.

DISCUSSION

The isolines (NIL) used in this study carried *qTSN4* or *qTSN12*, known to enhance leaf and panicle sizes but to reduce panicle number in some environmental situations (Fujita et al., 2013; Okami et al., 2015). The present study aimed at better characterizing this trade-off by comparing the NILs to their recurrent parents IRRI146 and IR64 regarding morphogenesis and C source-sink balance along the whole plant cycle and their behavior under low access to light.

The QTSN Affects Rice Morphogenesis and Physiology at Earlier Stage than Expected

A reduction of the rate of tiller emergence before PI (as early as 400°C days) was observed in this study in the presence of *qTSN* for both genetic backgrounds under both treatments. This was, however, not addressed in previous studies (Fujita et al., 2012, 2013) where the breeders mainly focused their attention on latter traits measured between flowering and maturity. Nevertheless, Okami et al. (2015) confirmed that tiller number was reduced with *qTSN4* under drought stress at vegetative stage as tillering rate per unit of above-ground biomass of the NIL was lower than that of the parent, which is in line with the present study. But in contrast to the present study, no difference between P and NIL was observed regarding tillering dynamics under well water conditions (Okami et al., 2015) and the ratio of main stem
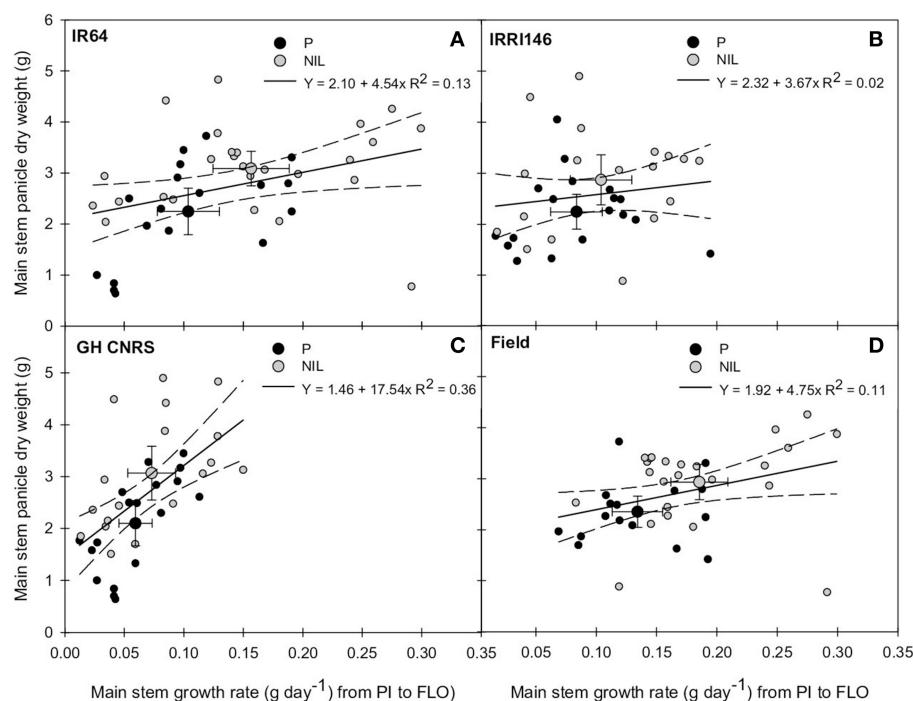
FIGURE 6 | Relationship between main stem growth rate from panicle initiation (PI) to flowering (FLO) and main stem panicle dry weight at maturity of the parent (black symbol) and the NIL (gray symbol), in IR64 background in GH-CNRS and field trials (A) in IRRI146 background in GH-CNRS and field trials (B) in GH-CNRS in IR64 and IRRI146 backgrounds (C) in field-IRRI experiment in IR64 and IRRI146 backgrounds (D). The values are mean ± SE. Regression curves are associated with confidence interval at $P = 0.05$. $n = 42$ for IR64 background and field-IRRI trial, $n = 34$ for IRRI146 background and GH-CNRS trial for regression curve.

leaf area to tiller number (Okami et al., 2012). Interestingly, in the present study, main stem growth rate between PI and FLO was inversely proportional to tillering rate before PI (**Figure 5**). Considering that reduction in tillering rate is expected to provide more assimilate available to the growing stems, the difference in main stem growth rate appears as a consequence of the change in early tillering rate. Several hypotheses subtending this correlation can be raised. On one hand, it can be hypothesized that *qTSN* implies a higher apical dominance due to hormonal signals, e.g., in relation to strigolactone (Jamil et al., 2012) or other hormones (ABA, IAA, GA3; Liu et al., 2011). This may be associated with a higher sensitivity to the red/far red ratio within the canopy maybe brought by *qTSN*. Indeed cessation of tillering in crops has been widely reported to be correlated with the increase, with crop age, in red/far-red ratio sensed by the plant within the canopy, even before any C limitation occurred within the plant (Ballaré and Casal, 2000; Ugarte et al., 2010). On the other hand, sink strength of growing stems should be higher at early stage in the presence of QTL, due to the initiation and pre-dimensioning of larger organs at meristem level, which may decrease tiller bud outgrowth. In this latter case, tillering would be reduced by a competition (or at least a signaling of competition) for resources active meristem and organs and tiller bud outgrowth, as pointed out by Rebolledo et al. (2012) across japonica rice genetic diversity. Whatever the hypothesis, the reduction in early tillering rate coincides with the appearance of leaf 8 in GH (rather leaf 7 in the field) so

with the initiation of leaf 10 or 11, based on the developmental pattern established by Nemoto et al. (1995). Interestingly in GH, the leaves with larger area in the presence of QTL were the 3–4 top leaves of the main stem, so those initiated right after leaf 11.

The size of organs was enhanced by *qTSN*4 on the main stem, from 3 to 4 leaves below the flag leaf up to the panicle. In addition to individual leaf area from the upper phytomers, internode length, and internode and peduncle thickness, were also enhanced with *qTSN*, while peduncle length and plant height were reduced. This finding was also reported in a recent study using the same genetic materials but comparing genotype behavior under drought and well-watered conditions (Okami et al., 2015). This is in line with Fujita et al. (2013) using the same genetic materials and Wu et al. (2011) using other genetic materials revealing that rice plants with larger culm diameter exhibited longer and wider flag leaf, more grains per panicle, as well as lower tiller number. Furthermore, Liu et al. (2008) confirmed that a thicker peduncle plays an important role in the determination of panicle size and grain yield potential.

This behavior in the presence of *qTSN* may be thus related to a stronger expression of apical dominance as the changes in organ size could not be detected at plant level: plant leaf area and shoot biomass at flowering were not different between parents and NILs, mainly due to the compensation between organ size and number. Nevertheless, the increase in C assimilation (in GH
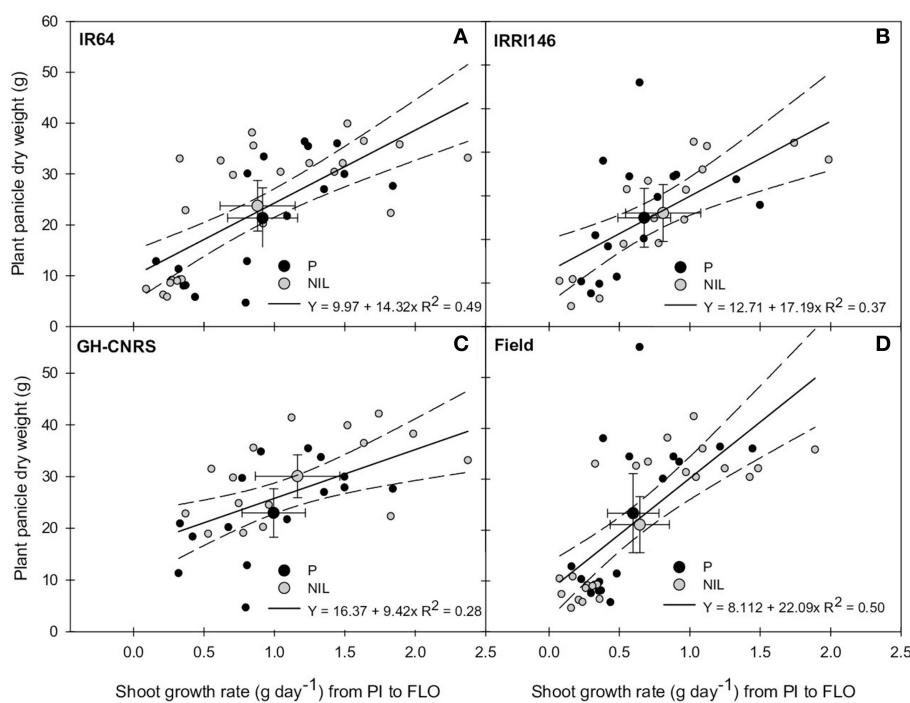
FIGURE 7 | Relationship between plant shoot growth rate from panicle initiation (PI) to flowering (FLO) and plant panicle dry weight at maturity of the parent (black symbol) and the NIL (gray symbol), in IR64 background in GH-CNRS and field trials (A) in IRRI146 background in GH-CNRS and field trials (B) in GH-CNRS in IR64 and IRRI146 backgrounds (C) in field-IRRI experiment in IR64 and IRRI146 backgrounds (D). The values are mean ± SE. Regression curves are associated with confidence interval at $P = 0.05$. $n = 42$ for IR64 background and field-IRRI trial, $n = 34$ for IRRI146 background and GH-CNRS trial for regression curve.

under shading) and in sugar (in particular starch) storage in stems (only in GH) suggested that *qTSN* enhances C availability within the main tiller at least during panicle development. Whether this is related to leaf anatomy and the elaboration of plant leaf area based on larger, thicker but fewer leaves needs to be confirmed. The fact that *qTSN4* co-localizes with NAL1, a gene involved in leaf anatomy, veining pattern and carboxylation, provides further insight to this *QTL* (Qi et al., 2008).

The QTSN Early Trade-Off Between Tiller and Main Stem Growth Partially Explains its Environment Dependent Effect on Plant Grain Production

The reduction of early tillering and its benefits for main stem growth rate in the presence of *qTSN* was dependent on the genetic background and the environment: it was more pronounced in GH-CNRS and in IR64. In addition, the positive effect of *qTSN* on plant grain production was mostly revealed in GH-CNRS while it was weak or inexistent in the field. This can be explained by the fact that in GH-CNRS fertile tiller (i.e., panicle) number was less reduced by *qTSN4* compared to that observed in the field, while its effect on panicle size was strong. The low effect in the field cannot be totally interpreted, but it is in line with a previous study (Okami et al., 2014) that revealed the absence

of difference in grain yield between parents and NILs of IR64 background during three summer seasons under flooded and aerobic conditions in Japan. The same authors, however, reported some differences under aerobic conditions in one season with low nitrogen supply (90 kg N ha^{-1}). These results support the present study suggesting that a positive effect of the *qTSN* may be expressed depending on the cropping environment, where stressing environments (low light, low N, low soil exploration as in pots in greenhouse) should favor the benefits of *qTSN* on panicle size and, if ever, on grain production.

Earlier Tiller Cessation and its Implication for Yield Potential

Increasing yield potential by inducing an earlier cessation of tiller production was already proposed in previous studies comparing performances of hybrids and inbreds (Bueno and Lafarge, 2009; Lafarge et al., 2010). An earlier cessation of tillering was then related to an earlier biomass accumulation within reproductive stems and to a higher biomass remobilization from internodes to panicles, as also promoted by a higher sensitivity to the red/far red ratio within the canopy (Ballaré and Casal, 2000). Interestingly, the present study also highlighted the correlation between an earlier tillering cessation and higher main stem biomass growth, in association with a larger sink size. In the present study, however, these correlations were weak at plant level in the field particularly, highlighting the complexity
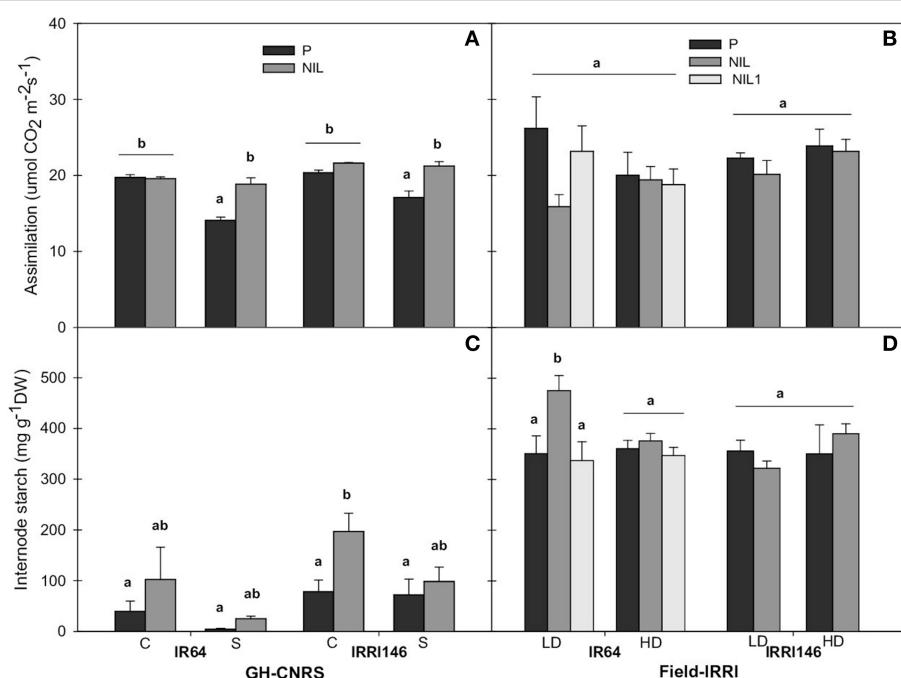
FIGURE 8 | Carbon assimilation (A,B) and internode starch concentration during panicle development (C,D) of parent (black), NIL (gray) and NIL1 (white) under control (C) and shading (S) in GH-CNRS (A,C), under low density (LD) and high density (HD) in field-IRRI (B,D). The values are mean ± SE. Results of Duncan test for multiple comparisons of each genotype per treatment at 5% level are shown in the letters above the bars. $n = 3$ in GH-CNRS, $n = 4$ in field-IRRI.

of the GxE interactions and trade-offs underlying the *qTSN* effect on plant grain production. It will be interesting to pay more attention on other fertile tillers to better understand *qTSN* impact on the whole plant growth and grain production. Nevertheless, this study reinforces the interest of developing genotypes optimizing tillering dynamics as long as yield potential is concerned.

CONCLUSION

The *qTSN* was confirmed in this study as a QTL potentially increasing panicle size due to an increase in stem growth rate, and in the size of the top leaves and internodes, at least at the main stem level. However, the trade-off between panicle size and panicle number was identified as the key node modulating the environment-dependent *qTSN* positive effect on plant grain production. This study revealed indeed that this trade-off was already visible at early stage through an earlier cessation of tiller production due to *qTSN* introgression, which coincided with the initiation at meristem level of phytomers with potentially larger leaves and internodes. Although it cannot be concluded if this early effect impacts directly tillering or organ dimensioning at meristem level, it seems worth going deeper in the understanding of the physiological regulation of this allele on plant functioning including contrasted cropping condition like limited radiation (i.e., during wet season in the tropics or N availability) in further studies.

AUTHOR CONTRIBUTIONS

TL, SY, MD, and DL participated in the design of the study; DA, AD participated in performing the research; AF and AC participated in biochemistry analysis; DA, TL, and DL participated in data analysis and wrote the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.01197

Figure S1 | The relative (NIL-P) of IRRI146 background in GH-CNRS under control. Leaf number on the main tiller **(A)** Tiller number per plant **(B)** Stem length of the main tiller **(C)** Individual leaf area at flowering and spikelet number per panicle at maturity **(D)** Peduncle and internode length at maturity **(E)** Peduncle and internode thickness at maturity **(F)**.

Figure S2 | The relative (NIL1-P) of IR64 background in field-IRRI under low density. Leaf number on the main tiller **(A)** Tiller number per plant **(B)** Stem length of the main tiller **(C)** Individual leaf area at flowering and spikelet number per panicle at maturity **(D)** Peduncle and internode length at maturity **(E)** Peduncle and internode thickness at maturity **(F)**.

Figure S3 | Relationship between plant shoot growth rate from panicle initiation (PI) to flowering (FLO) and plant grain dry weight at maturity of the parent (black symbol) and the NIL (gray symbol), in IR64 background in GH-CNRS and field trials **(A)** in IRRI146 background in GH-CNRS and field trials **(B)** in GH-CNRS in IR64 and IRRI146 backgrounds **(C)** in field-IRRI experiment in IR64 and IRRI146 backgrounds **(D)**. The values are mean ± SE. Regression curves are associated with confidence interval at $P = 0.05$. $n = 42$ for IR64 background and field-IRRI trial, $n = 34$ for IRRI146 background and GH-CNRS trial for regression curve.

Figure S4 | Carbon assimilation (A) and internode starch concentration during panicle development (B) of parent (black), NIL (gray) under isolated and crowded population in GH-IRRI. The values are mean ± SE. Results of Duncan test for multiple comparisons of each genotype per treatment at 5% level are shown in the letters above the bars. $n = 3$.

REFERENCES

Ashikari, M., Sakakibara, H., Lin, S. Y., Yamamoto, T., Takashi, T., Nishimura, A., et al. (2005). Cytokinin oxidase regulates rice grain production. *Science* 309, 741–745. doi: 10.1126/science.1113373

Ballaré, C. L., and Casal, J. J. (2000). Light signals perceived by crop and weed plants. *Field Crops Res.* 67, 149–160. doi: 10.1016/S0378-4290(00)00090-3

Boehringer, S. A. (1984). *Methods of Enzymatic Food Analysis using Single Reagents.* Mannheim: Boehringer Mannheim GmbH.

Borràs-Gelonch, G., Rebetzke, G. J., Richards, R. A., and Romagosa, I. (2012). Genetic control of duration of pre-anthesis phases in wheat (*Triticum aestivum* L.) and relationships to leaf appearance, tillering, and dry matter accumulation. *J. Exp. Bot.* 63, 69–89. doi: 10.1093/jxb/err230

Brennan, J. P., and Malabayabas, A. (2011). "Impacts of IRRI germplasm on the Philippines since 1985," in *International Rice Research Institute's Contribution to Rice Varietal Yield Improvement in South-East Asia* (Canberra, ACT: Australian Center for International Agricultural Research), 25–43. Impact Assesment Series Report. 74.

Bueno, C. S., and Lafarge, T. (2009). Higher crop performance of rice hybrids than of elite inbreds in the tropics: 1. hybrids accumulate more biomass during each phenological phase. *Field Crops Res.* 112, 229–237. doi: 10.1016/j.fcr.2009.03.006

Chen, S., Gao, R., Wang, H., Wen, M., Xiao, J., Bian, N., et al. (2015). Characterization of a novel reduced height gene (Rht23) regulating panicle morphology and plant architecture in bread wheat. *Euphytica* 203, 583–594. doi: 10.1007/s10681-014-1275-1

Chen, S., Zeng, F. R., Pao, Z. Z., and Zhang, G. P. (2008). Characterization of high-yield performance as affected by genotype and environment in rice. *J. Zhejiang Univ. Sci. B* 9, 363–370. doi: 10.1631/jzus.B0710603

Dingkuhn, M., Laza, M. R. C., Kumar, U., Mendez, K. S., Collard, B., Jagadish, K., et al. (2015). Improving yield potential of tropical rice: achieved levels and perspectives through improved ideotypes. *Field Crops Res.* 182, 43–59. doi: 10.1016/j.fcr.2015.05.025

Dong, X., Wang, X., Zhang, L., Yang, Z., Xin, X., Wu, S., et al. (2013). Identification and characterization of *OsEBS*, a gene involved in enhanced plant biomass and spikelet number in rice. *Plant Biotechnol. J.* 11, 1044–1057. doi: 10.1111/pbi.12097

Dong, Y. J., Kamiunten, H., Ogawa, T., Tsuzuki, E., Terao, H., Lin, D. Z., et al. (2004). Mapping of QTLs for leaf developmental behavior in rice (*Oryza sativa* L.). *Euphytica* 138, 169–175. doi: 10.1023/B:EUPH.0000046799.21410.13

Dreccer, M. F., Chapman, S. C., Ogbonnaya, F. C., Borgognone, M. G., and Trethowan, R. M. (2008). Crop and environmental attributes underpinning genotype by environment interaction in synthetic-derived bread wheat evaluated in Mexico and Australia. *Aust. J. Agric. Res.* 59, 447–460. doi: 10.1071/AR07220

Fujita, D., Ebron, L. A., Araki, E., Kato, H., Khush, G. S., Sheehy, J. E., et al. (2010). Fine mapping of a gene for low-tiller number, Ltn, in japonica rice (*Oryza sativa* L.) variety Aikawa 1. *Theor. Appl. Genet.* 120, 1233–1240. doi: 10.1007/s00122-009-1251-7

Fujita, D., Santos, R. E., Ebron, L. A., Telebanco-Yanoria, M. J., Kato, H., Kobayashi, S., et al. (2009). Development of introgression lines of an Indica-type rice variety, IR64, for unique agronomic traits and detection of the responsible chromosomal regions. *Field Crops Res.* 114, 244–254. doi: 10.1016/j.fcr.2009.08.004

Fujita, D., Tagle, A. G., Ebron, L. A., Fukuta, Y., and Kobayashi, N. (2012). Characterization of near-isogenic lines carrying QTL for high spikelet number with the genetic background of an indica rice variety IR64 (*Oryza sativa* L.). *Breed. Sci.* 62, 18–26. doi: 10.1270/jsbbs.62.18

Fujita, D., Trijatmiko, K. R., Tagle, A. G., Sapasap, M. V., Koide, Y., Sasaki, K., et al. (2013). *NAL1* allele from a rice landrace greatly increases yield in modern *indica* cultivars. *Proc. Natl. Acad. Sci. U.S.A.* 110, 20431–20436. doi: 10.1073/pnas.1310790110

Fujita, K., and Yoshida, S. (1984). Partitioning of photosynthates between panicle and vegetative organs of rice under different planting densities. *Soil Sci. Plant Nutr.* 30, 519–525. doi: 10.1080/00380768.1984.10434719

Gaju, O., Reynolds, M. P., Sparkes, D. L., Mayes, S., Ribas-Vargas, G., Crossa, J., et al. (2014). Relationships between physiological traits, grain number and yield potential in a wheat DH population of large spike phenotype. *Field Crops Res.* 164, 126–135. doi: 10.1016/j.fcr.2014.05.015

Hashida, Y., Aoki, N., Kawanishi, H., Okamura, M., Ebitani, T., Hirose, T., et al. (2013). A near isogenic line of rice carrying chromosome segments containing OsSPS1 of Kasalath in the genetic background of Koshihikari produces an increased spikelet number per panicle. *Field Crops Res.* 149, 56–62. doi: 10.1016/j.fcr.2013.04.020

Huang, X., Qian, Q., Liu, Z., Sun, H., He, S., Luo, D., et al. (2009). Natural variation at the DEP1 locus enhances grain yield in rice. *Nat. Genet.* 41, 494–497. doi: 10.1038/ng.352

Ikeda, K., Ito, M., Nagasawa, M., Kyozuka, J., and Nagato, Y. (2007). Rice *ABERRANT PANICLE ORGANIZATION* 1, encoding an F-box protein, regulates meristem fate. *Plant J.* 51, 1030–1040. doi: 10.1111/j.1365-313X.2007.03200.x

Itoh, Y., and Shimizu, H. (2012). Phyllochron dynamics during the course of late shoot development might be affected by reproductive development in rice (*Oryza sativa* L.). *Dev. Genes Evol.* 222, 341–350. doi: 10.1007/s00427-012-0419-3

Jamil, M., Charnikhova, T., Houshyani, B., van Ast, A., and Bouwmeester, H. (2012). Genetic variation in strigolactone production and tillering in rice and its effect on *Striga hermonthica* infection. *Planta* 235, 473–484. doi: 10.1007/s00425-011-1520-y

Khush, G. S. (1987). Rice breeding: past, present and future. *J. Genet.* 66, 195–216. doi: 10.1007/BF02927713

Komatsu, K., Maekawa, M., Ujiie, S., Satake, Y., Furutani, I., Okamoto, H., et al. (2003). LAX and SPA: major regulators of shoot branching in rice. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11765–11770. doi: 10.1073/pnas.1932414100

Lafarge, T. A., Broad, I. J., and Hammer, G. L. (2002). Tillering in grain sorghum over a wide range of population densities: identification of a common hierarchy for tiller emergence, leaf area development and fertility. *Ann. Bot.* 90, 87–98. doi: 10.1093/aob/mcf152

Lafarge, T., and Bueno, C. S. (2009). Higher crop performance of rice hybrids than of elite inbreds in the tropics: 2. does sink regulation, rather than sink size, play a major role? *Field Crops Res.* 114, 434–440. doi: 10.1016/j.fcr.2009.03.007

Lafarge, T., Seassau, C., Martin, M., Bueno, C., Clément-Vidal, A., Shreck, E., et al. (2010). Regulation and recovery of sink strength in rice plants grown under changes in light intensity. *Funct. Plant Biol.* 37, 413–428. doi: 10.1071/FP09137

Li, X., Qian, Q., Fu, Z., Wang, Y., Xiong, G., Zeng, D., et al. (2003). Control of tillering in rice. *Nature* 422, 618–621. doi: 10.1038/nature01518

Liu, G. L., Mei, H. W., Yu, X. Q., Zou, G. H., Liu, H. Y., Hu, S. P., et al. (2008). QTL analysis of panicle neck diameter, a trait highly correlated with panicle size, under well-watered and drought conditions in rice (*Oryza sativa* L.). *Plant Sci.* 174, 71–77. doi: 10.1016/j.plantsci.2007.09.011

Liu, W., Zhang, D., Tang, M., Li, D., Zhu, Y., Zhu, L., et al. (2013). THIS1 is a putative lipase that regulates tillering, plant height, and spikelet fertility in rice. *J. Exp. Bot.* 64, 4389–4402. doi: 10.1093/jxb/ert256

Liu, Y., Wang, Q., Ding, Y., Li, G., Xu, J., and Wang, S. (2011). Effects of external ABA, GA3 and NAA on the tiller bud outgrowth of rice is related to changes in endogenous hormones. *Plant Growth Regul.* 65, 247–254. doi: 10.1007/s10725-011-9594-x

Miura, K., Ikeda, M., Matsubara, A., Song, X. J., Ito, M., Asano, K., et al. (2010). *OsSPL14* promotes panicle branching and higher grain productivity in rice. *Nat. Genet.* 42, 545–592. doi: 10.1038/ng.592

Nemoto, K., Morita, S., and Baba, T. (1995). Shoot and root development in rice related to the phyllochron. *Crop Sci.* 35, 24–29. doi: 10.2135/cropsci1995.0011183X003500010005x

Okami, M., Kato, Y., Kobayashi, N., and Yamagishi, J. (2014). Agronomic performance of an IR64 introgression line with large leaves derived from New Plant Type rice in aerobic culture. *Eur. J. Agron.* 58, 11–17. doi: 10.1016/j.eja.2014.03.001

Okami, M., Kato, Y., Kobayashi, N., and Yamagishi, J. (2015). Morphological traits associated with vegetative growth of rice (*Oryza sativa* L.) during the recovery phase after early-season drought. *Eur. J. Agron.* 64, 58–66. doi: 10.1016/j.eja.2014.12.006

Okami, M., Kato, Y., and Yamagishi, J. (2012). Allometric relationship between the size and number of shoots as a determinant of adaptations in rice to water-saving aerobic culture. *Field Crops Res.* 131, 17–25. doi: 10.1016/j.fcr.2012.02.014

Okawa, S., Makino, A., and Mae, T. (2003). Effect of irradiance on the partitioning of assimilated carbon during the early phase of grain filling in rice. *Ann. Bot.* 92, 357–364. doi: 10.1093/aob/mcg147

Qi, J., Qian, Q., Bu, Q., Li, S., Chen, Q., Sun, J., et al. (2008). Mutation of the rice *narrow leaf1* gene, which encodes a novel protein, affects vein patterning and polar auxin transport. *Plant Physiol.* 147, 1947–1959. doi: 10.1104/pp.108.118778

Rebolledo, M. C., Dingkuhn, M., Clément-Vidal, A., Rouan, L., and Luquet, D. (2012). Phenomics of rice early vigour and drought response: are sugar related and morphogenetic traits relevant? *Rice* 5:22. doi: 10.1186/1939-8433-5-22

Streck, N. A., Lago, I., Bosco, L. C., de Paula, G. M., Oliveira, F. B., Gabriel, L. F., et al. (2009). Relationship between panicle differentiation and main stem leaf number in rice genotypes and red rice biotypes. *Sci. Agric.* 66, 195–203. doi: 10.1590/S0103-90162009000200008

Terao, T., Nagata, K., Morino, K., and Hirose, T. (2010). A gene controlling the number of primary rachis branches also controls the vascular bundle formation and hence is responsible to increase the harvest index and grain yield in rice. *Theor. Appl. Genet.* 120, 875–893. doi: 10.1007/s00122-009-1218-8

Tivet, F., Pinheiro, B. D. S., Raissac, M. D., and Dingkuhn, M. (2001). Leaf blade dimensions of rice (*Oryza sativa* L. and *Oryza glaberrima* Steud.). Relationships between tillers and the main stem. *Ann. Bot.* 88, 507–511. doi: 10.1006/anbo.2001.1447

Ugarte, C. C., Trupkin, S. A., Ghiglione, H., Slafer, G., and Casal, J. J. (2010). Low red/far-red ratios delay spike and stem growth in wheat. *J. Exp. Bot.* 61, 3151–3162. doi: 10.1093/jxb/erq140

Wang, W., Li, G., Zhao, J., Chu, H., Lin, W., Zhang, D., et al. (2014). DWARF TILLER1, a WUSCHEL-related homeobox transcription factor, is required for tiller growth in rice. *PLoS Genet.* 10:e1004154. doi: 10.1371/journal.pgen.1004154

Weng, X., Wang, L., Wang, J., Hu, Y., Du, H., Xu, C., et al. (2014). Grain number, plant height, and heading date7 is a central regulator of growth, development, and stress response. *Plant Physiol.* 164, 735–747. doi: 10.1104/pp.113.231308

Wu, L. L., Liu, Z. L., Wang, J. M., Zhou, C. Y., and Chen, K. M. (2011). Morphological, anatomical, and physiological characteristics involved in development of the large culm trait in rice. *Aust. J. Crop Sci.* 5, 1356–1363.

Zhang, B., and Yamagishi, J. (2010). Response of spikelet number per panicle in rice cultivars to three transplanting densities. *Plant Prod. Sci.* 13, 279–288. doi: 10.1626/pps.13.279

Check for updates

High-Throughput Non-destructive Phenotyping of Traits that Contribute to Salinity Tolerance in *Arabidopsis thaliana*

Mariam Awlia[1], Arianna Nigro[2], Jiří Fajkus[3], Sandra M. Schmoeckel[1], Sónia Negrão[1], Diana Santelia[2], Martin Trtílek[3], Mark Tester[1], Magdalena M. Julkowska[1]* and Klára Panzarová[3]*

[1] Division of Biological and Environmental Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, [2] Institute of Plant and Microbial Biology, University of Zurich, Zurich, Switzerland, [3] PSI (Photon Systems Instruments), Drásov, Czech Republic

Reproducible and efficient high-throughput phenotyping approaches, combined with advances in genome sequencing, are facilitating the discovery of genes affecting plant performance. Salinity tolerance is a desirable trait that can be achieved through breeding, where most have aimed at selecting for plants that perform effective ion exclusion from the shoots. To determine overall plant performance under salt stress, it is helpful to investigate several plant traits collectively in one experimental setup. Hence, we developed a quantitative phenotyping protocol using a high-throughput phenotyping system, with RGB and chlorophyll fluorescence (ChlF) imaging, which captures the growth, morphology, color and photosynthetic performance of *Arabidopsis thaliana* plants in response to salt stress. We optimized our salt treatment by controlling the soil-water content prior to introducing salt stress. We investigated these traits over time in two accessions in soil at 150, 100, or 50 mM NaCl to find that the plants subjected to 100 mM NaCl showed the most prominent responses in the absence of symptoms of severe stress. In these plants, salt stress induced significant changes in rosette area and morphology, but less prominent changes in rosette coloring and photosystem II efficiency. Clustering of ChlF traits with plant growth of nine accessions maintained at 100 mM NaCl revealed that in the early stage of salt stress, salinity tolerance correlated with non-photochemical quenching processes and during the later stage, plant performance correlated with quantum yield. This integrative approach allows the simultaneous analysis of several phenotypic traits. In combination with various genetic resources, the phenotyping protocol described here is expected to increase our understanding of plant performance and stress responses, ultimately identifying genes that improve plant performance in salt stress conditions.

Keywords: high-throughput phenotyping, *Arabidopsis thaliana*, salt stress, salinity tolerance, shoot-ion independent tolerance, kinetic chlorophyll fluorescence imaging, color segmentation

INTRODUCTION

Climate change and population growth place a twofold pressure on agricultural crop production. Crop yields need to be sustained and increased while grown in unfavorable environments (Godfray et al., 2010; Tester and Langridge, 2010; Tilman et al., 2011). To meet future food demands, breeding efforts are targeting more resource-efficient and stress-tolerant crops by combining large-scale plant phenotyping with genome sequencing in forward genetics studies. Phenotypic traits, including growth rate, size, shape, color, temperature and photosynthetic activity, are traditionally studied to evaluate plant performance under stress (Sirault et al., 2009; Zhang et al., 2012; Hairmansis et al., 2014; Chen et al., 2015; Ghanem et al., 2015). Plant breeding programs aimed at enhancing plant performance should investigate growth and photosynthetic activity in tandem because these processes are interdependent (Longenberger et al., 2009).

Advances in non-destructive image-based phenotyping technologies are enabling parallel studies of plant growth and photosynthetic performance over time (Munns et al., 2010; Dhondt et al., 2013; Hairmansis et al., 2014) using RGB and chlorophyll fluorescence (ChlF) imaging (Dhondt et al., 2013; Brown et al., 2014; Humplik et al., 2015). Traits related to plant growth, architecture and development have been quantified from digital color imaging, while leaf color is a simple, under-utilized trait that indicates plant health and leaf senescence (Berger et al., 2012). Kinetic ChlF imaging is a powerful tool for measuring plant photosynthetic capacity and provides valuable insights into the performance of photosynthetic apparatus (Oxborough, 2004; Baker, 2008). Light energy, captured by chlorophyll molecules, can undergo one of three fates: (1) be used to drive photosynthesis by photochemistry, (2) be dissipated as heat, or (3) be re-emitted as fluorescence. Because these three processes co-exist in close competition, the ChlF yield provides information on both the quantum efficiency of the plant's photochemistry and on the amount of heat dissipated. Under conditions of stress, the photochemical yield decreases, which in turn, causes heat dissipation and ChlF emissions to increase (Maxwell and Johnson, 2000; Murchie and Lawson, 2013). Although, high-throughput phenotyping of photosynthetic performance has previously been employed to study plant response to cold (Jansen et al., 2009; Humplik et al., 2015) and drought (Bresson et al., 2015), few studies have established an integrative approach that simultaneously analyzes plant growth and photosystem II (PSII) efficiency (Humplik et al., 2015). Systems for kinetic ChlF imaging have not been widely integrated into high-throughput phenotyping platforms, in contrast to single-level steady-state ChlF imaging. The latter only reflects chlorophyll content and not PSII activity. This means that using single-level steady-state ChlF can only discriminate between healthy, senescing and stressed leaves by the amount of chlorophyll they degrade (Campbell et al., 2015).

Soil salinity is a key stress factor that affects agriculture on a global scale (Munns and Tester, 2008; Cabot et al., 2014; Roy et al., 2014). In saline soil, plants accumulate ions in their shoots, compromising plant growth with ion toxicity, for example by reducing the rate of photosynthesis (Munns and Tester, 2008). During the early phase of salt stress, before ions accumulate significantly in the shoots, the osmotic phase of salinity tolerance takes place, which is referred to as shoot ion-independent tolerance (SIIT; Roy et al., 2014). During this phase, growth reduction is caused by decreased leaf emergence and expansion (Fricke et al., 2006; Munns and Tester, 2008; Berger et al., 2012). Mechanisms underlying ion sensing, cell cycle, cell expansion, and stomatal conductance are likely underlying this response (Ma et al., 2006; Stephan and Schroeder, 2014). Thus, performing regular growth measurements from when salt stress is introduced to the plant across an extended period of time provides the opportunity to discriminate between early (osmotic) and late (ionic) growth-related responses to salt stress. During the early phase of salt stress, the capacity of photosynthetic machinery is reduced (James et al., 2006; Chaves et al., 2009). In later phases, excessive photonic energy causes photochemical inactivation (Muranaka et al., 2002) that reduces PSII stability (Stepien and Johnson, 2009), limits stomatal gas diffusion and causes changes in carbon assimilation rates (Chaves et al., 2009), ultimately resulting in decreased photosynthetic activity. Additional limits to photosynthesis may be caused by the accumulation of unused organic compounds from carbon assimilation (Munns and Tester, 2008). Early responses of plants to salinity have previously been quantified by measuring rosette area, color, temperature, and photosynthetic activity using steady-state ChlF (Rajendran et al., 2009; Sirault et al., 2009; Berger et al., 2012; Chen et al., 2015). Here, we quantify how traits related to plant morphology, color and photosynthetic activity respond to salinity in one experimental setup to establish significant correlations among individual phenotypes.

We developed a phenotyping method that monitors plant responses to salt stress by evaluating plant growth, color and photosynthetic traits using an automated, high-throughput system. We grew Arabidopsis plants in soil and maintained them at 40, 60, or 80% of the soil-water holding capacity to achieve approximately 150, 100, and 50 mM NaCl, respectively. We established 100 mM NaCl as the optimal condition for salt treatment. To characterize the early and late plant responses to salt stress, we investigated RGB, greenness and photosynthesis-related traits. Traits of ChlF were clustered with relative plant performance values into groups corresponding to early and late responses to salt stress. This work provides the means for screening natural diversity panels and mapping populations to identify candidate genes underlying plant development and stress tolerance.

MATERIALS AND METHODS

Plant Material and Growing Conditions

Accessions of Arabidopsis Columbia-0 (Col-0) and C24 were used to establish the cultivation, salt treatment and phenotyping protocol. Thereafter, nine accessions of Arabidopsis [Col-0, C24, Canary Islands (Can), Coimbra (Co), Cape Verde Islands (Cvi), Landsberg *erecta* (Ler), Niederzenz (Nd), Rschew (Rsch) and Tenela (Te)] were used to optimize this protocol

and investigate the natural variation of plants in response to salt stress (Hannah et al., 2006). Seeds were sown into pots (70 mm × 70 mm × 65 mm, Poppelman TEKU DE) containing 60 g of freshly sieved soil (Substrate 2, Klasmann-Deilmann GmbH, Germany) and watered to full soil-water holding capacity. Seeds were stratified for 3 days at 4°C in the dark. All plants were grown in a climate controlled growth chamber (FS_WI, PSI, Czech Republic) with cool-white LED and far-red LED lighting. The protocol was setup with Col-0 and C24 grown in a 10 h/14 h 21°C/15°C light/dark cycle at a relative humidity of 60% and a photon irradiance of 250 μmol m^{-2} s^{-1}. The protocol was optimized for the nine accessions using a 12 h/12 h 22°C/20°C light/dark cycle with a relative humidity of 55% and an irradiance of 150 μmol m^{-2} s^{-1}. Seven days after stratification (DAS), seedlings of similar size were transplanted into soil that had been watered 1 day in advance to full soil-water holding capacity. Plants were cultivated in the growth chamber until most plants were at the 10-leaf stage (24 DAS for plants in the 10 h/14 h light/dark cycle and 21 DAS in the 12 h/12 h light/dark cycle). The growth timeline for Col-0 and C24 plants illustrates the implementation of the three watering regimes, the salt treatment and the phenotyping protocol (**Figure 1**).

Watering and Salt Treatment

Similar to Junker et al. (2015), we determined the soil water-holding capacity by filling 10 pots with 60 g of sieved soil and drying them for 3 days at 80°C to completely desiccate the soil. Soil was then saturated with water and left to drain for 1 day before weighing. Based on the soil-water content at 100% (130 g),

40, 60, and 80% of the soil-water holding capacity were found to weigh 52, 77, and 103 g, respectively. At 14 DAS, Col-0 and C24 seedlings were placed randomly in trays (5 × 4 pots per tray) and into the PlantScreen™ Compact System (PSI, Czech Republic) and automatically weighed and watered every other day to reach and maintain the reference weight corresponding to the desired soil-water contents (**Figure 1A**). Once at the 10-leaf stage at 24 DAS, nine replicates per accession were placed in 250 mM NaCl or dH$_2$O for 1 h to ensure full saturation of the soil. Pots were left to drain for 10 min before being placed in the phenotyping system. Effective NaCl concentration in the soil was estimated as 150, 100, and 50 mM NaCl in plants watered to 40, 60, and 80% soil-water holding capacity, respectively, representing conditions of severe, moderate, and mild salt stress (**Figure 1B**). For 11 days, with the exception of days 5 and 6 (**Figure 1C**), plants were transferred manually from the growth chamber to PlantScreen™ for image acquisition and then returned to the same positions inside the growth chamber. On the final day of imaging, the water content of the soil was found to be approximately 60–70%, indicating that plants had adequate soil moisture during the phenotyping period.

After analyzing the effects of the three watering regimes, 60% soil-water content and 100 mM NaCl were established as effective conditions for investigating early responses to salt stress. We then optimized the protocol, in terms of growth conditions and ChlF imaging, using nine accessions of Arabidopsis. At 18 and 20 DAS pots were watered to the target soil-water content. Once at the 10-leaf stage, at 21 DAS, eight replicates per accession were placed in a 250 mM NaCl solution or dH$_2$O for 1 h to
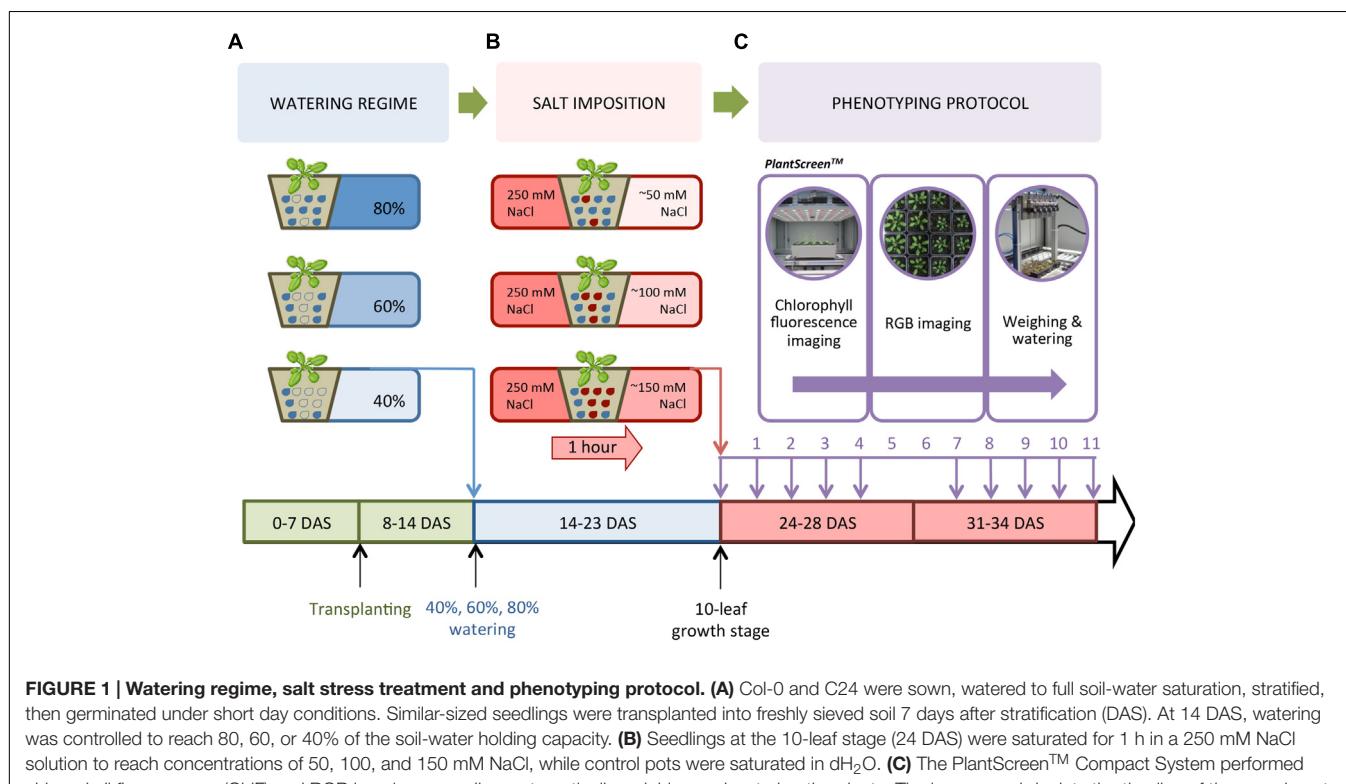
FIGURE 1 | Watering regime, salt stress treatment and phenotyping protocol. (A) Col-0 and C24 were sown, watered to full soil-water saturation, stratified, then germinated under short day conditions. Similar-sized seedlings were transplanted into freshly sieved soil 7 days after stratification (DAS). At 14 DAS, watering was controlled to reach 80, 60, or 40% of the soil-water holding capacity. **(B)** Seedlings at the 10-leaf stage (24 DAS) were saturated for 1 h in a 250 mM NaCl solution to reach concentrations of 50, 100, and 150 mM NaCl, while control pots were saturated in dH$_2$O. **(C)** The PlantScreen™ Compact System performed chlorophyll fluorescence (ChlF) and RGB imaging, as well as automatically weighing and watering the plants. The lower panel depicts the timeline of the experiment.

ensure saturation of the soil. Plants were imaged for 7 days with no additional watering. The final soil-water content was approximately 70–80%.

High-Throughput Phenotyping

Control and salt-treated plants were automatically phenotyped for RGB and kinetic ChlF traits using PlantScreen[TM] (Supplementary Figures S1 and S2) from 1 h after introducing salt stress. The phenotyping was conducted for 11 days to develop the protocol and 7 days to investigate natural variation among the nine accessions. Trays were transported within PlantScreen[TM] on conveyor belts between the light-isolated imaging cabinets, weighing and watering station and the dark/light acclimation chamber. A single round of measuring consisted of an initial 15 min dark-adaptation period inside the acclimation chamber, followed by ChlF and RGB imaging, weighing and watering. Pixel count, color and fluorescence intensity were evaluated from the images. A total time of 1 h and 40 min was required to measure 10 trays (200 plants). The PlantScreen[TM] Analyzer software (PSI, Czech Republic) was used to automatically process the raw data.

RGB Imaging and Processing

Trays were loaded into the imaging cabinet of the PlantScreen[TM] platform with three RGB cameras (one top and two side views) mounted on robotic arms, each supplemented with an LED-based light source to ensure homogenous illumination of the imaged object. To assess plant growth and morphological traits, RGB images (resolution 2560×1920 pixels) of 5×4 plants per tray were captured using the GigE uEye UI-5580SE-C/M 5 Mpx Camera (IDS, Germany) from the top view only. Light conditions, plant position and camera settings were fixed throughout the experiments. The PlantScreen[TM] software required three steps to extract features from the RGB images. (1) Basic processing applied in real-time involving correction for barrel (fisheye) distortion, tray detection, cropping of individual pots, background subtraction to remove non-plant pixels from the images, filtration and artifact removal to produce binary (black and white) and RGB representations of each plant [binary images represent the plant's surface (white) and background (black)]; (2) morphologic analysis, requiring separation of the background from the plant shoot tissue allowing the pixel number per plant and rosette area to be counted; and (3) analysis of greenness using background-subtracted RGB images to evaluate the color. For this step, the images were color-segmented to represent and evaluate rosette coloring (Supplementary Figure S1). The morphometric parameters area, perimeter, roundness, compactness, rotational mass symmetry, eccentricity and slenderness of leaves were computed from the RGB image processing and have been listed and defined in Supplementary Table S1.

Plant Growth-Related Parameters

To evaluate the effect of salt stress on early and late plant growth rates (GR), we examined the increase in projected rosette area over time by fitting a linear function to two time intervals. The regression coefficient of the fitted function was determined and used as a trait in the statistical analysis. Relative effects of salt stress were calculated by dividing the estimated growth rates (GR) in salt conditions by the average in control conditions ($GR_{(salt)}/GR_{(control)}$). The calculation was performed for each accession over two time intervals (0–4 days and then 7–11 days when developing the protocol, and 0–3 days and then 4–7 days when examining natural variation among accessions; **Figures 2D and 5C**). This ratio has been termed the shoot ion-independent tolerance (SIIT) index, which is used as a measure of plant salinity tolerance (Roy et al., 2014). The effects of salt stress were determined by performing an analysis of variance (ANOVA) per accession and treatment with Tukey's *post hoc* test of significance for each RGB trait (p-value < 0.05).

Color Segmentation and Evaluation of Greenness

Using color segmentation, we analyzed the change in rosette coloring. We calibrated the analysis by using RGB images from both control and salt-stressed conditions and from the start, middle and end of the phenotyping period to obtain an unbiased color scale (Supplementary Figure S1). Values in the RGB channels, from each pixel corresponding to the plant's surface area, were extracted to serve as a dataset for k-means clustering. Nine clusters were sufficient for optimal color differentiation and all input pixels were partitioned according to their Euclidean distance in the RGB color space. The RGB coordinates of cluster centroids were used as base hues to evaluate greenness. Original pixel color was approximated from the nearest cluster centroid, yielding color-segmented images. To calculate the relative hue abundance independent of the rosette area, pixel counts of individual hue values were divided by the rosette area of the same plant on the same day (**Figure 3**). The effect of salt stress was determined by performing an ANOVA per accession and treatment with Tukey's *post hoc* test of significance for each greenness hue (p-value < 0.05).

Kinetic Chlorophyll Fluorescence Imaging and Processing

To assess the effect of salt stress on photosynthetic performance, ChlF measurements were acquired using an enhanced version of the FluorCam FC-800MF pulse amplitude modulated (PAM) chlorophyll fluorometer (PSI, Czech Republic). The ChlF imaging station was mounted on a robotic arm with an LED light panel and a high-speed charge-coupled device camera (pixel resolution of 720×560, frame rate 50 fps and 12-bit depth) positioned in the middle of the light panel (Supplementary Figure S2). The LED panel was equipped with 3×64 orange-red (618 nm) and 64 cool-white LEDs (6500 K), distributed equally over 75×75 cm. This resulted in a $\pm 5\%$ maximum deviation from the mean across the imaged area of 35×35 cm. Modulated light of known wavelength was applied to detect the ChlF signal. Three types of light sources were used: (1) PAM short-duration measuring flashes (33 μs) at 618 nm, (2) orange-red (618 nm) and cool-white (6500 K) actinic lights with maximum irradiance 440 μmol m^{-2} s^{-1} and (3) saturating cool-white light with maximum irradiance 3000 μmol m^{-2} s^{-1}.

Plant trays were automatically loaded into the light-isolated imaging cabinet of PlantScreen[TM] with a top-mounted LED
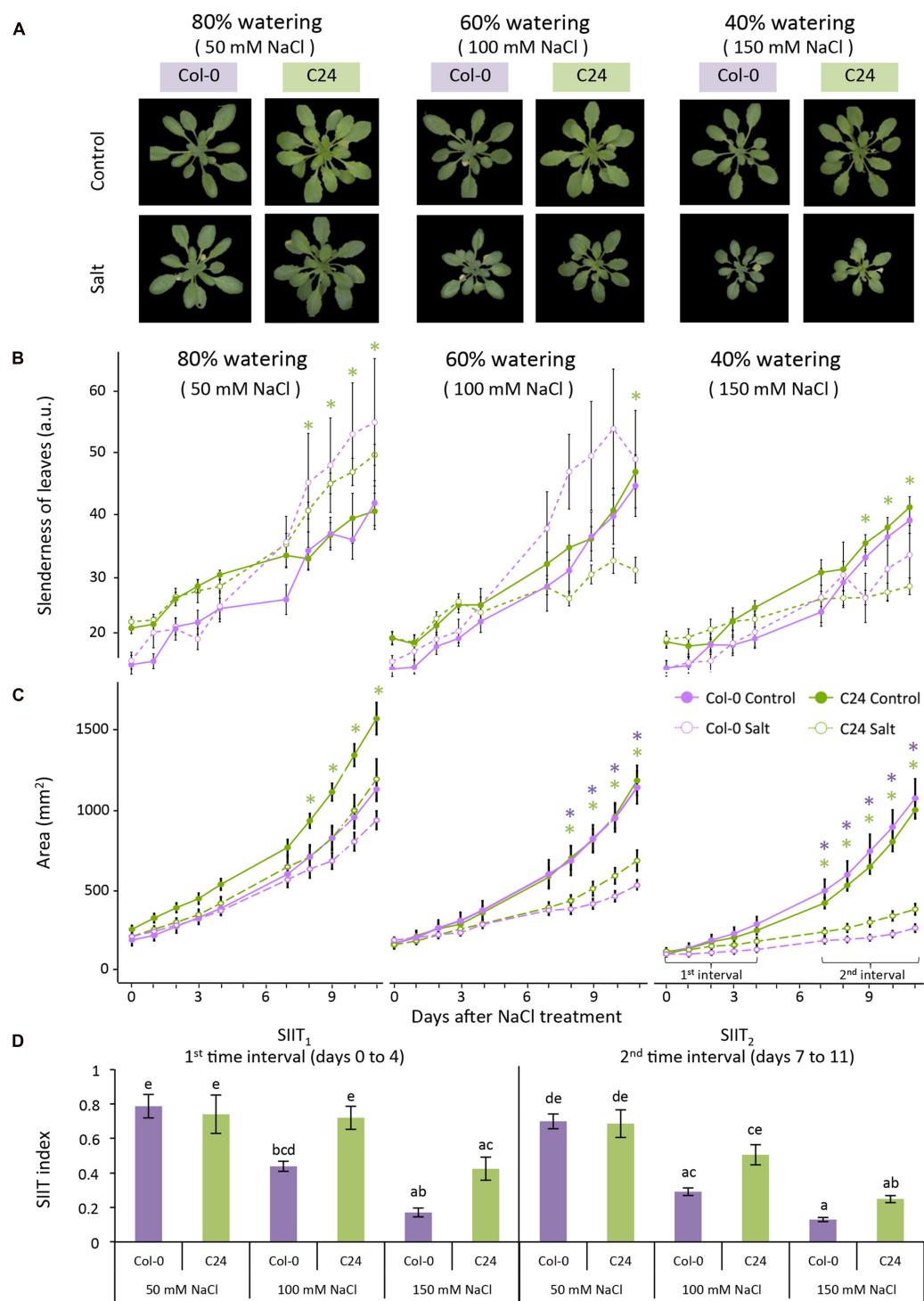
FIGURE 2 | Col-0 and C24 growth-related responses to salt stress. (A) RGB images of control (upper panel) and 11 days salt-stressed (lower panel) plants maintained at 80, 60, and 40% of the soil water-holding capacity, which correspond to approximately 50, 100, and 150 mM NaCl, respectively. **(B)** Leaf slenderness and **(C)** projected rosette area over time for Col-0 (purple lines) and C24 (green lines) in control (solid lines) and salt-stressed (dashed lines) conditions. Values represent the averages of nine replicates per accession and treatment. Error bars represent standard error. The significant differences between control and salt treatment per accession are indicated with * and ** for p-values below 0.05 and 0.01, respectively. **(D)** The shoot ion-independent tolerance (SIIT) index was calculated as the ratio of growth rates in salt-stressed conditions to average growth rates observed in control conditions per accession and treatment over two time intervals. Values represent the average of nine biological replicates per accession and treatment. Error bars represent standard error. Different letters are used to indicate the significant differences between the accessions and conditions as tested with one-way ANOVA with *post hoc* Tukey's test ($p < 0.05$).
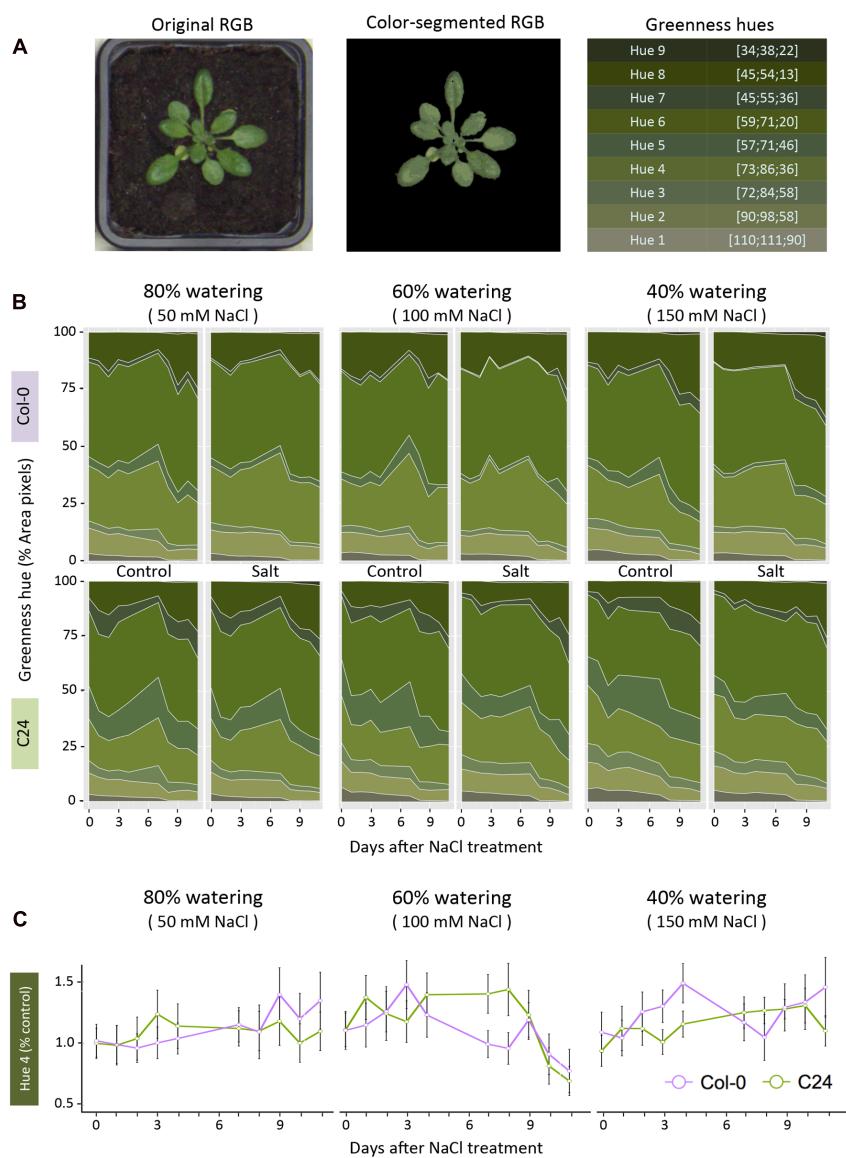
FIGURE 3 | Quantification of change in rosette color in response to salt stress. (A) RGB images were color-segmented into nine green hues, which are summarized with their RGB coordinates. **(B)** Dynamic relative changes in greenness hue abundance are presented across the phenotyping period, with 50 mM NaCl (left panel), 100 mM NaCl (middle panel) and 150 mM NaCl (right panel), respectively. The hues ordered in the right panel of A are presented as a portion of the rosette area (pixel counts) in 80, 60, and 40% watering with 50, 100, and 150 mM NaCl, respectively. Values shown represent the averages calculated over nine biological replicates per accession and treatment. **(C)** Salt-induced changes in hue 4, calculated as relative to control conditions, for Col-0 (purple) and C24 (green). Values represent the averages of nine replicates per accession and treatment. Error bars represent standard error. The significant differences between control and salt-stressed treatments per accession are indicated with * and ** for p-values below 0.05 and 0.01, respectively, as calculated with one-way ANOVA with *post hoc* Tukey's test.

light panel. After the 15 min dark-adaptation period, when PSII reaction centers open, the trays were automatically transported to the ChlF imaging cabinet. A 5 s flash of light was applied to measure the minimum level of fluorescence in the dark-adapted state (F_o), followed by a saturation pulse of 800 ms (with an irradiance of 1200 μmol m^{-2} s^{-1}) used to determine the maximum fluorescence in the dark-adapted state (F_m). Plants were relaxed in the dark for 17 s and then subjected to 70 s of cool-white actinic lights to drive photosynthesis and measure the

peak rise in fluorescence (F_p). These conditions were used in both ChlF imaging techniques using quenching kinetics and light curve protocol.

For quenching kinetics protocol, additional saturation pulses were applied at 8, 18, 28, 48, 68 s during actinic illumination, corresponding to L1, L2, L3, L4, and Lss states at a constant photon irradiance of 210 μmol m^{-2} s^{-1} (Supplementary Figure S3). These ChlF signals were used to acquire the maximum fluorescence in the light-adapted state (F_m'), and the

level of ChlF measured just before the saturation pulse was considered the steady-state fluorescence in the light-adapted state (F_t). Further responses to dark-relaxation were measured by switching the actinic light off for 100 s and applying saturating pulses at 30, 60, and 90 s, corresponding to D1, D2, D3 states (Supplementary Figure S2). The PlantScreen[TM] Analyzer software performed the automated ChlF feature extraction by mask application, background subtraction and parameter calculation based on the fluorescence levels of F_o, F_m, F_p, F_t and F_m', which were estimated by integrating pixel-by-pixel values across the entire rosette (Supplementary Figure S2; Supplementary Table S2). Minimum fluorescence in the light-adapted stated (F_o') was calculated according to Oxborough and Baker (1997).

For the examination of natural variation in the nine Arabidopsis accessions we optimized ChlF imaging by quantifying the rate of photosynthesis at different photon irradiances using the light curve protocol (Henley, 1993; Rascher et al., 2000) which was proven to provide detailed information on ChlF under stress (Brestic and Zivcak, 2013). A 5 s flash of light was applied to measure the minimum fluorescence, followed by a saturation pulse of 800 ms (with an irradiance of 1200 μmol m^{-2} s^{-1}) to determine the maximum fluorescence in the dark-adapted state. Next, 60 s intervals of cool-white actinic light at 95, 210, 320, 440 μmol m^{-2} s^{-1} corresponding to L1, L2, L3, and L4, respectively, were applied. A saturation pulse was applied at the end of the period of actinic light to acquire the maximal fluorescence in the light-adapted state (Supplementary Figure S6A). The ChlF signal measured just before the saturation pulse was taken as the steady-state fluorescence value in the light-adapted state. The ChlF parameters were extracted and processed as described above for data collected from day 0 to day 7 of salt treatment (**Figure 6**).

Statistical Analysis on Chlorophyll Fluorescence-Related Responses to Salt Stress

An ANOVA with Tukey's *post hoc* test of significance (*p*-value < 0.05) was used to evaluate the differences in ChlF between control and salt-stressed plants. Trait values specific to accession, day and condition were divided by the overall average per trait to analyze the fluctuations in the ChlF traits, which were due to both plant development and salt treatment (**Figures 4** and **6**). Principal component analysis (PCA) was performed on 20 ChlF traits collected from Col-0 and C24 under the different adapted states and saturating pulses (L1 to L4, Lss and D1 to D3) to reduce data dimensionality (Supplementary Figure S5; Supplementary Table S3). Eight ChlF traits (F_v/F_m, F_v'/F_m', ΦP, qP, ΦNO, ΦNPQ, qN and NPQ; Lazar, 2015) measured on day 7 at 440 μmol m^{-2} s^{-1} were clustered with $SIIT_1$ and $SIIT_2$ ($GR_{salt}/GR_{control}$ for each time interval) using the Ward linkage method. This was performed to study the relationships between the ChlF traits and relative changes in growth rate under salt stress among the nine accessions (**Figure 7**). Normalization was done by dividing the relative trait values by the overall average per trait. Mann–Whitney *U*-test with a continuity correction were performed on the ChlF parameters captured by the light curve protocol, and *p*-values were calculated using treatment as a grouping variable per accession for each day of the phenotyping period (Supplementary Table S4).

RESULTS

Salt Stress Affected Growth and Morphology-Related Traits Over Time

To define the most suitable screening conditions to study early responses to salt stress, we first aimed to optimize the salt treatment (**Figure 1**). Three watering regimes were used to control the soil-water content (40, 60, and 80%; **Figure 1A**). The salt solution was diluted according to the watering regime to the final concentrations of 150, 100, or 50 mM NaCl in the soil corresponding to severe, moderate and mild salt stress, respectively (**Figure 1B**) and plants were phenotyped using RGB and ChlF imaging (**Figure 1C**). We then analyzed the phenotypes of Col-0 and C24 plants to discern the conditions most suitable for screening early salt-induced changes without greatly compromising plant health (**Figure 2**).

Investigation of rosette morphology revealed that Col-0 developed more slender leaves than C24 under control conditions (**Figure 2B**; Supplementary Table S1), but under severe salt stress, differences were less pronounced. Changes in roundness and compactness of Col-0 and C24 leaves were apparent after 2–3 days of salt treatment, whereas changes in rotational mass symmetry, eccentricity and slenderness of Col-0 leaves were recorded after 1 day of salt treatment (Supplementary Table S1; **Figure 2C**). C24 plants showed significant decreases in rosette area with salt treatment at day 8 in mild and moderate stress conditions and at day 7 in severe stress conditions. Mild stress levels did not cause a significant decrease in rosette area in Col-0 plants but a marked decrease was observable at day 8 with moderate and at day 7 with severe salt stress (**Figure 2C**).

Growth rates were estimated in each accession and condition by splitting the growth period into two intervals, 0–4 days and from 7 to 11 days, and fitting two linear functions to the increase in rosette area over time (**Figure 2C**). Growth rates of salt-treated plants were smaller in both intervals than those of control plants. This reduction was more pronounced in the second interval (Supplementary Table S1), allowing the discrimination between early and late responses to salt stress. The ratio of GR_{salt} to $GR_{control}$ was used to describe the SIIT index and was calculated to assess salinity tolerance in the early ($SIIT_1$) and late ($SIIT_2$) phases of salt stress (**Figure 2D**). There was a clear decrease in SIIT values in both Col-0 and C24 plants with increasing salt stress levels. Both $SIIT_1$ and $SIIT_2$ of C24 were higher than those of Col-0 under moderate and severe salt stress conditions in both intervals, indicating that C24 has higher salinity tolerance. Differences between control and salt-stressed plants and between Col-0 and C24 were most pronounced under moderate and severe conditions of salt stress. In addition, control plants grown in the 40% soil-water content were much smaller than control plants grown in other watering regimes (**Figure 2C**), suggesting that these plants were likely to be suffering from drought stress. Mild salt stress had no effect on the growth of Col-0 and only
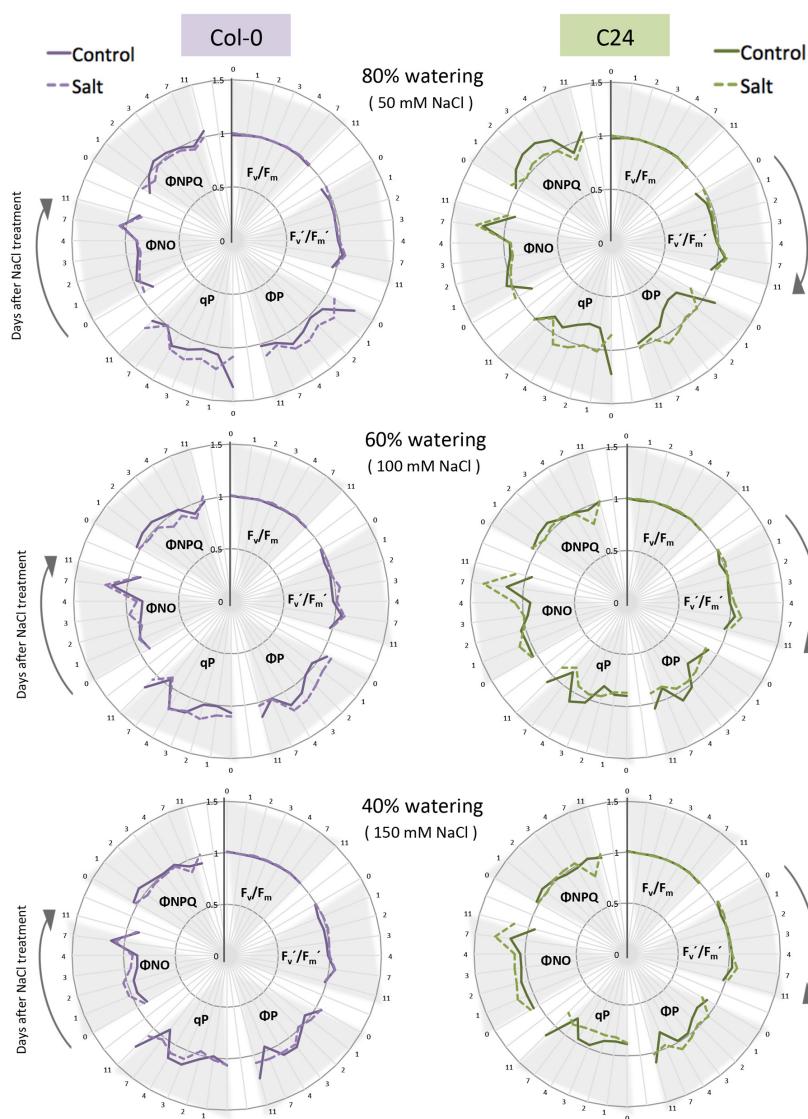
FIGURE 4 | The effect of watering regime and salt treatment on ChlF parameters. Maximum quantum yield of PSII photochemistry for the dark-adapted (F_v/F_m) and light-adapted (F_v'/F_m') states, the quantum yield of PSII photochemistry for the light-adapted state (ΦP), the photochemical quenching coefficient that estimates the fraction of open PSII reaction centers (qP), the quantum yield of constitutive non-regulatory non-photochemical dissipation processes (ΦNO) and the quantum yield of regulatory non-photochemical quenching (ΦNPQ) were measured using the quenching protocol for Col-0 (purple) and C24 (green) plants in control (solid line) and salt-stressed (dashed line) conditions. Time in days is indicated on the outer rim of the graph. Individual ChlF traits are presented in clusters indicated by gray triangles. Values represent the average of nine biological replicates per accession and treatment divided by the overall average per trait. Values were measured for plants grown in 80, 60, and 40% soil-water contents in control or salt-stressed conditions.

a slight effect was observed later in C24 plants (**Figure 2B**). We established that plant growth and performance was best assessed using a 60% watering regime, which resulted in moderate salt stress of 100 mM NaCl.

Color Segmentation of RGB Images Illustrates Changes in Rosette Greenness Over Time

We examined the features of pixel color in the RGB images to identify changes in rosette greenness under salt stress. Color information was extracted from pixels corresponding to the imaged rosettes. RGB images were color-segmented into nine green hues and analyzed for their relative abundance as a percentage of rosette area (**Figure 3A**). This strategy enabled the number of pixels representing each hue to be normalized for rosette area and compared between accessions and treatments (**Figure 3B**). Hues 1, 2, 5, and 8 changed over time without marked differences between treatments, while hues 3, 4, 6, 7, and 9 differed from control conditions after only 1 day of exposure to salt stress (**Figure 3B**). We calculated the ratio for each hue, between control and salt-treated plants, to observe
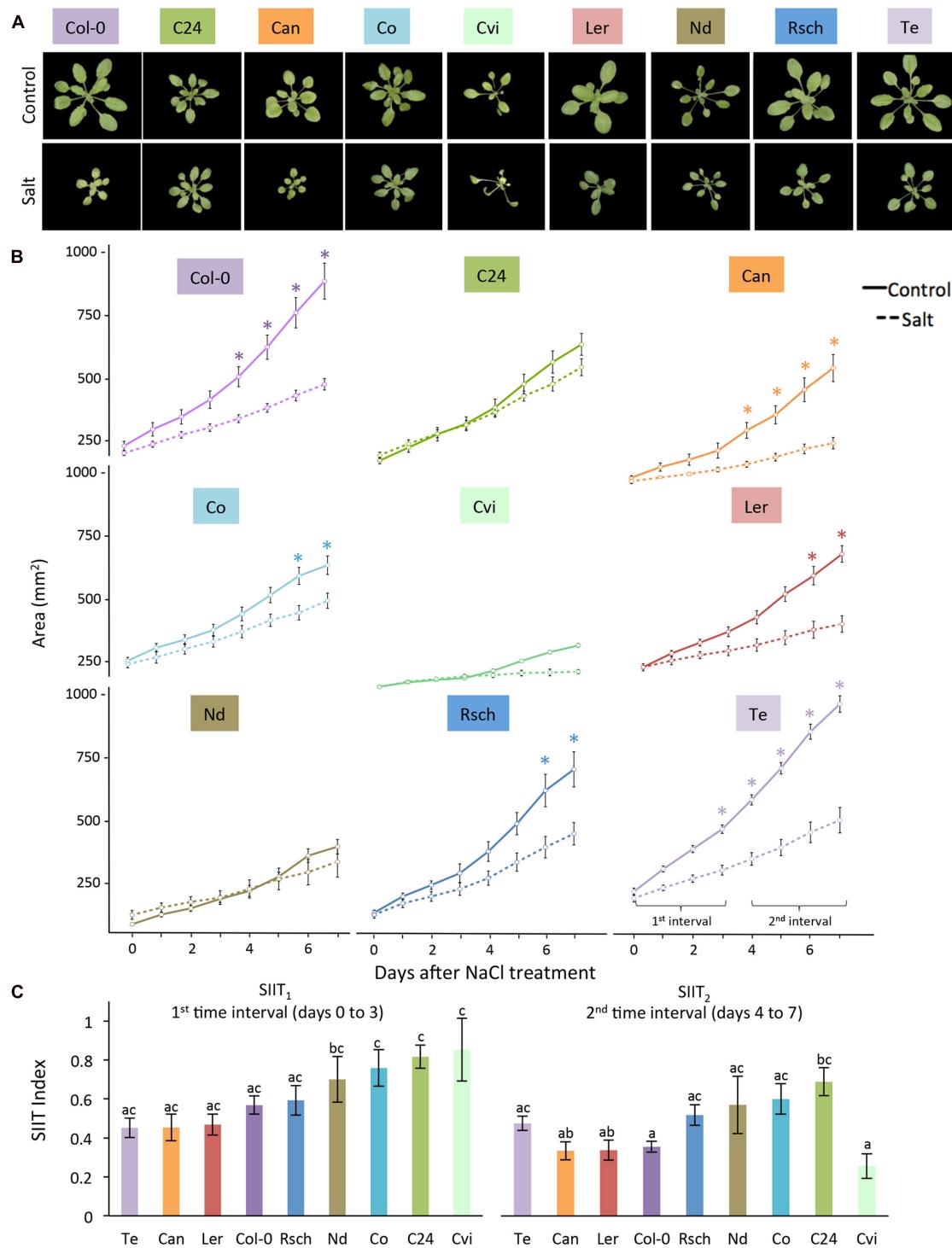
FIGURE 5 | Natural variation in growth-related responses of nine Arabidopsis accessions under salt stress. (A) RGB images of control (upper panel) and 7 days salt-stressed (lower panel) plants. **(B)** Projected rosette area over time for plants grown in control (solid lines) and salt-stressed (dashed lines) conditions. Values represent the average of eight biological replicates per accession and treatment. Error bars represent standard error. Significant differences between control and salt stress treatment are indicated with * and ** for p-values below 0.05 and 0.01, respectively. **(C)** Shoot ion-independent tolerance ($SIIT_1$ and $SIIT_2$) values were calculated from averages of eight replicates per accession and treatment. Error bars represent standard error. Different letters indicate significant differences between accessions as tested with one-way ANOVA with *post hoc* Tukey's test ($p < 0.05$).

**FIGURE 6 | Natural variation in ChlF-related responses of nine Arabidopsis accessions under salt stress.** Maximum quantum yield of PSII photochemistry for the dark-adapted (F_v/F_m) and light-adapted (F_v'/F_m') states, the quantum yield of PSII photochemistry for the light-adapted state (ΦP), the photochemical quenching coefficient that estimates the fraction of open PSII reaction centers (qP), the quantum yield of constitutive non-regulatory non-photochemical dissipation processes (ΦNO) and the quantum yield of regulatory non-photochemical quenching (ΦNPQ) measured by the light curve protocol; control (solid line) and salt-stressed (dashed line) conditions. Time in days is indicated on the outer rim of the graph. Individual ChlF traits are presented in clusters indicated by gray triangles. Values represent the average of eight replicates per accession and treatment divided by the overall average per trait.

the salt-induced changes among between the accessions and treatments. We presented the results for hue 4, showcasing the differences between Col-0 and C24 throughout the phenotyping period (**Figure 3C**).

Chlorophyll Fluorescence Imaging Captures the Early and Late Changes in Photosynthetic Performance in Response to Salt Stress

To further explore the photosynthetic performance of control and salt-treated plants, we used ChlF parameters measured by the PAM method and quenching kinetics. From the measured

fluorescence transient states, the basic ChlF parameters were derived (i.e., F_o, F_m, F_m', F_t, F_v, and F_p), which were used to calculate the quenching coefficients (i.e., qP, NPQ, PQ, and qN) and other parameters characterizing plant photosynthetic performance (i.e., F_o', F_v/F_m, ΦP, F_v'/F_m', ΦNO, ΦNPQ and Rfd; these parameters are summarized in Supplementary Table S2). The quenching kinetics protocol allowed to detect the shifts in the ChlF curves of Col-0 and C24 salt-stressed plants as early as 1 h after moderate or severe stress (Supplementary Figure S3).

We then selected six ChlF parameters that reflect the photosynthetic function of PSII (Lazar, 2015): the maximum quantum yield of PSII photochemistry in the dark-adapted (F_v/F_m), and the light-adapted (F_v'/F_m') states, the coefficient
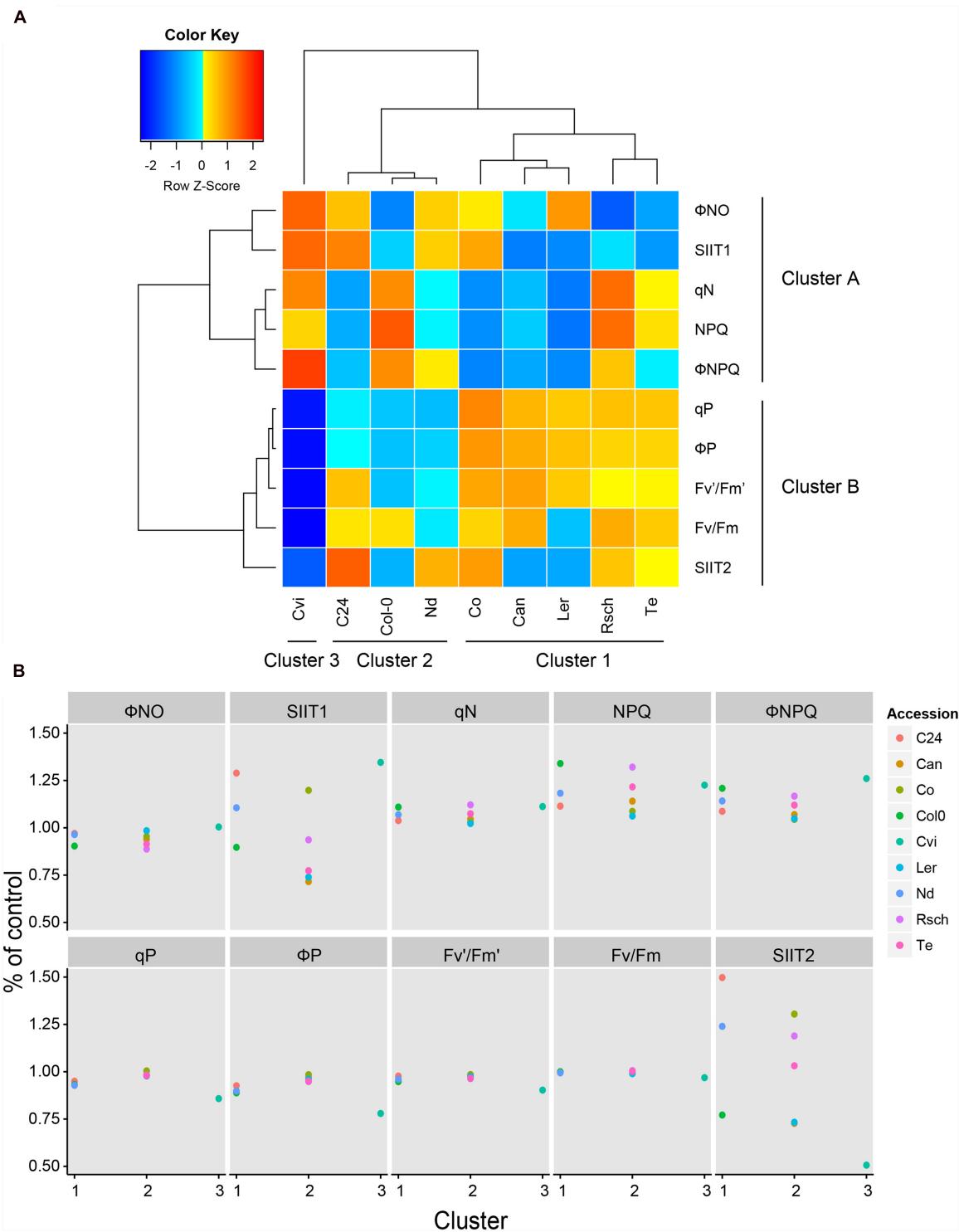
FIGURE 7 | Clustering of ChlF parameters with SIIT values revealed the early and late responses to salt stress. (A) Using the Ward Linkage method, the shoot-ion independent tolerance (SIIT) values, $SIIT_1$, $SIIT_2$ and eight ChlF traits, measured at the highest photon irradiance (L4) at day 7, were clustered. Three cluster groups were identified for the nine accessions, while two clusters were observed to form across the phenotypic traits. The trait values for individual accessions presented in the heat map were normalized by the z-Fisher transformation per trait. **(B)** The effect of salt treatment on ChlF parameters and SIIT values across the three cluster groups identified in (A). $SIIT_1$, $SIIT_2$ and the selected ChlF parameters were calculated relative to control conditions and divided by the overall average per trait. Values represent average of eight biological replicates per accession and treatment.

of photochemical quenching that estimates the fraction of open PSII reaction centers (qP), the actual quantum yield of PSII photochemistry in the light-adapted state, the proportion of light absorbed by the chlorophyll associated with PSII that is used in photochemistry (ΦP), the quantum yield of constitutive non-light-induced dissipation consisting of ChlF emission and heat dissipation (ΦNO) and the quantum yield of regulatory light-induced heat dissipation (ΦNPQ) for days 0, 1, 2, 3, 4, 7, and 11 of the phenotyping period (**Figure 4**). The traits related to maximum quantum yield (F_v/F_m and F_v'/F_m') were not significantly different between control and salt-treated plants (**Figure 4**; Supplementary Figure S4; Supplementary Table S2), suggesting that PSII was not damaged during the course of the experiment. The other four ChlF parameters (qP, ΦP, ΦNO, and ΦNPQ) varied with time in both control and salt-stressed conditions (**Figure 4**).

Change in nine ChlF parameters, with respect to control conditions, was used to indicate the effect of salt stress on quenching processes and PSII efficiency (Supplementary Table S2; Supplementary Figure S4): traits measured in the light-adapted state were most affected (L1, L2, L3, L3, and Lss), while traits measured in the dark-adapted state (D1, D2, and D3) did not vary between treatments or accessions, except in the case of NPQ, qN, RFD, and ΦNPQ under severe stress (Supplementary Figure S4). F_v'/F_m' was unchanged in all salt treatments and in both accessions and ΦNO displayed only slight changes indicating that those parameters were robust in response to salt stress. With this protocol we were able to detect rapid changes in ΦP, qP, PQ, NPQ, qN, RFD, and ΦNPQ in C24, but not in Col-0, after only 1 day of salt treatment (Supplementary Figure S4). To explore the ChlF parameters even further, we used PCA to classify the observed trends (Supplementary Figure S5). PCA performed on the 20 ChlF traits, under the eight different adapted states and saturating pulses (L1 to L4, Lss and D1 to D3), showed that the five PCs explained 85% of the variation (Supplementary Figure S5). PC1 described accession-specific trends and PC2 contained traits relevant to treatment with salt (Supplementary Table S3).

Natural Variation of Growth-Related Traits was Quantified in Response to Salt Stress

Because moderate salt stress elicited significant changes between control and salt-treated plants without symptoms of severe stress, it was used to investigate natural variation among the nine accessions of *Arabidopsis thaliana* (**Figure 5A**). Plants were cultivated with a longer light period at higher temperatures and lower photon irradiance than the initial conditions for Col-0 and C24. Phenotyping of the plants was conducted in the same manner through 7 days after salt treatment using RGB (**Figure 5A**) and ChlF imaging with the light curve protocol (Supplementary Figure S6). The rosette area of Te was the most significantly reduced by salt stress starting from day 3, while Col-0, Can, Co and Ler showed significant reductions later. Rosette areas of C24, Nd, and Cvi were not significantly reduced (**Figure 5B**). Natural variation was

evident across all SIIT values in both intervals, indicating differences in salinity tolerance among the nine accessions. Lower $SIIT_2$ than $SIIT_1$ values indicated that plants became less tolerant over time; however, this difference was only significant for Cvi. This demonstrates that we were able to assess natural variation, not only in the growth reduction magnitude, but also in the timing of the responses to salt stress.

Light Curve Chlorophyll Fluorescence Imaging Captured Early Responses to Salt Stress

ChlF parameters were measured at four photon irradiances and were calculated as described previously (Supplementary Table S2). Differences in most ChlF parameters between control and salt-stressed plants were observed within 24 h of introducing salt at the highest photon irradiance (**Figure 6**). We found that the maximum quantum yield of photosynthesis in the dark-adapted state (F_v/F_m) was not affected by salt stress in Col-0. Rapid responses to salt stress were observed in photochemical and non-photochemical quenching, as ΦNPQ rapidly increased, followed by a decrease in non-regulatory heat dissipation ΦNO. The increase in heat dissipation via xanthophyll-mediated non-photochemical quenching (ΦNPQ) coincided with a significant decrease in the photochemical quenching coefficient (qP) and inhibition of the PSII operating efficiency (ΦP). Maximum quantum yield in the light-adapted state (Fv'/Fm') decreased in response to salt stress, but not as severely as did the other ChlF traits (Supplementary Figure S6B; Supplementary Table S4).

Because the highest actinic photon irradiance (L4) provided the most discriminative power for the quantification of early salt-induced changes in ChlF parameters, it was chosen to assess the natural variation in photosynthetic activity. Comparison among the nine accessions using the six key ChlF traits revealed that F_v/F_m did not differ between control and salt-treated plants, with the exception of Cvi (**Figure 6**; Supplementary Table S4). Increased ΦNPQ, upon exposure to salt stress, was observed to varying degrees (**Figure 6**; Supplementary Table S4). We also observed natural variation in the salt-induced decrease of ΦP, qP, and ΦNO over time. Overall, we observed that salt stress results in rapid and substantial increase in non-photochemical processes (i.e., the dissipation of heat in the PSII antennae), which correlates with reduced PSII quantum efficiency and photochemical quenching under stress (**Figure 6**; Supplementary Figure S6C).

Eight ChlF traits and two SIIT values were used to cluster the nine accessions using the Ward linkage method (**Figure 7A**). The accessions clustered into three groups (1–3) while the phenotypic traits were classified into two clusters (A and B). Clustering of traits revealed that non-photochemical quenching-related parameters (NPQ, ΦNO, ΦNPQ, and qN) are more prominent during the early stage (cluster A), while quantum yield-related parameters (F_v/F_m, F_v'/F_m', ΦP, and qP) corresponded to the later stage of exposure to salt stress (cluster B). Eight accessions were grouped into two clusters (clusters 1 and 2) and distinct responses of Cvi placed it separately (cluster

3; **Figure 7**). Accessions in cluster 2 (C24, Nd, and Col-0) showed the least pronounced responses to salt stress during the early phase of exposure to salt stress in terms of $SIIT_1$ (**Figure 5C**) and significant decline in photosynthetic activity (ΦP and qP; **Figures 6** and **7**; Supplementary Table S4). Finally, the accessions belonging to cluster 1 (Rsch, Te, Ler, Can, and Co) were characterized by rapid reduction in growth rate in the early phase of salt stress (**Figure 5C**) and less prominent changes in ChlF parameters (**Figures 6** and **7**; Supplementary Table S4). Hence, using ChlF parameters and SIIT values, we were able to distinguish between the processes affected in the early and late responses of plants to salt stress.

DISCUSSION

Recent advances in high-throughput phenotyping have allowed the parallel screening of multiple quantitative traits measuring plant growth and performance under stress conditions. In this study, we used RGB and ChlF measures, with rosette coloring, to dissect the complex responses of plants to salt stress. We developed a phenotyping protocol to monitor early physiological changes in response to salt stress involving growth, rosette morphology and photosynthetic performance. To make these evaluations, we determined the RGB, greenness and ChlF traits most responsive to salinity. The 60% soil-water content (100 mM NaCl) was the most suitable condition for studying early plant responses to salt stress without causing growth arrest (**Figure 2**) or premature leaf senescence (**Figure 3**).

Investigation of the RGB images revealed that salt stress caused little change in rosette morphology (**Figure 2B**; Supplementary Table S1); however, these phenotypes should not be overlooked as leaf slenderness and rosette compactness are known to play important roles in heat dissipation and transpiration rate (Bridge et al., 2013). Similarly, differences in rosette greenness were predominantly related to development and accession rather than to treatment with salt (**Figure 3**). Nevertheless, we did observe an increase in darker hues in plants grown in severe and moderate salt stress conditions than in those grown in mild stress, which could be due to accumulation of anthocyanin (Van Oosten et al., 2013). Pronounced changes in both lighter and darker hues were previously reported to occur in the later phases of salt stress response than the phenotyping period used in this study (Ben Abdallah et al., 2016). Therefore, color segmentation and quantification of green hues could provide valuable information regarding plant development and stress-related responses, especially when combined with quantitative pigment-content analysis.

We used automated image-processing pipelines to examine the salt-induced changes in rosette area over time by fitting a linear function to describe the growth over time (**Figures 2C** and **5B**; Supplementary Table S1). We used $SIIT_1$ and $SIIT_2$ as indicators of plant salinity tolerance in two time intervals (**Figures 2C** and **5B**) finding lower $SIIT_2$ than $SIIT_1$ values all accessions studied. The $SIIT_2$ values of Cvi decreased dramatically, potentially due to its early increase

in non-photochemical processes, which are represented by the traits NPQ, ΦNPQ and qN, with concurrent decreases in photochemical efficiency, represented by the traits F_v'/F_m', ΦP, and qP. These steps are usually followed by increase in constitutive non-light-induced dissipation (ΦNO) and a drop in the maximal quantum efficiency of PSII in dark-adapted state (F_v/F_m; Mishra et al., 2012). The observed trend of lower $SIIT_2$ values could be due to accumulation of ions in photosynthetic tissues, putting additional constraint on photosynthesis and subsequently plant growth (Munns and Tester, 2008).

Although, maximum quantum yield was commonly used for assessing plant performance under stress (Jansen et al., 2009; Bresson et al., 2015), we found that F_v/F_m seems to be a robust parameter (**Figures 4** and **6**), being affected only under severe stress and not reflecting early salt stress responses, which is in agreement with previous reports (Baker and Rosenqvist, 2004). Other parameters quantifying photochemical and non-photochemical processes displayed more dynamic responses to salt stress (**Figures 4** and **6**). To identify the ChlF traits most responsive to salt stress, we performed PCA resulting in extracting PC1 and PC2, corresponding to the differences observed between treatment and accession, respectively (Supplementary Figure S5). According to PC1, F_o, F_m, F_v, F_t, and F_p were the traits that changed the most in response to salt treatment. Based on traits contained in PC2, we found that the differences observed in F_v/F_m, NPQ, qN, and qP were more accession-specific (Supplementary Table S3). Based on our results, early ΦNPQ, qP, and ΦP responses to salt stress could be used to distinguish salt-tolerant from salt-sensitive plants (**Figures 4** and **6**). Similar results have been found for studies of drought and salt stress (Baker and Rosenqvist, 2004; Stepien and Johnson, 2009; Mishra et al., 2012). In conclusion, we observed largely different ChlF responses to salt stress among the accessions, evidencing that different accessions use different strategies to tolerate salt stress.

In this study, we demonstrated that phenotyping multiple quantitative traits in one experimental setup can provide new insights into the dynamics of plant responses to stress. These traits can be used to assess plant natural variation and to cluster accessions based on the magnitude and timing of their response to stress (**Figure 7**). Our work identified a set of phenotypes that can serve as markers for early responses to salt stress. These phenotypic markers can be used to study mutant populations, natural diversity panels and responses to other stress conditions, such as drought, cold or nutrient-deficiency, which would reveal the scope of their influence on tolerance to stress. Integrating thermal imaging into the phenotyping pipeline, along with quantifying water-use and transpiration-use efficiency, would provide a more comprehensive understanding of plant responses and development under stress. The protocol presented here can also be used to study non-model plants and crop species with more complicated 3D morphology, ultimately capturing a broad range of phenotypic traits. These traits could then be used in combination with forward genetics studies to identify genes underlying early responses to salt stress with the goal of providing new target genes for crop improvement.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01414

REFERENCES

Baker, N. R. (2008). Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annu. Rev. Plant Biol.* 59, 89–113. doi: 10.1146/annurev.arplant.59.032607.092759

Baker, N. R., and Rosenqvist, E. (2004). Applications of chlorophyll fluorescence can improve crop production strategies: an examination of future possibilities. *J. Exp. Bot.* 55, 1607–1621. doi: 10.1093/jxb/erh196

Ben Abdallah, S., Aung, B., Amyot, L., Lalin, I., Lachaal, M., Karray-Bouraoui, N., et al. (2016). Salt stress (NaCl) affects plant growth and branch pathways of carotenoid and flavonoid biosyntheses in Solanum nigrum. *Acta Physiol. Plant.* 38:72. doi: 10.1007/s11738-016-2096-8

Berger, B., de Regt, B., and Tester, M. (2012). Trait dissection of salinity tolerance with plant phenomics. *Methods Mol. Biol.* 913, 399–413. doi: 10.1007/978-1-61779-986-0_27

Bresson, J., Vasseur, F., Dauzat, M., Koch, G., Granier, C., and Vile, D. (2015). Quantifying spatial heterogeneity of chlorophyll fluorescence during plant growth and in response to water stress. *Plant Methods* 11:23. doi: 10.1186/s13007-015-0067-5

Brestic, M., and Zivcak, M. (2013). "PSII fluorescence techniques for measurement of drought and high temperature stress signal in crop plants: protocols and applications," in *Molecular Stress Physiology of Plants*, eds R. G. Rout and B. A. Das (New Delhi: Springer), 87–131.

Bridge, L. J., Franklin, K. A., and Homer, M. E. (2013). Impact of plant shoot architecture on leaf cooling: a coupled heat and mass transfer model. *J. R. Soc. Interface* 10:20130326. doi: 10.1098/rsif.2013.0326

Brown, T. B., Cheng, R., Sirault, X. R., Rungrat, T., Murray, K. D., Trtilek, M., et al. (2014). Traitcapture: genomic and environment modelling of plant phenomic data. *Curr. Opin. Plant Biol.* 18, 73–79. doi: 10.1016/j.pbi.2014.02.002

Cabot, C., Sibole, J. V., Barcelo, J., and Poschenrieder, C. (2014). Lessons from crop plants struggling with salinity. *Plant Sci.* 226, 2–13. doi: 10.1016/j.plantsci.2014.04.013

Campbell, M. T., Knecht, A. C., Berger, B., Brien, C. J., Wang, D., and Walia, H. (2015). Integrating image-based phenomics and association analysis to dissect the genetic architecture of temporal salinity responses in rice. *Plant Physiol.* 168, 1476–1489. doi: 10.1104/pp.15.00450

Chaves, M. M., Flexas, J., and Pinheiro, C. (2009). Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Ann. Bot.* 103, 551–560. doi: 10.1093/Aob/Mcn125

Chen, T. W., Kahlen, K., and Stutzel, H. (2015). Disentangling the contributions of osmotic and ionic effects of salinity on stomatal, mesophyll, biochemical and light limitations to photosynthesis. *Plant Cell Environ.* 38, 1528–1542. doi: 10.1111/pce.12504

Dhondt, S., Wuyts, N., and Inze, D. (2013). Cell to whole-plant phenotyping: the best is yet to come. *Trends Plant Sci.* 18, 433–444. doi: 10.1016/J.Tplants.2013.04.008

Fricke, W., Akhiyarova, G., Wei, W. X., Alexandersson, E., Miller, A., Kjellbom, P. O., et al. (2006). The short-term growth response to salt of the developing barley leaf. *J. Exp. Bot.* 57, 1079–1095. doi: 10.1093/Jxb/Erj095

Ghanem, M. E., Marrou, H., and Sinclair, T. R. (2015). Physiological phenotyping of plants for crop improvement. *Trends Plant Sci.* 20, 139–144. doi: 10.1016/j.tplants.2014.11.006

Godfray, H. C., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., et al. (2010). Food security: the challenge of feeding 9 billion people. *Science* 327, 812–818. doi: 10.1126/science.1185383

Hairmansis, A., Berger, B., Tester, M., and Roy, S. J. (2014). Image-based phenotyping for non-destructive screening of different salinity tolerance traits in rice. *Rice* 7:16. doi: 10.1186/s12284-014-0016-3

Hannah, M. A., Wiese, D., Freund, S., Fiehn, O., Heyer, A. G., and Hincha, D. K. (2006). Natural genetic variation of freezing tolerance in Arabidopsis. *Plant Physiol.* 142, 98–112. doi: 10.1104/pp.106.081141

Henley, W. J. (1993). Measurement and interpretation of photosynthetic light-response curves in algae in the context of photoinhibition and diel changes. *J. Phycol.* 29, 729–739. doi: 10.1111/j.0022-3646.1993.00729.x

Humplik, J. F., Lazar, D., Husickova, A., and Spichal, L. (2015). Automated phenotyping of plant shoots using imaging methods for analysis of plant stress responses - a review. *Plant Methods* 11:29. doi: 10.1186/s13007-015-0072-8

James, R. A., Munns, R., Von Caemmerer, S., Trejo, C., Miller, C., and Condon, T. (2006). Photosynthetic capacity is related to the cellular and subcellular partitioning of Na, K and Cl in salt-affected barley and durum wheat. *Plant Cell and Environment* 29, 2185–2197. doi: 10.1111/J.1365-3040.2006.01592.X

Jansen, M., Gilmer, F., Biskup, B., Nagel, K. A., Rascher, U., Fischbach, A., et al. (2009). Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via GROWSCREEN FLUORO allows detection of stress tolerance in *Arabidopsis thaliana* and other rosette plants. *Funct. Plant Biol.* 36, 902–914. doi: 10.1071/Fp09095

Junker, A., Muraya, M. M., Weigelt-Fischer, K., Arana-Ceballos, F., Klukas, C., Melchinger, A. E., et al. (2015). Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. *Front. Plant Sci.* 5:770. doi: 10.3389/fpls.2014.00770

Lazar, D. (2015). Parameters of photosynthetic energy partitioning. *J. Plant Physiol.* 175, 131–147. doi: 10.1016/j.jplph.2014.10.021

Longenberger, P. S., Smith, C. W., Duke, S. E., and McMichael, B. L. (2009). Evaluation of chlorophyll fluorescence as a tool for the identification of drought tolerance in upland cotton. *Euphytica* 166, 25–33. doi: 10.1007/s10681-008-9820-4

Ma, S. S., Gong, Q. Q., and Bohnert, H. J. (2006). Dissecting salt stress pathways. *J. Exp. Bot.* 57, 1097–1107. doi: 10.1093/Jxb/Erj098

Maxwell, K., and Johnson, G. N. (2000). Chlorophyll fluorescence - a practical guide. *J. Exp. Bot.* 51, 659–668. doi: 10.1093/jexbot/51.345.659

Mishra, K. B., Iannacone, R., Petrozza, A., Mishra, A., Armentano, N., La Vecchia, G., et al. (2012). Engineered drought tolerance in tomato plants is reflected in chlorophyll fluorescence emission. *Plant Sci.* 182, 79–86. doi: 10.1016/j.plantsci.2011.03.022

Munns, R., James, R. A., Sirault, X. R., Furbank, R. T., and Jones, H. G. (2010). New phenotyping methods for screening wheat and barley for beneficial responses to water deficit. *J. Exp. Bot.* 61, 3499–3507. doi: 10.1093/jxb/erq199

Munns, R., and Tester, M. (2008). Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* 59, 651–681. doi: 10.1146/annurev.arplant.59.032607.092911

Muranaka, S., Shimizu, K., and Kato, M. (2002). A salt-tolerant cultivar of wheat maintains photosynthetic activity by suppressing sodium uptake. *Photosynthetica* 40, 509–515. doi: 10.1023/A:1024335515473

Murchie, E. H., and Lawson, T. (2013). Chlorophyll fluorescence analysis: a guide to good practice and understanding some new applications. *J. Exp. Bot.* 64, 3983–3998. doi: 10.1093/jxb/ert208

Oxborough, K. (2004). Imaging of chlorophyll a fluorescence: theoretical and practical aspects of an emerging technique for the monitoring of photosynthetic performance. *J. Exp. Bot.* 55, 1195–1205. doi: 10.1093/jxb/erh145

Oxborough, K., and Baker, N. R. (1997). Resolving chlorophyll a fluorescence images of photosynthetic efficiency into photochemical and non-photochemical components - calculation of qP and Fv'/Fm' without measuring Fo'. *Photosynth. Res.* 54, 135–142. doi: 10.1023/A:1005936823310

Rajendran, K., Tester, M., and Roy, S. J. (2009). Quantifying the three main components of salinity tolerance in cereals. *Plant Cell Environ.* 32, 237–249. doi: 10.1111/J.1365-3040.2008.01916.X

Rascher, U., Liebig, M., and Luttge, U. (2000). Evaluation of instant light-response curves of chlorophyll fluorescence parameters obtained with a portable chlorophyll fluorometer on site in the field. *Plant Cell and Environ.* 23, 1397–1405. doi: 10.1046/j.1365-3040.2000.00650.x

Roy, S. J., Negrao, S., and Tester, M. (2014). Salt resistant crop plants. *Curr. Opin. Biotechnol.* 26, 115–124. doi: 10.1016/j.copbio.2013.12.004

Sirault, X. R. R., James, R. A., and Furbank, R. T. (2009). A new screening method for osmotic component of salinity tolerance in cereals using infrared thermography. *Funct. Plant Biol.* 36, 970–977. doi: 10.1071/Fp09182

Stephan, A. B., and Schroeder, J. I. (2014). Plant salt stress status is transmitted systemically via propagating calcium waves. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6126–6127. doi: 10.1073/Pnas.1404895111

Stepien, P., and Johnson, G. N. (2009). Contrasting responses of photosynthesis to salt stress in the glycophyte Arabidopsis and the halophyte Thellungiella: role of the plastid terminal oxidase as an alternative electron sink. *Plant Physiol.* 149, 1154–1165. doi: 10.1104/Pp.108.132407

Tester, M., and Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science* 327, 818–822. doi: 10.1126/science.1183700

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20260–20264. doi: 10.1073/Pnas.1116437108

Van Oosten, M. J., Sharkhuu, A., Batelli, G., Bressan, R. A., and Maggio, A. (2013). The *Arabidopsis thaliana* mutant air1 implicates SOS3 in the regulation of anthocyanins under salt stress. *Plant Mol. Biol.* 83, 405–415. doi: 10.1007/s11103-013-0099-z

Zhang, X., Hause, R. J. Jr., and Borevitz, J. O. (2012). Natural genetic variation for growth and development revealed by high-throughput phenotyping in *Arabidopsis thaliana*. *G3* (Bethesda) 2, 29–34. doi: 10.1534/g3.111.001487

Genotype × Environment Interactions of Yield Traits in Backcross Introgression Lines Derived from *Oryza sativa* cv. Swarna/*Oryza nivara*

Divya Balakrishnan *, Desiraju Subrahmanyam, Jyothi Badri, Addanki Krishnam Raju, Yadavalli Venkateswara Rao, Kavitha Beerelli, Sukumar Mesapogu, Malathi Surapaneni, Revathi Ponnuswamy, G. Padmavathi, V. Ravindra Babu and Sarla Neelamraju

Crop Improvement Section, ICAR- National Professor Project, ICAR- Indian Institute of Rice Research, Hyderabad, India

Advanced backcross introgression lines (BILs) developed from crosses of *Oryza sativa* var. Swarna/*O. nivara* accessions were grown and evaluated for yield and related traits. Trials were conducted for consecutive three seasons in field conditions in a randomized complete block design with three replications. Data on yield traits under irrigated conditions were analyzed using the Additive Main Effect and Multiplicative Interaction (AMMI), Genotype and Genotype × Environment Interaction (GGE) and modified rank-sum statistic (*YSi*) for yield stability. BILs *viz.,* G3 (14S) and G6 (166S) showed yield stability across the seasons along with high mean yield performance. G3 is early in flowering with high yield and has good grain quality and medium height, hence could be recommended for most of the irrigated locations. G6 is a late duration genotype, with strong culm strength, high grain number and panicle weight. G6 has higher yield and stability than Swarna but has Swarna grain type. Among the varieties tested DRRDhan 40 and recurrent parent Swarna showed stability for yield traits across the seasons. The component traits thousand grain weight, panicle weight, panicle length, grain number and plant height explained highest genotypic percentage over environment and interaction factors and can be prioritized to dissect stable QTLs/ genes. These lines were genotyped using microsatellite markers covering the entire rice genome and also using a set of markers linked to previously reported yield QTLs. It was observed that wild derived lines with more than 70% of recurrent parent genome were stable and showed enhanced yield levels compared to genotypes with higher donor genome introgressions.

Keywords: BILs, stability, AMMI, GGE, *Oryza nivara*, yield traits

INTRODUCTION

Improving rice production per unit area and per unit time will be a major challenge in future due to the expanding population of rice consumers in the world. The average yield of existing cultivars reached a plateau and now research is directed toward wild relatives of *Oryza* to explore novel genes that can improve yield traits. Wild relatives were widely explored as donors for stress resistance and less exploited for yield improvement because of non-preferable agronomic traits linked with them. Wild rice genotypes provide a diverse range of allelic variation due to their adaption to a wide range of environmental conditions. Wild and related genotypes are valuable resources to explore novel variations to widen the genetic background of cultivated rice (Brar and Khush, 1997; Tanksley and McCouch, 1997; Swamy and Sarla, 2008; Wickneswari et al., 2012). Introgression of chromosomal segments from wild species into cultivated species can also generate de-novo variations in the new genetic background (Wang et al., 2005).

Back cross introgression lines developed from wild and adapted genotypes are useful in diversifying existing germplasm in more usable form and also in discovering novel genes/QTLs. As BILs have maximum genome of recurrent parent with few donor segments, it is advantageous to use them for precise estimation of quantitative traits. Fixed BILs can be replicated and can be used to study their environment interactions. The evaluation of the BILs for stability is very important especially when it is derived from an interspecific cross, as it takes more time to attain stability in the new back ground. Utilization of stable BILs will accelerate varietal development due to the presence of novel genes in an adapted parental background (Jeuken and Lindhout, 2004).

As grain yield is a complex quantitative trait, with high environmental interaction; selection of genotypes based on performance in single environment is not effective for varietal identification (Shrestha et al., 2012). It is essential to carry out selection based on yield stability evaluation than average performance in multiple environment conditions (Kang, 1993; Tariku et al., 2013; Islam et al., 2015). Selection of genotypes for stability and adaptability is required prior to recommendation in case of a crop such as rice which is grown in diverse ecologies. Stability is the suitability of a variety over a wide range of environments while adaptability is the better survival of a genotype over any specific environment. This can be attained through either genetic or physiological homeostasis of genotypes for environmental fluctuations (Singh and Narayanan, 2006). For cultivation in large area it is stability for yield traits which is desirable but for achieving maximum productivity, it is adaptability to best target environments that is preferred.

Effects of genotype, environment and genotype × environment interaction determine the phenotypic performance and its general and specific adaptation to different environments (Falconer and Mackey, 1996). This information is required for planning better selection strategies and to identify the best environment to select genotypes for grain yield (Gauch and Zobel, 1996; Kang, 1998). Several studies have been conducted on stability performance for grain yield of rice for different

ecosystems (Cooper et al., 1999; Wade et al., 1999; Ouk et al., 2007; Anandan et al., 2009; Kumar et al., 2012; Tariku et al., 2013; Liang et al., 2015; Katsura et al., 2016). Many such studies showed that genotype × environment interaction was more significant than genotypic main effects (Henderson et al., 1996; Cooper and Somrith, 1997; Wade et al., 1997, 1999; Cooper et al., 1999; Inthapanya et al., 2000).

There are several methods to study stability and genotype × environment interactions of traits through conventional analysis. Different models were proposed on stability variance, ecovalence, regression coefficient analysis or principal component analysis (PCA) (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966; Perkins and Jinks, 1968; Freeman and Perkins, 1971; Shukla, 1972; Kang, 1993). Kang (1993) proposed yield stability static (Ysi) by combining yield and stability as a single selection criterion by modifying rank Sum method. However, additive main effects and multiplicative interaction (AMMI) model and the genotype main effects and genotype × environment interaction effects (GGE) model are more popular methods. This method is followed to quantify the genotype environment interaction through PCA and graphical representation and has been widely applied in the multi-environment cultivar trials (Kempton, 1984; Crossa et al., 1990; Gauch and Zobel, 1997).

A panel of 14 BILs derived from Swarna/ *Oryza nivara* was studied along with 9 high yielding rice varieties of different duration and these 23 lines were screened in three seasons. Genotypic characterization of these BILs was conducted with genome wide polymorphic markers and markers linked to yield QTLs. The objectives of this study were (1) to identify the yield potential of backcross introgression lines in comparison with existing popular varieties (2) to identify stable high yielding BILs and their parental genome percentage (3) to prioritize the component traits important for further genetic dissection and improvement.

MATERIALS AND METHODS

Location

Field experiments were conducted at Indian Institute of Rice Research, Hyderabad (17° 19′ N and 78° 29′ E) at an altitude of 549 m above mean sea level during two wet seasons *Kharif*-2013(E1), (*Kharif*-2014) (E2) and one dry season *Rabi*-2014(E3). Crop was grown in alkaline vertisol with a pH of 7.94 at irrigated field conditions. Details of meteorological conditions during the crop growth period are presented in **Table 1**.

Plant Material

Studies were conducted at IIRR to develop wild introgression lines between *O. sativa cv.* Swarna and accessions of wild relative, *Oryza nivara* (Kaladhar et al., 2008; Swamy et al., 2011). The developed BILs were advanced to BC_2F_6 generation and further purified by single panicle selection method upto BC_2F_8. From two sets of BILs consisting of 94 lines from Swarna / *O. nivara* (81848) (S lines) and 104 lines from Swarna / *O. nivara* (81832) (K lines), a panel of 14 BILs at BC_2F_8 generation were selected based on their preferable phenotypic traits (**Table 2**).

TABLE 1 | Weather parameters during crop season.

Season	Months	Temperature (°C)			R.H. (%)		Rainfall (mm)	Rainy days	Sunshine (Hrs.)	Wind speed (Km/Hr)	Evaporation (mm)	Crop stage
		Max.	Min.	Mean temp.	I	II						
Kharif-2013	Jul-13	32.64	23.92	28.25	81.83	56.83	150.20	9.00	4.56	10.55	4.69	Sowing
	Aug-13	28.49	21.94	25.21	89.90	75.93	158.10	1.00	3.39	6.36	4.02	Transplanting
	Sep-13	31.05	20.59	25.82	87.07	64.33	110.60	8.00	5.71	3.05	4.69	Vegetative stage
	Oct-13	29.99	19.73	24.86	88.77	63.33	253.20	9.00	5.45	2.87	3.86	Heading / observations on Pn
	Nov-13	28.42	14.38	21.40	86.27	50.43	31.00	2.00	6.66	1.72	2.69	Harvesting
	Mean	30.54	21.54	26.04	86.89	65.11	168.03	6.75	4.78	5.71	4.32	
Rabi-2014	Dec-13	28.02	10.09	19.05	83.10	36.58	0.00	0.00	8.87	1.75	2.70	Sowing
	Jan-14	28.69	13.25	20.97	84.74	40.29	0.00	0.00	8.17	2.49	3.07	Transplanting
	Feb-14	31.20	16.55	23.88	78.36	32.75	0.00	0.00	8.99	2.93	4.62	Vegetative stage
	Mar-14	33.23	20.36	26.80	79.58	36.39	56.80	5.00	7.35	2.69	4.61	Vegetative stage
	Apr-14	37.60	22.04	29.82	76.73	36.07	72.60	2.00	7.72	2.02	6.02	Heading / observations on Pn
	May-14	37.74	23.87	30.81	66.10	33.84	40.10	3.00	8.39	3.47	7.11	Heading / observations on Pn
	Jun-14	37.02	24.60	30.81	68.57	45.07	53.60	3.00	7.90	10.17	8.14	Harvesting
	Mean	33.69	19.22	26.45	77.10	35.87	44.62	2.00	8.13	2.72	5.09	
Kharif-2014	Jul-14	31.80	23.60	27.70	82.32	62.40	108.00	8.00	3.20	12.40	4.00	Sowing
	Aug-14	32.00	23.30	27.65	85.90	62.70	184.10	10.00	5.40	7.10	3.00	Transplanting
	Sep-14	30.30	22.40	26.35	87.43	60.80	60.60	7.00	5.10	5.20	3.00	Vegetative stage
	Oct-14	31.60	19.90	25.75	83.45	49.80	69.20	3.00	6.40	2.40	4.80	Heading / observations on Pn
	Nov-14	30.60	16.80	23.70	79.23	40.90	10.60	1.00	7.00	1.80	4.70	Harvesting
	Mean	31.26	21.20	26.23	83.67	55.32	86.50	5.80	5.42	5.78	3.90	

As BILs have a range of flowering duration from 77 to 120 days popular varieties IR64, Jaya, MTU1010, MTU1081, NLR34449, Sahbhagi Dhan, Swarna, Tellahamsa, and Tulasi with different flowering duration were grown as checks under irrigated conditions. These BILs were evaluated for yield and related traits in irrigated conditions over a period of three seasons (2013–2014) along with checks. As there is considerable variation in the duration among the BILs, per day productivity was computed to compare genotypes with different duration.

Field Experimental Details

Seeds were sown in nursery beds, and 25 days old seedlings were transplanted, with single seedling per hill in all the field trials. The planting density was 33.3 hills m^{-2}, with 20 cm row spacing and 15 cm intra-row spacing with five rows of 21 plants each constituting a replication. Normal package of practice and fertilizer application was followed; weeds, insects, and diseases were controlled by using standard herbicides and pesticides as required to avoid yield loss. The experimental plots were arranged in a randomized complete block design with three replications each containing 105 plants. These same parameters were followed uniformly across the seasons.

Phenotyping

These genotypes were screened for various yield contributing traits in all the seasons following Standard Evaluation System

(IRRI, 2013). The observations on yield and morpho-agronomic traits were recorded from the field experiments.

Statistical Analysis

Analysis of variance was computed for individual environment, then a combined analysis of variance was performed, considering both environments and genotypes as fixed using PB tools (Version 1.4, http://bbi.irri.org/products) and R (R Core Team, 2012). Significance of all effects was tested against mean square of error. The performance of BILs was tested over three seasons and was assessed using stability models *viz*, (1) yield-stability statistic (*YSi*) (Kang, 1993), (2) Additive Main effects and Multiplicative Interaction (AMMI) (Gauch and Zobel, 1997), and (3) GGE Biplot or Site Regression model (Yan and Kang, 2003). These models were used to interpret and visualize the stability and GEI patterns. In the AMMI model, only the GEI term is absorbed in the multiplicative component, whereas in the GGE model, the main effects of genotypes (G) plus the GEI are absorbed into the multiplicative component. Yield-stability (*YSi*) statistic was developed by Kang (1993) to be used as a selection criterion when G × E interaction is significant. The stability-variance was determined following modified Shukla's (1972) method and genotypes with significant stability variance were considered unstable. The stability variance was integrated with yield to obtain the *YSi* statistic as outlined by Kang and Magari (1995). Simultaneous selection of high yielding and stable genotypes is possible through this method.

TABLE 2 | Description of the genotypes used in the study.

Code	Genotype	Type	Pedigree	Year of release	Origin
G1	14_3	BIL	Swarna / *O. nivara*	–	IIRR, India
G2	148S	BIL	Swarna / *O. nivara*	–	IIRR, India
G3	14S	BIL	Swarna / *O. nivara*	–	IIRR, India
G4	166_1S	BIL	Swarna / *O. nivara*	–	IIRR, India
G5	166_2S	BIL	Swarna / *O. nivara*	–	IIRR, India
G6	166S	BIL	Swarna / *O. nivara*	–	IIRR, India
G7	248S (DRR Dhan 40)	BIL	Swarna / *O. nivara*	2013	IIRR, India
G8	24K	BIL	Swarna / *O. nivara*	–	IIRR, India
G9	250K	BIL	Swarna / *O. nivara*	–	IIRR, India
G10	3_1K	BIL	Swarna / *O. nivara*	–	IIRR, India
G11	65S	BIL	Swarna / *O. nivara*	–	IIRR, India
G12	70S	BIL	Swarna / *O. nivara*	–	IIRR, India
G13	75S	BIL	Swarna / *O. nivara*	–	IIRR, India
G14	7K	BIL	Swarna / *O. nivara*	–	IIRR, India
G15	Swarna	*indica*	Vasistha/Mahsuri	1982	ARI, Maruteru
G16	IR64	*indica*	IR5657-33-2-1/ IR2061-465-1-5-5	1991	IRRI
G17	Jaya	*indica*	TN-1/T-141	1968	
G18	MTU1010	*indica*	Krishnaveni/IR64	2000	ARI, Maruteru
G19	MTU1081	*indica*	Ajaya/BPT5204		ARI, Maruteru
G20	NLR34449	*indica*			
G21	Sahbhagi Dhan	*indica*	IR55419-4*2/WayRarem	2009	CRRI and IRRI
G22	Tellahamsa	*indica*	HR12/TN1	1971	
G23	Tulasi	*indica*	Rasi/Finegora	1988	IIRR, India

The AMMI model (Gauch, 1988) was used in analyzing the stability and interaction for yield traits. The AMMI model is a combination of analysis of variance (ANOVA) and principal component analysis (PCA). The G × E interaction was evaluated with the AMMI model by considering the first two principal components. ANOVA model was used to analyze the trait data with main effects of genotype and environment without the interaction, then, a principal component analysis was integrated using the standardized residuals. These residuals include the experimental error and the effect of the GEI. The analytical model can be written as

$$Y_{ij} = \mu + \delta_i + \beta_j + \sum_{k=1}^{K} \lambda_k \delta_{ik} \beta_{jk} + \varepsilon_{ij}$$

Where Y_{ij} is the mean yield of ith genotype in j^{th} environment, μ is the overall mean, δ_i is the genotypic effect, β_j is the environment effect, λ_k is the singular value for PC axis k, δik is the genotype eigenvector value for PC axis n, β_{jk} is the environment eigenvector value for PC axis k and ε_{ij} is the residual error assumed to be normally and independently distributed (0, $\sigma2/r$), $\sigma2$ is the pooled error variance and r is the number of replicates.

GGE biplots display both G (genotype) and GE (genotype environment) variation (Kang, 1993) for genotype evaluation. The GGE biplot is based on the sites regression (SREG) linear bilinear model (Cornelius et al., 1996; Crossa and Cornelius, 1997; Crossa et al., 2002). The sites regression model as a multiplicative model in the bilinear terms shows the main effects of cultivars plus the cultivar × environment interaction (GGE) and the model is

$$Y_{ij} - \mu_j = \sum_{k=1}^{t} \lambda_k \delta_{ik} \beta_{jk} + \varepsilon_{ij}$$

The GGE biplot graphically represents G and GEI effect present in the multi-location trial data using environment centered data. GGE biplots were used to evaluate (1) mega environment analysis (which-won-where pattern), where genotypes can be recommended to specific mega environments. (2) Genotype evaluation, where stable specific genotypes can be recommended across all locations and (3) location evaluation, explains discriminative power of target locations for genotypes under study.

Sum of square percentage was computed as percentage of sum of squares of components of stability analysis of variance per total sum of squares to know the contribution of each component viz., genotype, environment and GEI. Correlation analysis was performed with Statistical Tool for Agricultural Research (STAR) using Pearson's correlation coefficient method. Significance levels are indicated as: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

Genotyping
Molecular screening was conducted to identify the presence of reported QTLs in the BILs and also to identify recurrent

parent genome percentage. Leaves of 20 days-old seedlings were collected from the field and CTAB (Cetyl Trimethyl Ammonium Bromide) method was followed for DNA extraction (Doyle and Doyle, 1987). Polymorphic SSR markers with genome wide distribution (**Figure 1**) from universal core genetic map (Orjuela et al., 2010) were used for genotyping (Supplementary Table 1). PCR reactions were carried out in Thermal cycler (Veriti PCR, Applied Biosystems, USA) with the total reaction volume of 10 µl containing 15 ng of genomic DNA, 1X assay buffer, 200 µM of dNTPs, 1.5 mM MgCl$_2$, 10 pmol of forward and reverse primer and 1 unit of Taq DNA polymerase (Thermo Scientific, U.S.A). PCR cycles were programmed as follows: initial denaturation at 94°C for 5 min followed by 35 cycles of 94°C for 45 s, 55°C for 30 s, 72°C for 45 s and a final extension of 10 min at 72°C. Amplified products were resolved in 4% metaphor agarose gels prepared in 0.5 X TAE buffer and electrophoresis was conducted at 120V for 2 h. Gels were stained with ethidium bromide and documented using gel documentation system (Alfa imager, U.S.A). Amplified fragments were scored for the presence (1) or absence (0) for each primer genotype combination. The SSR genotypic data generated in the population were analyzed using the software, GGT ver.2.0. The graphical representations and comparisons were made among the 23 lines on linkage group basis and also the entire genome level on individual basis.

RESULTS

Yield and Yield Related Traits

Wide range of variation was observed for yield traits among the genotypes and across the environments. Combined analysis of variance of three environmental data showed significant genotypic and genotype × environment interactions for all the traits except for 1000 grain weight where the G × E interactions were not significant. In the three environments *Kharif* 2013, *Rabi* 2014 and *Kharif* 2014 the variation in seasonal average was

observed for DFF, GY, BY PH, TN, and BM. In three seasons; broad genotypic variation was observed and genotypic average ranged for DFF (77.37 to 133.04); GY (4.73 to 24.63); BY (0.33 to 2.10); PH (65.98 to 148.43); TN (6.12 to 20.70); GN (93.70 to 314.55); PL (18.09 to 25.19); PW (0.98 to 3.87); TGW (12.50 to 26.07); SF (60.57 to 97.57); BM (11.26 to 54.12); HI (0.20 to 0.55) and per day productivity (0.04 to 0.19) among the BILs under study (**Table 3**). The data obtained from the three replications was assessed and compared with high yielding checks in each season. Considering the three season average, among BILs G3 scored highest grain yield, harvest index and per day productivity and G6 scored highest bulk yield compared to Swarna and on par with other checks. G2 was of shortest duration and showed desirable yield traits such as panicle length, 1000 grain weight, spikelet fertility and plant height compared with checks. It was early in flowering with lowest unfilled grains in all the seasons. Derived lines from G6 i.e., G5 had highest grain number, filled grains and panicle weight. G8 was identified as having highest average biomass and dry matter production among BILs, G13, and G1 for high tiller number and G14 for maximum days to maturity. High yielding check MTU1010 showed highest grain yield, bulk yield, harvest index and per day productivity than BILs for three seasons average.

The BILs were screened for seedling vigor in two seasons *Kharif* and *Rabi* in field conditions. Seedling vigor for BILs was obtained both for *Kharif* and *Rabi* season in terms of plant height and tiller number from the data taken on 40 days after transplanting and 70 days after transplanting. BILs seeds were subjected to germination test and vigor index analysis *in vitro* by paper towel method (ISTA, 1999). Seedling vigor was also assessed based on paper towel method using the data for shoot length and root length from 7 days and 14 days and germination studies. In both the seasons G2 had highest seedling vigor in terms of plant height. G14 was best for number of tillers in *Kharif* and G7 in *Rabi* season. G13(75S) showed highest vigor in terms
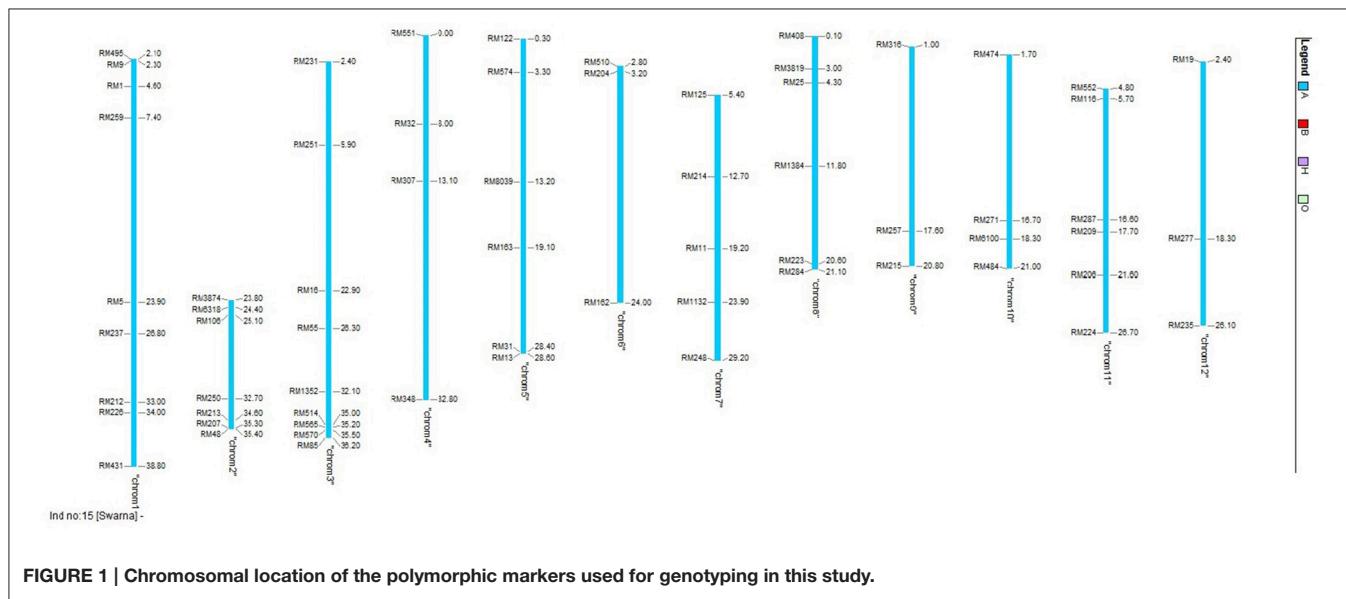
FIGURE 1 | Chromosomal location of the polymorphic markers used for genotyping in this study.

TABLE 3 | The mean performance of genotypes under the study across the seasons.

SL. NO	GENOTYPE	DFF	GY	BY	PH	TN	GN	PL	PW	TGW	SF	TDM	HI	YPD
G1	14_3	108.67	6.18	0.43	74.66	13.89	122.96	21.58	1.25	13.72	66.13	19.04	0.33	0.05
G2	148S	92.00	15.21	1.18	139.73	11.40	112.56	24.28	2.64	24.99	93.60	39.73	0.38	0.13
G3	14S	103.67	20.87	1.56	82.58	12.20	111.06	22.04	2.25	23.49	89.28	37.09	0.54	0.15
G4	166_1	112.33	11.00	1.39	95.84	9.62	169.42	22.23	2.63	18.39	79.70	34.84	0.31	0.08
G5	166_2	108.33	15.11	1.21	87.89	8.20	261.76	21.35	3.65	18.18	63.53	37.85	0.39	0.11
G6	166S	112.44	18.90	1.67	94.62	10.38	183.97	22.07	3.16	19.63	81.99	41.78	0.44	0.13
G7	248S	117.00	14.61	1.42	83.83	12.01	123.22	20.25	1.96	16.48	71.27	37.64	0.38	0.10
G8	24K	119.89	16.58	1.35	106.67	9.44	154.33	19.72	2.73	18.67	83.11	50.53	0.35	0.11
G9	250K	117.67	13.02	0.94	92.60	11.96	128.21	22.04	2.08	17.89	81.59	36.04	0.35	0.09
G10	3_1K	115.78	12.26	1.19	83.51	13.22	140.03	19.63	1.85	16.62	76.47	34.47	0.35	0.08
G11	65S	111.22	12.96	1.14	111.69	13.56	149.81	20.44	1.92	14.18	83.57	38.75	0.33	0.09
G12	70S	116.33	12.63	1.26	76.34	13.68	127.18	18.75	1.77	17.51	89.00	35.97	0.34	0.08
G13	75S	114.67	13.31	1.11	72.42	14.38	112.17	19.25	1.90	17.04	86.21	29.65	0.44	0.09
G14	7K	122.11	16.67	1.11	90.59	12.81	162.06	22.53	2.12	16.21	72.84	47.04	0.36	0.11
G15	Swarna	120.72	18.13	1.36	86.56	13.82	156.99	20.84	2.10	16.91	75.57	45.86	0.41	0.13
G16	IR64	104.11	17.71	1.50	86.53	13.53	130.00	23.17	2.37	24.30	83.72	38.52	0.50	0.15
G17	JAYA	112.78	16.94	1.62	98.22	11.27	152.03	22.42	2.56	23.09	77.16	42.30	0.45	0.13
G18	MTU1010	97.11	21.19	1.86	95.53	11.40	144.44	22.39	2.47	24.28	78.81	39.84	0.54	0.17
G19	MTU1081	102.11	19.09	1.28	95.24	9.98	230.55	23.56	3.28	15.70	81.06	38.13	0.50	0.14
G20	NLR34449	104.56	14.72	1.38	76.64	13.84	154.89	19.24	1.72	13.42	77.51	31.49	0.47	0.11
G21	Sahbhagi Dhan	97.11	15.96	1.49	97.02	9.89	155.56	23.12	2.70	21.46	83.60	33.82	0.47	0.13
G22	Tellahamsa	93.11	15.46	1.37	100.38	11.60	137.69	23.28	2.75	23.14	84.22	34.99	0.44	0.12
G23	Tulasi	95.11	17.77	1.59	91.20	12.67	137.80	20.47	2.44	23.27	79.01	36.41	0.49	0.14
	Mean	108.64	15.49	1.32	92.19	11.95	150.38	21.51	2.36	19.07	79.95	37.47	0.42	0.11
	Max	122.11	21.19	1.86	139.73	14.38	261.76	24.28	3.65	24.99	93.60	50.53	0.54	0.17
	Min	92.00	6.18	0.43	72.42	8.20	111.06	18.75	1.25	13.42	63.53	19.04	0.31	0.05
	Variance	85.23	11.47	0.08	205.96	2.97	1291.03	2.42	0.31	13.46	50.24	39.37	0.00	0.00
	SD	9.23	3.39	0.29	14.35	1.72	35.93	1.56	0.56	3.67	7.09	6.27	0.07	0.03

of tiller number consistently across the seasons. In terms of vigor, G6 and derived lines were better compared to checks. Productive tillers were highest in G1 at the time of harvest. Among the checks Tulasi and Sahbhagi Dhan showed comparatively higher vigor and BILs outperformed popular checks. G2 showed highest seedling vigor in paper towel screening method.

Stability Analysis

Observations on the yield traits for all three seasons were then subjected to combined analyses through Yield-stability statistic (YSi) (Kang, 1993), AMMI and GGE biplot models. In the analysis, each combination of season with location was considered as an environment. Analysis of variance was first conducted for each environment. Pooled data of 3 seasons was subjected to stability analysis using PB tools and R software and specific genotypic adaptation, general genotypic adaptation and specific population adaptation to different seasons were identified.

It was found that G3 was the most stable genotype in the selection ranks for GY followed by G6, Swarna, G5, G7, and G14 based on combined analysis of yield and stability using

YSi statistic. Similarly for bulk yield, G6 scored highest rank followed by G4, G7, G12, G3, G10, and G2. Among the BILs G6 and G3 showed non-significant stability variance and high average yield, so they may be considered for further multilocation trials. Number of genotypes selected based on YSi ranking varied among the traits. 9 genotypes were found to be superior based on YSi scoring for GY and PH; 8 for DTM, FG,TGW, BM, TDMPD;7 for BY, DFF, TN,PTN, GN,PL,SF, and YPD and 5 genotypes were selected based on high trait mean and stability for HI. G6 was found to be stable for 14 traits under study except DFF, TN, PTN, and BM followed by Swarna which was stable for 13 yield contributing traits. G14(7K) showed stability and high mean for 12 traits, G3(14s) for 11 traits, G2(148S) and G8(24K) for 10 traits and G7(248S) for 9 traits (**Table 4**).

Sum of YSi scores for each yield and contributing trait was computed to identify the overall ranking of genotypes and G6 scored highest followed by Swarna, G3, G14, G8, and G2. Overall ranking varied if we select among the contributing traits. The varieties which scored highest for DFF, DTM belong to late duration as the highest values were considered for calculation. Similarly for plant height, the tallest varieties scored highest YSi

TABLE 4 | YSi Ranking of each genotype based on trait means and significance of stability variance.

SL. NO	YSi	GY	BY	DFF	DTM	PH	TN	PTN	GN	FG	PL	PW	TGW	SF	BM	TDM	HI	TDMPD	YPD	Rank sum	Ysi Sum
G1	14_3	-2	-2	1	-8	-1	15+	12+	-3	-10	2	-10	-2	-9	-2	-2	0	-2	-2	2	-25
G2	148S	3	5+	-10	-10	18+	4	4	0	-3	18+	6+	10+	18+	12+	12+	6	16+	10+	10	119
G3	14S	18+	8+	-9	-9	1	9+	11+	-10	2	4	9+	17+	17+	1	6+	18+	11+	18+	11	122
G4	166_1	-8	13+	-3	2	10+	1	0	7+	7+	15+	5+	12+	0	11+	3	-9	3	-8	8	61
G5	166_2	10+	-1	0	2	6	-1	-1	10+	10+	9+	14+	11+	-2	-4	6+	12+	5	10+	9	96
G6	166S	16+	18+	-2	5+	12+	3	3	9+	9+	14+	9+	16+	10+	5	13+	16+	12+	16+	14	184
G7	248S	9+	11+	14+	-6	3	6	5	-6	-9	3	4+	1	-8	7+	7+	11+	8+	9+	9	69
G8	24K	4	3	12+	9+	16+	0	2	3	6+	-6	8+	13+	11+	10+	10+	5	10+	4	10	120
G9	250K	-3	-8	15+	7+	11+	5	9+	5+	4+	12+	2	8+	9+	-2	-3	6	-3	-4	9	70
G10	3_1K	2	6+	11+	8+	-2	11+	11+	3	-2	1	-7	3	-3	3	2	4	2	1	5	54
G11	65S	4	4	-4	-2	17+	8+	13+	10+	12+	-3	3	-1	4	13+	11+	1	6+	2	8	98
G12	70S	-1	10+	1	8+	-4	5	7+	-4	-1	-2	-9	7+	16+	10+	0	2	0	-2	7	46
G13	75S	-2	2	1	5+	-2	9+	6	-1	1	-1	2	5	14+	-5	-8	15+	-8	-3	4	30
G14	7K	9+	-5	10+	18+	0	10+	6	14+	11+	8+	8+	0	2	9+	8+	-1	6+	7+	12	120
G15	Swarna	15+	4	9+	14+	5	148	13+	8+	10+	6+	7+	4	-4	15+	15+	5	15+	15+	13	170
Trait Mean		14.50	1.22	112.86	142.61	91.97	12.04	11.32	147.71	115.93	21.13	2.27	17.99	79.59	23.25	37.75	0.38	0.26	0.10		
Ysi Mean		4.93	4.53	3.27	2.87	6.00	6.60	6.73	3.00	3.13	5.33	3.40	6.93	5.00	5.53	5.33	6.07	5.40	4.87		
L.S.D. (0.05)		2.45	0.19	1.93	1.71	3.57	2.14	2.09	11.47	11.70	0.83	0.23	0.74	4.09	3.36	5.05	0.04	0.04	0.15		

+ indicates the selection based on stability and trait performance.

ranking. So selection of rank and its direction can be decided based on the requirements of target ecosystems. The ranking by *YSi* statistic based on predicted means for stability parameters for grain yield is shown in **Table 5**.

General Genotypic Adaptation

AMMI and GGE biplot explained the general genotypic adaptation or stability across genotypes (**Figure 2**). To visualize the performance of different genotypes in a given environment, biplots were used. The relative ranking of different genotypes on the biplots is based on its projection onto the O-axis in AMMI Biplot and GGE biplot was used to diagnose the G × E interaction effects on each yield contributing trait. The results of the AMMI model analysis are interpreted on the basis of AMMI1 biplot where the graph is plotted with the main effect and first multiplicative axis term (PC1) for both genotypes and environments. Greater the Principal Component Axis (PC1) scores, either negative or positive, indicated the specific adaptation of a genotype to certain environments. The more the PC1 scores approximate to zero, the more stable the

genotype among the environments under study. The AMMI biplot showed 81.3% fitness in the model for grain yield, and 60.9% for bulk yield. Among the BILs G8, G2, G3, G14, G11, and check Swarna (G15) exhibited high yield with high main (additive) effects showing positive PC1 score. BIL G10 showed less environmental interaction while three environments showed high interaction for GY. Consequently, for BY, *Kharif* 2014 (E2) showed high interaction but genotypes G2, G6, and G12 were identified with low environmental interactions and were considered best across the seasons for the trait. Based on AMMI analysis G10(3-1K) was the most stable genotype for BM; G6 and G3 for BY; G14 for DFF, DTM; G5 for FG, GN, TN; G3 for HI, GY, YPD; G2 for PH, TN, SF, 1000GW; G3 for PTN; G14 for TDM, TDMPD and G12 for TN. GGE biplot also showed similar results for stability of genotypes in trait expression across environments (Supplementary Figure 1). Genotypic variation was observed for each trait in case of adaptability to specific environments. *Kharif* environment was most favorable for high yielding BILs such as G3, G5, G2, and G14 while *Rabi* was favorable for G8 and G6 as they appeared most responsive

TABLE 5 | Stability analysis of variance grain yield of genotypes across 3 environments.

	d.f.	Sum of squares	Mean squares	F	p.value	SS%
TOTAL	44	3902.642				
GENOTYPES	14	1610.939	115.067	3.08	0.005	41.27816
ENVIRONMENTS	2	1245.561	622.7802	63.82	< 0.001	31.91583
INTERACTION	28	1046.142	37.3622	3.83	< 0.001	26.80601
HETEROGENEITY	14	229.0894	16.3635	0.28	0.988	5.870111
RESIDUAL	14	817.0531	58.3609	5.98	< 0.001	20.9359
POOLED ERROR	84		9.758			

SL. NO	SPY	Mean yield	Yield rank	Adj.rank	Adjusted (Y)	Stability variance	Stability rating (S)	YSi (Y+S)	Superior lines
G1	14_3	6.18	1	−3	−2	19.83	0	−2	
G2	148S	15.21	10	1	11	55.68**	−8	3	
G3	14S	20.87	15	3	18	29.39	0	18	✓
G4	166_1	11.00	2	−2	0	88.049**	−8	−8	
G5	166_2	15.11	9	1	10	2.00	0	10	✓
G6	166S	18.90	14	2	16	23.43	0	16	✓
G7	248S	14.61	8	1	9	28.14	0	9	✓
G8	24K	16.58	11	1	12	93.86**	−8	4	
G9	250K	13.02	6	−1	5	56.16**	−8	−3	
G10	3_1K	12.26	3	−1	2	−2.30	0	2	
G11	65S	12.96	5	−1	4	13.63	0	4	
G12	70S	12.63	4	−1	3	44.78*	−4	−1	
G13	75S	13.31	7	−1	6	50.25**	−8	−2	
G14	7K	16.67	12	1	13	44.01*	−4	9	✓
G15	Swarna	18.13	13	2	15	13.50	0	15	✓
Yield Mean		14.50							
Ysi Mean		4.93							
LSD		5.07							

*P < 0.05, **P < 0.01.
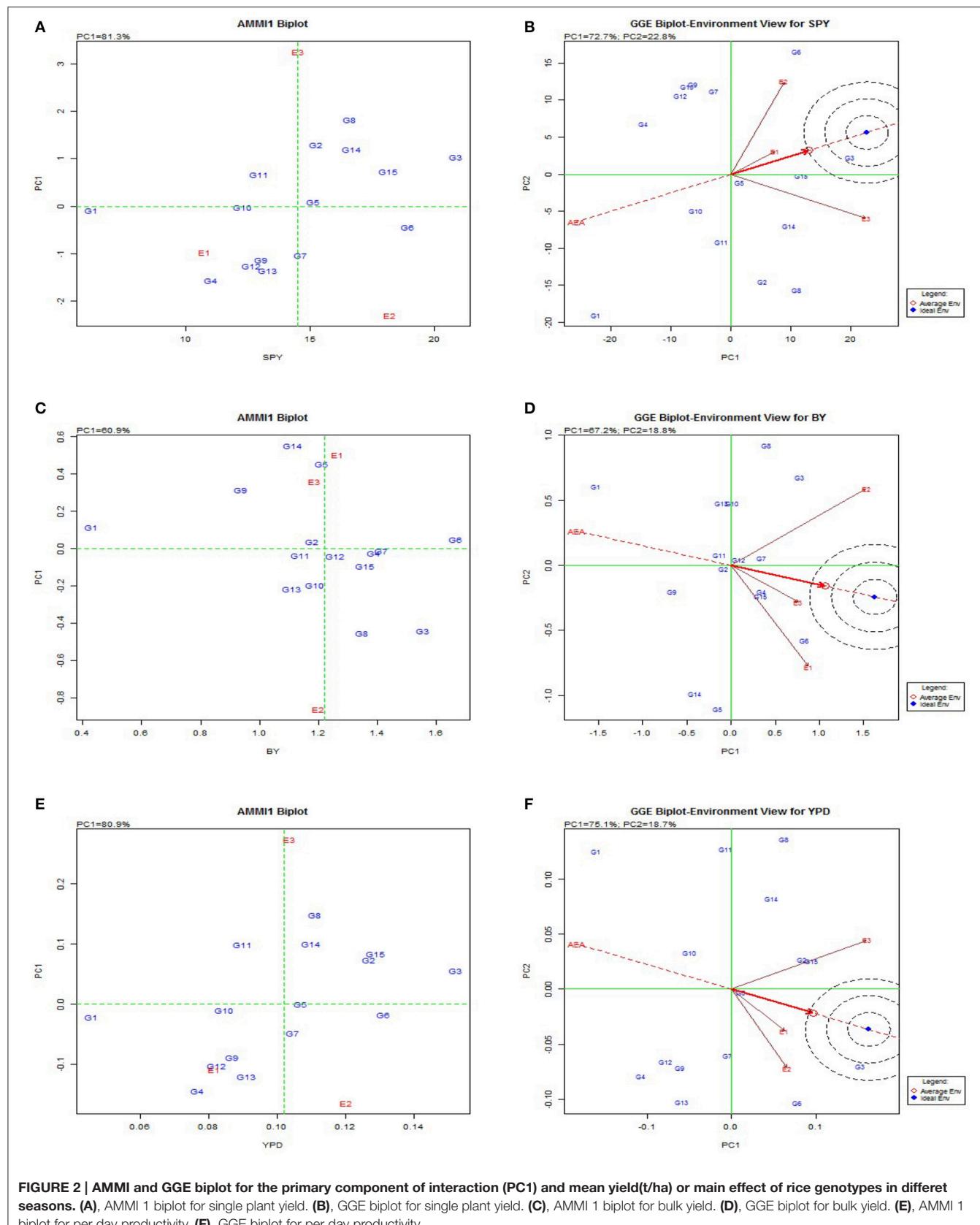
FIGURE 2 | AMMI and GGE biplot for the primary component of interaction (PC1) and mean yield(t/ha) or main effect of rice genotypes in differet seasons. (A), AMMI 1 biplot for single plant yield. **(B)**, GGE biplot for single plant yield. **(C)**, AMMI 1 biplot for bulk yield. **(D)**, GGE biplot for bulk yield. **(E)**, AMMI 1 biplot for per day productivity. **(F)**, GGE biplot for per day productivity.

for yield contributing traits in these respective environments. The environments E1 and E2 were more responsive for the traits BM, FG, GN, PL, PW, and SF and environment E3 was responsive for traits GY, TDM, YPD, TDMPD, DFF, DTM, and HI.

From the biplot graph of AMMI, it was inferred that interactions of environments are highly varied and all the three environments were highly interactive for most of the yield traits. E3 (*Rabi* season) appeared to be a favorable environment for BM and SF; E1 for BY, GN, PL, and E2 for PTN and GY. Genotypes G3(14S), G6(166S), and G9(250K) showed low interaction effects and hence they can be considered stable. In case of GY, G3, and G6 had high mean values and hence they can be recommended for all the environments. The genotypes with high interaction are suitable for specific environments, genotypes with high mean and positive interaction are suited for favorable environments and those with high mean and negative interaction are suited for unfavorable environments for the respective traits. A line that passes through the origin and is perpendicular to the O-axis in the biplots separates genotypes that yielded above the mean (G3, G4, G15, G7, G6, G8, G12, and G4) that would possibly yield above average in all the seasons and genotypes that yielded below average (G1, G14, G5, G10, G14, G9, and G1). The released varieties Tulasi, MTU1010, Swarna and Sahbhagi Dhan used as checks performed well across the seasons.

Specific Genotypic Adaptation

Genotypic evaluation was conducted and based on GGE biplot which-won-where pattern and adaptation showed specific genotypic adaptation to limited environment conditions or the adaptability of genotypes for each environment (**Figure 3**). The same genotype performed best across three seasons for plant height (G2), 1000 grain weight (G2) (Supplementary Figure 2), harvest index (G3) and per day productivity (G3). This shows that these traits have stable expression across the seasons with limited environmental influence and the selected genotypes are most stable for the particular trait. The traits like biomass, days to flowering, days to maturity, filled grains, grain number, panicle weight, spikelet fertility, total dry matter, tiller number showed same genotype performed better in two *Kharif* (wet season) seasons under study but another genotype appeared to be best for *Rabi* (dry season). So these traits are showing seasonal variation and genotypic performance depends on environmental conditions. Traits like bulk yield, productive tiller number and grain yield showed no seasonal dependence on genotypic performance.

The polygon was drawn joining cultivars that are located farthest from the origin so that all other cultivars are contained in the polygon. Perpendicular lines to the sides of the polygon divide the biplot into sectors. Each sector has a vertex cultivar which is present in the corner of the polygon. The vertex cultivar is the best performing cultivar in the environments that share the sector with it. Vertex genotypes are G3, G4 and G7 at E1, E2 and E3 respectively. In case of checks, it is inferred that cultivar Sahbhagi Dhan is suited to *Rabi* season and Tulasi to *Kharif* season. The analysis indicated that G3, G12, and G4 were suitable

BILs for cultivation in irrigated environment as they had the highest ranking in biplot and in predicted means.

Association Analysis

Multiple correlations between different yield and yield related traits was conducted for all the three seasons (**Figure 4**) and it was observed that grain yield has high significant association with panicle weight, 1000 grain weight, total dry matter, per day productivity and harvest index. Days to fifty percent flowering showed negative correlation with bulk yield, grain yield, spikelet fertility, 1000 grain weight, per day productivity and harvest index. Number of primary branches and secondary branches showed positive association both with filled and unfilled grains. Harvest index directly depended on grain yield, per day productivity, filled grains and panicle weight. In season wise correlation conducted among the yield traits, DFF showed highly significant association with DTM; TN with productive tiller number, panicle length and filled grain number; total number of grains with panicle weight in all the three seasons. Single plant yield showed stable and significant association with 1000 grain weight, biomass, harvest index and per day productivity across the seasons.

Genotyping the BILs

All the BILs were screened using universal core genetic map for rice (Orjuela et al., 2010) and the genotypic data was analyzed using the Graphical Genotypes software (GGT 2.0) (van Berloo, 2008). 74 polymorphic SSR loci out of 165 genome wide core set microsatellite (SSR) markers were used for characterisation of BILs. On an average, percentage of recurrent genome of BILs varied from 36.8% (G2) to 90.6% (G14). The most stable and high yielding BILs, G6 (70.8%) and G3 (72.6%) had about 70% of recurrent parent genome and 20% donor parent genome with less than 10% of heterozygous segments. Further, G2 showed less percentage of recurrent parent genome and donor genome with maximum number of null alleles and recombination. Average percentage of recurrent parent genome in the genotypes was 74.7% and donor genome was 12.5%. Heterozygous segments average was 1.7% and null alleles were 9.7% and recombination was 18.5% (**Figure 5**).

Another panel of SSR markers which were linked to the reported QTLs from the same population was also used to screen these BILs. It was observed that many of the BILs have these reported QTLs in either homozygous or heterozygous condition. Yield QTL *yldp1.4* from *O. nivara* was present in most of the BILs.G6 had *O. nivara* allele of *yldp1.4* and its derived lines G4 and G5 had this QTL in completely heterozygous stage. G2 had four QTLs *yldp1.4 yldp2.3*, *nsp1.2* and *dtm 2.7* and G1 had three QTLs *yldp9.1*, *dtm9.3* and *nfg1.2*.

DISCUSSION

Pre-breeding and utilization of wild accessions are gaining importance in plant breeding programs for the identification of novel genes to improve yield levels of existing cultivars. Complex quantitative traits such as yield, with multiple contributing traits are highly influenced by environment interaction effects.
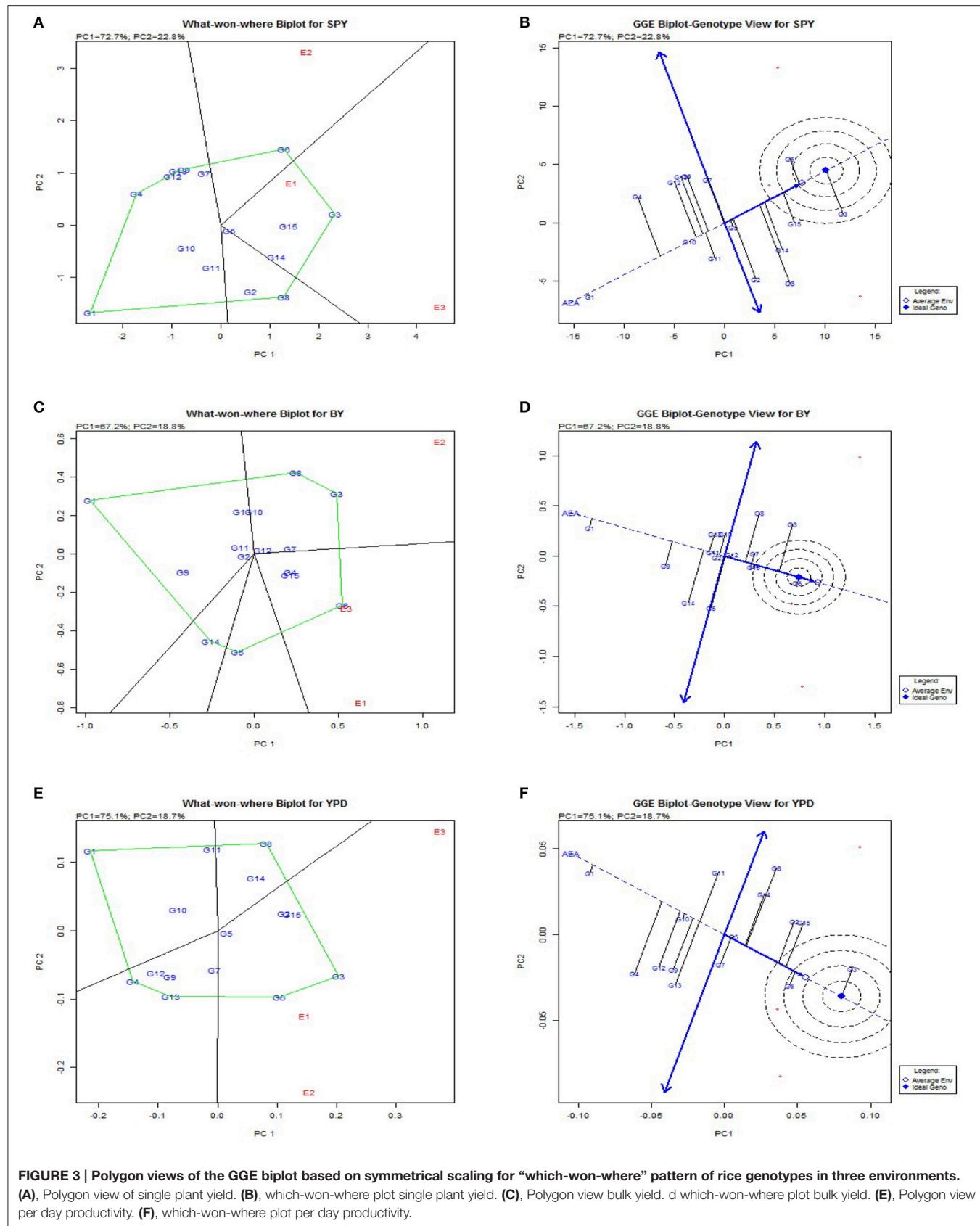
FIGURE 3 | Polygon views of the GGE biplot based on symmetrical scaling for "which-won-where" pattern of rice genotypes in three environments.
(A), Polygon view of single plant yield. **(B)**, which-won-where plot single plant yield. **(C)**, Polygon view bulk yield. d which-won-where plot bulk yield. **(E)**, Polygon view per day productivity. **(F)**, which-won-where plot per day productivity.

FIGURE 4 | Graphic representation of the correlation matrices of yield traits in three seasons.

A-Swarna allele, B- *Oryza nivara* allele, H- Heterozygous, O- other than parental alleles

FIGURE 5 | Graphical representations of donor introgression in the backcross introgression lines using GGT software.
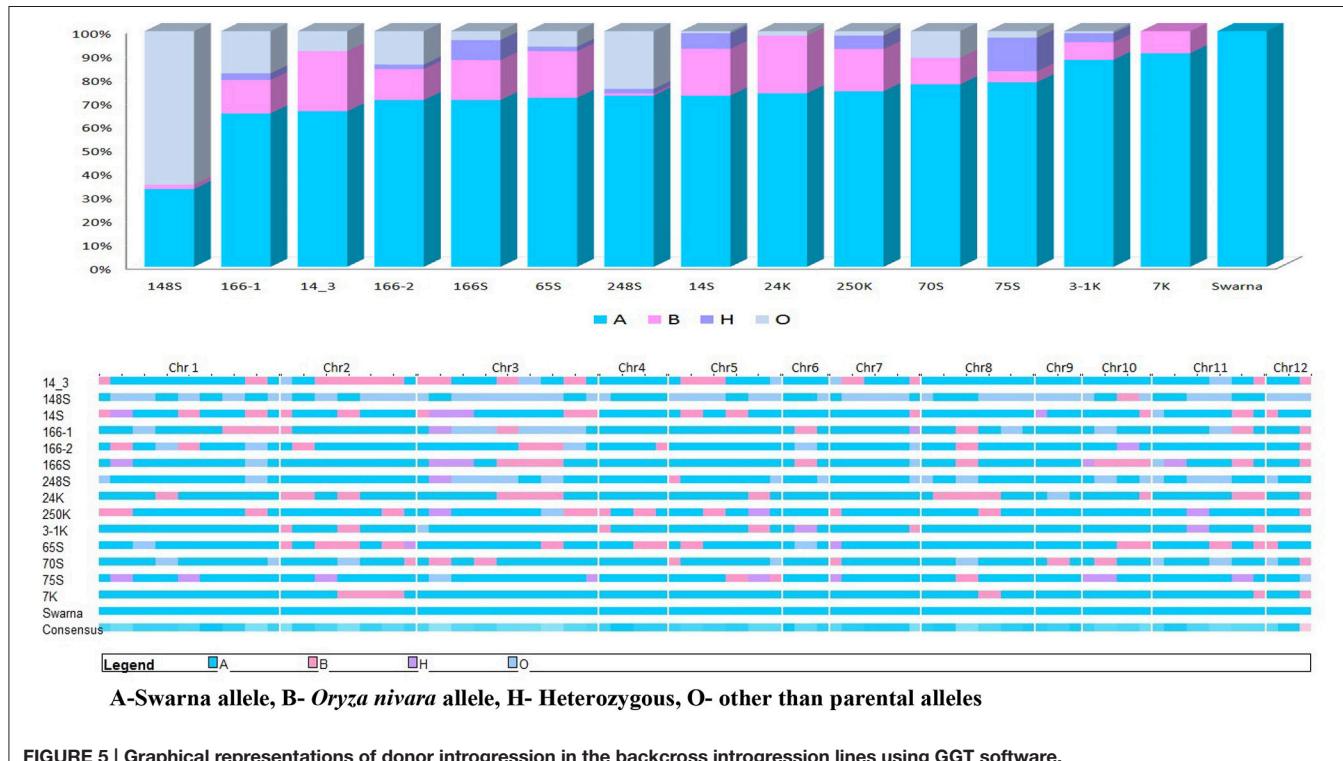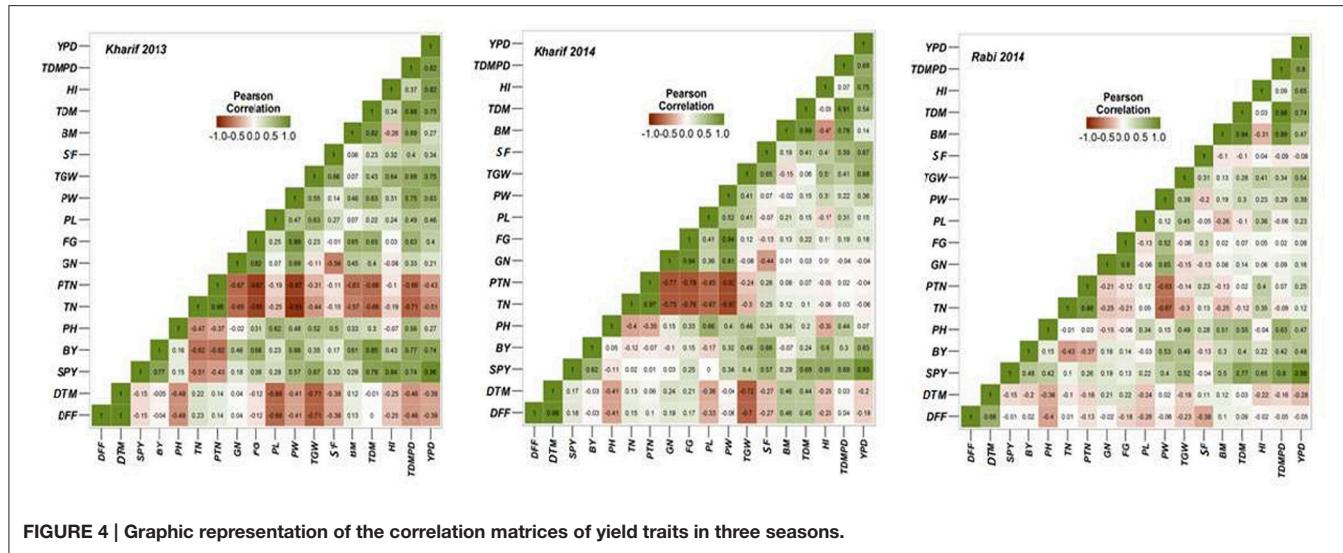
Wide spread cultivation of rice in various agro ecological environments and the unpredicted effects of climate change makes the cultivation of stable and adaptable genotypes more desirable (Bose et al., 2012; Vanave et al., 2014). Stability and GEI studies are very important for the efficient breeding and adoption in multi-environment conditions (Kempton et al., 1997; Atlin et al., 2000; IRRI, 2006; Liang et al., 2015).

The yield of rice genotypes fluctuates considerably with change in environmental conditions (Bose et al., 2014).

Segregation and appearance of wild traits in advanced generations are common phenomena in interspecific crosses. Keeping this in view, the present study was aimed at identification and characterization of a set of BILs developed from the same parental cross, for three seasons and stability was analyzed for key component traits for yield. Most of the multi location trials and multi-year testing focus only on the stability of grain yield but here we studied other yield contributing and component traits also.

Stability analysis models like *YSi* statistics, AMMI and GGE biplots are very useful in selecting lines with high homeostasis for broad target environments and were utilized in multilocation trials and in coordinated variety testing programmes. Prasad et al. (2001) studied stability and yield performance of mega varieties using the data from All India Coordinated Rice Improvement Programme (AICRIP) as well as international trials for a period of 25 years and identified four mega environments for testing the varieties for yield potential in India. Many studies have used GGE biplot analysis mainly for mega-environment evaluation, cultivar evaluation, and assessment of varietal stability (Kang, 1993; Yan and Hunt, 2001; Yan and Kang, 2003; Dehghani et al., 2006; Navabi et al., 2006; Blanche et al., 2007; Ding et al., 2007; Jalata, 2011; Mohammadi et al., 2012; Rakshit et al., 2012; Amiri et al., 2015). Balestre et al. (2010); Nassir (2013) studied stability and adaptability of upland rice genotypes by the GGE biplot method based on the predicted genotypic and phenotypic values. The simultaneous selection for high mean and stability results in the selection of superior genotypes with non significant stability variance and it enhances quality of selection. This method was successfully utilized in most of the crops including rice (Wade et al., 1999; Ouk et al., 2007; Tariku et al., 2013) especially for assessing grain yield.

In this study, three seasons data was subjected to correlation analysis and the traits which are associated significantly were discussed for stability analysis. Stability analysis models helped identification of superior genotypes with both high mean yield and stability. Different stability analysis models showed that G3 is the most stable genotype for grain yield followed by G6 and Swarna. G5 and G6 were identified to be stable and ideal genotypes for bulk yield, grain number and number of filled grains, followed by G3, G7, G15, G8, and G12. G14 was identified as most stable genotype for biomass, flowering duration, panicle length and total dry matter production. G2 showed stability and high mean value for 1000 grain weight, spikelet fertility, plant height and panicle length; it was identified as stable genotype for days to fifty percent flowering and days to maturity with lowest mean value indicating the most stable short duration BIL followed by G3. Some of the superior genotypes for yield specific traits with less stability across the seasons can be stabilized with limited back cross approach (Singh and Huerta-Espino, 2004) with an adapted cultivar.

An ideal genotype would be one that has both high mean yield and high stability. The position of an "ideal" genotype is closer to the direction of the mean environment and has a zero projection onto the perpendicular AEC ordinate. G2 and G4 showed high mean ranking and were identified as the best performing lines in terms of both mean yield and stability across environments in the irrigated ecosystem. Based on adaptation map G2 is best adapted for environment E1, G4 for E2 and G7 for E3. Response plot for mean yield also indicated the same results across the seasons.

Among the genotypes recurrent parent Swarna was stable across the seasons for traits like bulk yield, biomass, days to maturity, number of filled grains, panicle length, productive tiller number, panicle weight. The mega variety Swarna, which is popular in major rice growing countries like India, Bangladesh, Philippines and Thailand; is known for its adaptability in wide range of environments (Prasad et al., 2001). The performance of the BILs was on par or above the mega varieties like Swarna and most popular cultivars of different durations. As the BILs and checks belong to different maturity groups, per day productivity (YPD) was considered to compare their yield. The frequency distribution of the pooled data of three seasons for each trait followed a bell shaped curve with Swarna placed on the peak of the curve (**Figure 6**). However, G6 and G3 proved significantly superior in yield over the recurrent parent Swarna and on par with the best check MTU1010. Graphical representation of the molecular marker data has relevance in studying the genome constitution of the recombinant population (Young and Tanksley, 1989). It was observed that the BILs had more than 70% of recurrent parent genome. Tian et al. (2006) reported that the high-yielding ILs contained relatively less introgressed segments than the low-yielding ILs in a set of 159 ILs derived from *Oryza rufipogon* in *indica* cultivar Guichao2 back ground using 126 polymorphic SSRs.

Plant height and 1000 grain weight were the most stable traits across the season with minimal genotypic variation and with PC1values of 96.5% and 97.8% respectively in GGE biplot (**Figure 7**). The explained SS (%) factor was calculated comparing sum of square (SS) from AMMI ANOVA showing the percentage contribution of genotype, environment and interaction effects in phenotypic expression of each trait. It was observed that grain yield was contributed mainly by genotype (41.28%), followed by environment (31.92%), and their interaction (26.81%). The percentage of explanation of phenotype by genotypic contribution was high for 1000 grain weight (90.40%) and plant height (85.99) and environment effect was high for tiller number (47.56), days to flowering (47.24) and grain yield (31.92) while interaction effect was high for bulk yield and filled grains per panicle (**Figure 8**).

The G × E interactions for most of the yield traits under study were significant but some traits showed stable genotypic performance across the environments. The seasonal variation between the highest value and lowest value was observed for traits BM, BY, DFF, DTM, TN, and GY but the difference was minimal for traits GN, PL, PW, TGW, SF, HI, and YPD (Supplementary Figure 3). For the widely varying traits with high GEI, additional agronomic management is also required along with crop improvement for trait stability. Crossover GE and dissimilarity between environments for discriminating genotypes were very low in case of GN, PH, PL, PW, SF, and TGW but were moderate in case of all other traits. Identification of stably expressing contributing traits is very essential for crop improvement for any major trait than combining multiple traits which fluctuate across the environments. It was observed that 1000 grain weight is the trait contributing most to yield and is stably inherited as well as stably associated with grain yield. So while considering the grouping of the genotypes for multi location yield trials, this trait should also to be considered along with duration for a reliable comparison and analysis. The future breeding programs need to focus on improvement of stably performing traits with high heritability

FIGURE 6 | Frequency distribution of yield and yield related traits in Swarna / *O. nivara* derived BILs arrow indicates Swarna mean value.

FIGURE 7 | The factor explained SS(%) was calculated comparing sum of square (SS) from AMMI ANOVA showing the percentage contribution of genotype environment and interaction effects in phenotypic expression of each trait across environments.

to develop stable high yielding genotypes. Dalvi et al. (2007), Panwar et al. (2008), Waghmode and Mehta (2011), and Padmavathi et al. (2013) studied the GE interaction for grain quality in rice and the stability of grain quality is also important in multilocation trials if quality is the selection criteria.

In this study, we used two *Kharif* and one *Rabi* season data and there was significant seasonal variation observed for most of the traits. In case of yield GEI were higher in *Kharif* season than in *Rabi* season and similar results were reported by Atlin et al. (2000). GGE biplots indicated that *Kharif* 2013 was the ideal season to select genotypes for BM, HI, PL, PTN, SF, GY, TDM, TGW, *Kharif* 2014 for DFF, FG, GN, PH, PW, and *Rabi* 2014 for BY and TGW. *Kharif* seasons were most ideal seasons to select BILs for yield contributing traits; especially *Kharif* 2014 was discriminative as well as representative among the seasons and was suitable for selecting genotypes with general adaptation. *Rabi* season was the most discriminating and least representative environment for testing genotypes and is useful in selecting only specifically adapted genotypes. *Kharif* 2013 was found least discriminating but most representative among the seasons for most of the traits. The highest environmental averages for all the yield traits were observed in either of the two *Kharif* seasons. The maximum value was also observed in the *Kharif* seasons except for traits like SF, BM, and YPD which showed maximum value in *Rabi*. The significant difference due to environment indicated the existence of genotypic differences in adaptability. Genotypes also differed considerably with respect to their stability for yield traits. Similar observations on GEI were made by Gauch and Zobel (1996); Wade et al. (1999); Ouk et al. (2007); Das et al. (2009); Sreedhar et al. (2011); Tariku et al. (2013); Akter et al. (2014) on multi environment studies using rice genotypes. All the three models showed similar results and utilization of *Ysi* statistic is advantageous and complements the AMMI and GGE method for selecting stable and high yielding genotypes (Nassir and Ariyo, 2011). Kumar et al. (2012) used different models for stability analysis and the correlations between the

stability rankings of entries produced by the GGE model and the parameters of Shukla, AMMI, showed very high rank correlation coefficients.

Advanced BILs with stable yield traits can be grown in several environments to study QTL x environment interactions and these lines can be used in breeding programmes as well as to develop varieties in relatively less duration (Jeuken and Lindhout, 2004). Further studies will focus on (i) development of BIL x BIL mapping population from the stable identified genotypes (ii) identification of QTL for stable contributing traits for yield and (iii) development of varieties from selected stable BILs through multi location variety trials.

CONCLUSIONS

The study showed the importance of genotype × environment interaction and stability analysis for evaluation of genotypic yield potential. Wild introgression lines derived from *O. nivara* in Swarna background were studied for stability for yield related traits in *Kharif* and *Rabi* season. The stability and adaptability studies using AMMI, GGE biplot and *Ysi* statistics indicated G3(14S) and G6(166S) as the most stable BILs with high yield performance. The percentage of explanation of genotype on phenotype was high for 1000 grain weight and plant height and environment effect was high for tiller number, days to flowering and grain yield and interaction effect was high for bulk yield and filled grains per panicle. DRR Dhan 40, an elite BIL and recently released variety showed yield stability with high mean performance and the mega varieties which were used as checks also showed yield stability across the seasons. It was observed that wild derived lines with about 70% of recurrent parent genome were more stable and showing enhanced yield levels. Thus, more emphasis should be devoted in future breeding programs to pre breeding and to develop genotypes with wider adaptation. Stability analysis and GEI may
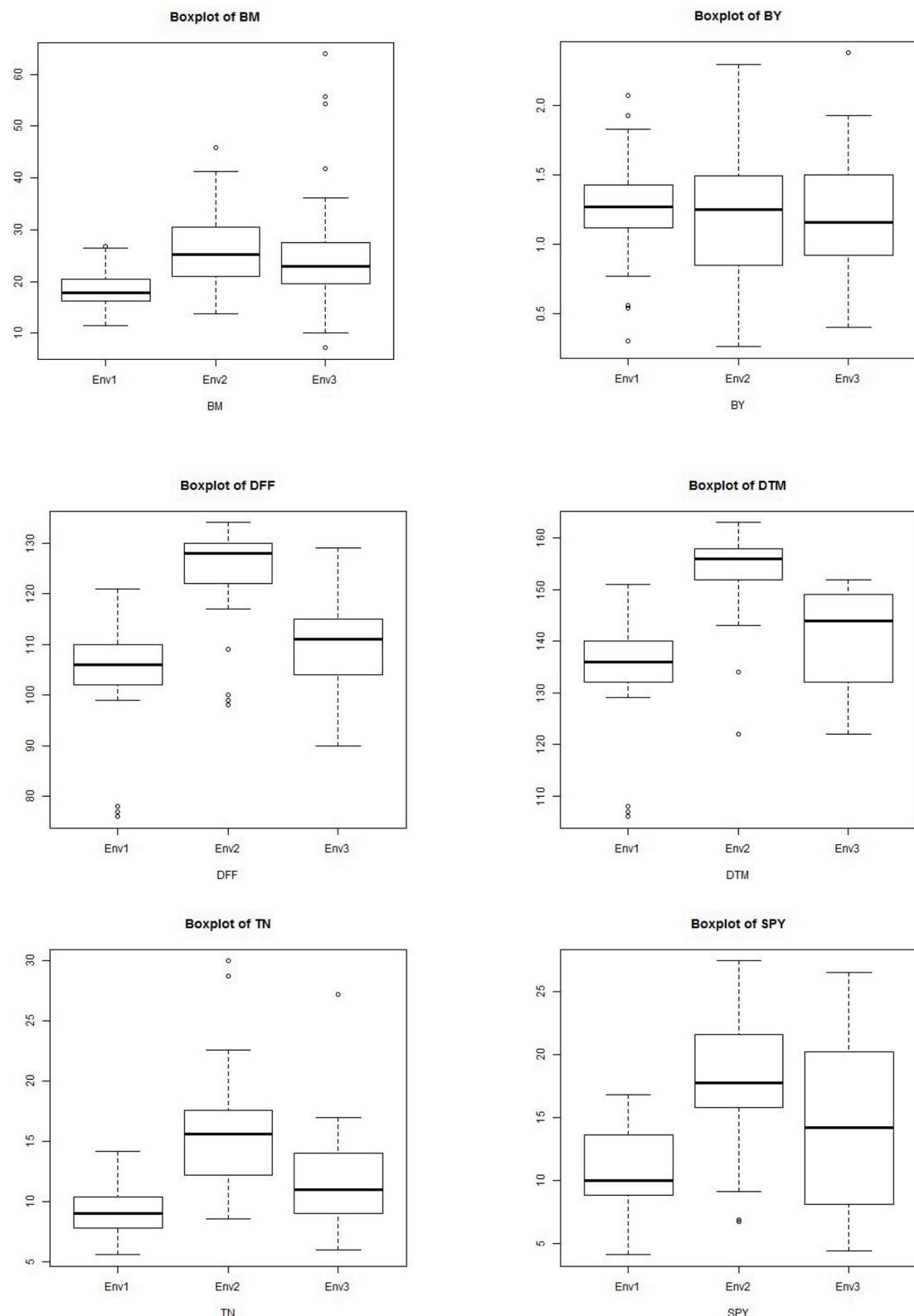
FIGURE 8 | Boxplot showing the differences in yield and related traits among subpopulations. Box edges represent the upper and lower quantile with median value shown as bold line in the middle of the box. Whiskers represent 1.5 times the quantile of the data. Individuals falling outside the range of the whiskers shown as open dots.

be further extended widely for stress resistance, quality as well as nutrient composition for precise identification of superior genotypes.

ETHICS STATEMENT

The authors declare that the experiments comply with the current laws of the country in which they were performed and in compliance with ethical standards.

AUTHOR CONTRIBUTIONS

DB, SN conceived and planned the work. Phenotypic and genotypic screening was performed by JB, AR, RY, KB, SM, MS, DB, RP, and GP. DS, DB, analyzed the data. DB wrote the manuscript. DS and SN revised and proofread the manuscript. Facilities provided at Indian Institute of Rice Research by VB.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016. 01530

REFERENCES

Akter, A., Jamil, H. M., Umma, K. M., Islam, M. R., Hossain, K., and Mamunur, R. M. (2014). AMMI Biplot analysis for stability of grain yield in hybrid rice (*Oryza sativa* L.). *J. Rice Res.* 2, 126–169. doi: 10.4172/jrr.1000126

Amiri, R., Bahraminejad, S., Sasani, S., Jalali-Honarmand, S., and Fakhri, R. (2015). Bread wheat genetic variation for grain's protein, iron and zinc concentrations as uptake by their genetic ability. *Eur. J. Agrono.* 67, 20–26. doi: 10.1016/j.eja.2015.03.004

Anandan, A., Eswaran, R., Sabesan, T., and Prakash, M. (2009). Additive main effects and multiplicative interactions analysis of yield performances in rice genotypes under coastal saline environments. *Adv. Biol. Res.* 3, 43–47.

Atlin, G. N., Baker, R. J., McRae, K. B., and Lu, X. (2000). Selection response in subdivided target regions. *Crop Sci.* 40, 7–13. doi: 10.2135/cropsci2000.4017

Balestre, M., Santos, V. B., Soares, A. A., and Reis, M. S. (2010). Stability and adaptability of upland rice genotypes. *Crop Breed. Appl. Biotechnol.* 10, 357–363. doi: 10.1590/S1984-70332010000400011

Blanche, S. B., Myers, G. O., and Kang, M. S. (2007). GGE biplots and traditional stability measures for interpreting genotype by environment interactions. *J. Crop Imp.* 20, 123–135. doi: 10.1300/J411v20n01_07

Bose, L. K., Jambhulkar, N. N., and Singh, O. N. (2014). Additive main effects and multiplicative interaction (AMMI) analysis of grain yield stability in early duration rice. *J. Anim. Plant Sci.* 24, 1885–1897.

Bose, L. K., Nagaraju, M., and Singh, O. N. (2012). Genotype × Environment interaction and stability analysis of lowland rice genotypes. *J. Agric. Sci.* 57, 1–8. doi: 10.2298/JAS1201001B

Brar, D. S., and Khush, G. S. (1997). Alien introgression in rice. *Plant Mol. Biol.* 35, 35–47. doi: 10.1023/A:1005825519998

Cooper, M., Rajatasereekul, S., Immark, S., Fukai, S., and Basnayake, J. (1999). Rainfed lowland rice breeding strategies for northeast Thailand. I. Genotypic variation and genotype × environment interactions for grain yield. *Field Crops Res.* 64, 131–151. doi: 10.1016/S0378-4290(99)00056-8

Cooper, M., and Somrith, B. (1997). "Implications of genotype-by-environment interactions for yield adaptation of rainfed lowland rice: influence of flowering date on yield variation," in *Breeding Strategies for Rainfed Lowland Rice in Drought-Prone Environments,* eds S. Fukai, M. Cooper, and J. Salisbury (Canberra, ACT: Australian Centre for International Agricultural Research), 104–114.

Cornelius, P. L., Crossa, L., and Seyedsadr, M. S. (1996). "Statistical test and estimators of multiplicative models for genotype-by-environment interaction," in *Genotype-by-Environment Interaction,* eds M. S. Kang and H. G. Gauch (Boca Raton, FL: CRC Press),199–234.

Crossa, J., and Cornelius, P. L. (1997). Sites regression and shifted multiplicative model clustering of cultivar trials sites under heterogeneity of variances.

Crop Sci. 37, 406–415. doi: 10.2135/cropsci1997.0011183X0037000 20017x

Crossa, J., Cornelius, P. L., and Yan, W. (2002). Biplots of linear-bilinear models for studying crossover genotype × environment interaction. *Crop Sci.* 42, 619–633. doi: 10.2135/cropsci2002.0619

Crossa, J., Gauch, H. G., and Zobel, R. W. (1990). Additive main effects and multiplicative interactions analysis of two international maize cultivar trials. *Crop Sci.* 30, 493–500. doi: 10.2135/cropsci1990.0011183X003000030003x

Dalvi, V. V., Patel, D. V., Vashi, P. S., and Kshirsagar, R. M. (2007). Genotype × Environment interaction for yield and its components in rice hybrids. *J. Maharastra Agric.Univ.* 32, 25–28.

Das, S., Misra, R. C., and Patnaik, M. C. (2009). G×E interaction of mid-late rice genotypes in LR and AMMI model and evaluation of adaptability and yield stability. *Environ Ecol.* 27, 529–535.

Dehghani, H., Ebadi, A., and Yousefi, A. (2006). Biplot analysis of genotype by environment interaction for barley yield in Iran. *Agron. J.* 98, 388–393. doi: 10.2134/agronj2004.0310

Ding, M., Tier, B., and Yan, W. (2007). "Application of GGE biplot analysis to evaluate genotype (G), environment (E) and G×E interaction on *P. radiata*: case study," in *Australasian Forest Genetics Conference, 11–14 April 2007,* the Old Woolstore, Hobart, Tasmania, Australia.

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochem. Bull.* 19, 11–15.

Eberhart, S. A., and Russell, W. A. (1966). Stability parameters for comparing varieties. *Crop Sci.* 6, 36–40. doi: 10.2135/cropsci1966.0011183X0006000 10011x

Falconer, D. S., and Mackey, T. F. C. (1996). *Introduction to Quantitative Genetics, 4th Edn.* Harlow: Addison-Wesley Longman.

Finlay, K. W., and Wilkinson, G. N. (1963). The analysis of adaptation in a plant breeding programme. *Aust. J. Agric. Res.*14, 742–754. doi: 10.1071/AR9630742

Freeman, G. H., and Perkins, J. M. (1971). Environmental and genotype environmental components of variability. VIII. Relations between genotypes grown in different environments and measures of these environments. *Heredity* 27, 15–23. doi: 10.1038/hdy.1971.67

Gauch, G. H., and Zobel, R. W. (1996). "AMMI analysis of yield trials," in, *Genotype by Environment Interaction,* eds M. S. Kang, and H. G. Gauch (Boca Raton, FL: CRC Press), 85–122.

Gauch, H. G. (1988). Model selection and validation for yield trials with interaction. *Biometrics.* 44, 705–715. doi: 10.2307/2531585

Gauch, H. G., and Zobel, R. W. (1997). Identifying mega-environments and targeting genotypes. *Crop Sci.* 37, 311–326. doi: 10.2135/cropsci1997.001118 3X003700020002x

Henderson, S. A., Fukai, S., Jongdee, B., and Cooper, M. (1996). "Comparing simulation and experimental approaches to analysis of genotype by

environment interactions for yield in rainfed lowland rice," in *Plant Adaptation and Crop Improvement*, eds M. Cooper, and G. L. Hammer (Wallingford: CAB International, in Association with IRRI and ICRISAT), 443–464.

Inthapanya, P. S., Sihavong, P., Sihathep, V., Chanphengsay, M., and Fukai, S., Basnayake, J. (2000). Genotypic performance under fertilized and non-fertilized conditionsin rainfed lowland rice. *Field Crops Res.* 65, 1–14. doi: 10.1016/S0378-4290(99)00065-9

International Seed Testing Association (ISTA) (1999). International rules for seed testing. *Seed Sci. Technol.* 27, 340.

IRRI (2006). *Bringing Hope, Improving Lives: Strategic Plan 2007–2015.* LosBanos, CA: IRRI.

IRRI (2013). *Standard Evaluation System (SES) for Rice, 5th Edn.* Philippines: RRI.

Islam, M. R., Sarker, M. R. A., Sharma, N., Rahman, M. A., Collard, B. C. Y., Gregorio, G. B., et al. (2015). Assessment of adaptability of recently released salt tolerant rice varieties in coastal regions of South Bangladesh. *Field Crops Res.* 190, 34–43. doi: 10.1016/j.fcr.2015.09.012

Jalata, Z. (2011). GGE-biplot analysis of multi-environment yield trials of barley (Hordeium vulgare L.) genotypes in south eastern Ethiopia highlands. *Int. J. Plant Breed. Genet.* 5, 59–75. doi: 10.3923/ijpbg.2011.59.75

Jeuken, M. J. W., and Lindhout, P. (2004). The development of lettuce backcross inbred lines (BILs) for exploitation of the Lactuca saligna (wild lettuce) germplasm. *Theor Appl. Genet.* 109, 394–401. doi: 10.1007/s00122-004-1643-7

Kaladhar, K., Swamy, B. P. M., Babu, A. P., Reddy, C. S., and Sarla, N., (2008). Mapping quantitative trait loci for yield traits in BC2F2 population derived from Swarna / O.nivara cross. *Rice Genet. Newsl.* 24, 34–36.

Kang, M. S. (1993). Simultaneous selection for yield and stability in crop performance trials: consequences for growers. *Agron. J.* 85, 754–757. doi: 10.2134/agronj1993.00021962008500030042x

Kang, M. S. (1998). Using genotype-by-environment interaction for crop cultivar development. *Adv. Agron.* 62, 199–252. doi: 10.1016/S0065-2113(08)60569-6

Kang, M. S., and Magari, R. (1995). STABLE: A basic program for calculating stability and yield stability statistics. *Agron. J.* 87, 276–277. doi: 10.2134/agronj1995.00021962008700020023x

Katsura, K., Tsujimoto, Y., Oda, M., Matsushima, K., Inusah, B., Dogbe, W., et al. (2016). Genotype-by-environment interaction analysis of rice (*Oryza* spp.) yield in a flood plain ecosystem in West Africa. *Eur. J. Agron.* 73, 152–159. doi: 10.1016/j.eja.2015.11.014

Kempton, R. A. (1984). The use of bi-plots in interpreting variety by environment interactions. *J. Agric. Sci.* 103, 123–135. doi: 10.1017/S0021859600043392

Kempton, R. A., Fox, P. N., and Cerezo, M. (1997). *Statistical Methods for Plant Variety Evaluation.* (London: Chapman and Hall), 1–8.

Kumar, A., Verulkar, S. B., Mandal, N. P., Variar, M., Shukla, V. D., Dwivedi, J. L., et al. (2012). High-yielding, drought-tolerant, stable rice genotypes for the shallow rainfed lowland drought-prone ecosystem. *Field Crops Res.* 133, 37–47. doi: 10.1016/j.fcr.2012.03.007

Liang, S., Ren, G., Liu, J., Zhao, X., Zhou, M., McNeil, D., et al. (2015). Genotype-by-environment interaction is important for grain yield in irrigated lowland rice. *Field Crops Res.* 180, 90–99. doi: 10.1016/j.fcr.2015.05.014

Mohammadi, M., Karimizadeh, R., Hosseinpour, T., Falahi, H. A., Khanzadeh, H., Sabaghnia, N., et al. (2012). Genotype × Environment interaction and stability analysis of seed yield of durum wheat genotypes in dryland conditions. *Not. Sci. Biol.* 4, 57–64. doi: 10.15835/nsb.4.3.7807

Nassir, A. L. (2013). Genotype × Environment analysis of some yield components of upland rice (*Oryza sativa* L.) under two ecologies in Nigeria. *Int. J. Plant Breed. Genet.* 7, 105–114. doi: 10.3923/ijpbg.2013.105.114

Nassir, A. L., and Ariyo, O. J. (2011). Genotype × Environment interaction and yield-stability analyses of rice grown in tropical inland swamp. *Not. Bot. Hort. Agrobot. Cluj.* 39, 220–225. doi: 10.15835/nbha3915591

Navabi, A., Yang, R.-C., Helm, J., and Spaner, D. M. (2006). Can spring wheat-growing mega or niche-adapted genotypes? *Crop Sci.* 46, 1107–1116. doi: 10.2135/cropsci2005.06-0159

Orjuela, J., Garavit, A., Bouniol, M., Arbelaez, J. D., Moreno, L., Kimball, J., et al. (2010). A universal core genetic map for rice. *Theor. Appl. Genet.* 120, 563–572. doi: 10.1007/s00122-009-1176-1

Ouk, M., Basnayake, J., Tsubo, M., Fukai, S., Fischer, K. S., Kang, S., et al. (2007). Genotype-by-environment interactions for grain yield associated with water availability at flowering in rainfed lowland rice. *Field Crops Res.* 101, 145–154. doi: 10.1016/j.fcr.2006.10.003

Padmavathi, P. V., Satyanarayana, P. V., Ahamed, L. M., and Chamundeswari, N. (2013). Stability analysis of quality traits in rice hybrids. *Oryza* 50, 208–213.

Panwar, L. L., Joshi, V. N., and Ali, M. (2008). Genotype × Environment interaction in scented rice. *Oryza* 45, 103–109.

Perkins, J. M., and Jinks, J. L. (1968). Environmental and genotype environmental components of variability. III. Multiple lines and crosses. *Heredity* 23, 339–356. doi: 10.1038/hdy.1968.48

Prasad, G. S. V., Prasadarao, U., Rani, N. S., Rao, L. V. S., Pasalu, I. C., and Muralidharan, K. (2001). Indian rice varieties released around the world. *Curr. Sci.* 80, 1508–1511.

Rakshit, S., Ganapathy, K. N., Gomashe, S. S., Rathore, A., Ghorade, R. B., Nagesh Kumar, M. V., et al. (2012). GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data. *Euphytica* 185, 465–479. doi: 10.1007/s10681-012-0648-6

R Core Team (2012). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. ISBN 3-900051-07-0

Shrestha, S., Asch, F., Dusserre, J., Ramanantsoanirina, A., and Brueck, H. (2012). Climate effects on yield components as affected by genotypic responses to variable environmental conditions in upland rice systems at different altitudes. *Field Crops Res.* 134, 216–228. doi: 10.1016/j.fcr.2012.06.011

Shukla, G. K. (1972). Some statistical aspects of partitioning genotype environmental components of variability. *Heredity* 29, 237–245. doi: 10.1038/hdy.1972.87

Singh, P., and Narayanan, S. S. (2006). *Biometrical Techniques in Plant Breeding.* New Delhi: Kalyani Publishers, 316.

Singh, R. P., and Huerta-Espino, J. (2004). "The use of 'Single-backcross, selected bulk breeding approach for transferring minor genes based rust resistance into adapted cultivars," in *Proceedings of 54th Australian Cereal Chemistry Conference and 11th Wheat Breeders Assembly, 21–24 September 2004,* eds C. K. Black, J. F. Panozzo, G. J. Rebetzke (Canberra, ACT), 48–51

Sreedhar, S., Reddy, T. D., and Ramesha, M. S. (2011). Genotype × Environment interaction and stability for yield and its components in hybrid rice cultivars (*Oryza sativa* L.). *Int. J. Plant Breed. Genet.* 5, 194–208. doi: 10.3923/ijpbg.2011.194.208

Swamy, B. P. M., Kaladhar, K., Ramesha, M. S., Viraktamath, B. C., and Sarla, N. (2011). Molecular mapping of QTLs for yield and yield-related traits in *Oryza sativa* cv Swarna × *O. nivara* (IRGC81848) backcross population. *Rice Sci.* 18, 178–186. doi: 10.1016/S1672-6308(11)60025-5

Swamy, B. P. M., and Sarla, N. (2008). Yield enhancing quantitative trait loci (QTLs) from wild species. *Biotechnol. Adv.* 26, 106–120. doi: 10.1016/j.biotechadv.2007.09.005

Tanksley, S. D., and McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277, 1063–1066. doi: 10.1126/science.277.5329.1063

Tariku, S., Taddesse, L., Bitew, M., and Asfaw, M. (2013). Genotype by environment interaction and grain yield stability analysis of rice (*Oryza sativa* L.) genotypes evaluated in north western Ethiopia. *Net. J. Agric. Sci.* 1, 10–16.

Tian, F., Li, D. J., Fu, Q., Zhu, Z. F., Fu, Y. C., Wang, X. K., et al. (2006). Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. *Theor. Appl. Genet.* 112, 570–580. doi: 10.1007/s00122-005-0165-2

Vanave, P. B., Apte, U. B., Kadam, S. R., and Thaware, B. L. (2014). Stability analysis for straw and grain yield in rice (*Oryza sativa* L.) *Electron. J. Plant Breed.* 5, 442–444.

van Berloo, R. (2008). GGT 2.0 Versatile software for visualization and analysis of genetic data. *J. Hered.* 99, 232–236. doi: 10.1093/jhered/esm109

Wade, L. J., McLaren, C. G., Criseno, L., Amarante, S. T., Sarawgi, A. K., Kumar, R., et al. (1997). "Genotype-by-environment interactions: RLRRC Experience," in *Breeding Strategies for Rainfed Lowland Rice in Drought-prone Environments,* eds S. Fukai, M. Cooper, and J. Salisbury (Canberra, ACT: Australian Centre for International Agricultural Research), 115–125

Wade, L. J., McLaren, C. G., Quintana, L., Harnpichitvitaya, D., Rajatasereekul, S., Sarawgi, A. K., et al. (1999). Genotype by environment interactions across diverse rainfed lowland rice environments. *Field Crops Res.* 64, 35–50. doi: 10.1016/S0378-4290(99)00049-0

Waghmode, B. D., and Mehta, H. D. (2011). Genotype × environment interaction and stability analysis in hybrid rice. *Crop Improv.* 38, 6–12.

Wang, Y. M., Dong, Z. Y., Zhang, Z. J., Lin, X. Y., Shen, Y., Zhou, D. W., et al. (2005). Extensive de novo genomic variation in rice induced by introgression from wild rice (*Zizania latifolia* Griseb.). *Genetics* 170, 1945–1956. doi: 10.1534/genetics.105.040964

Wickneswari, R., Bhuiyan, M. A. R., Kalluvettankuzhy, K. S., Lim, L. S., Thomson, M. J., Narimah, M. K., et al. (2012). Identification and validation of quantitative trait loci for agronomic traits in advanced backcross breeding lines derived from *Oryza rufipogon* × *Oryza sativa* cultivar MR219. *Plant Mol. Biol. Rep.* 30, 929–939. doi: 10.1007/s11105-011-0404-4

Yan, W., and Hunt, L. A. (2001). Interpretation of genotype × environment interaction for winter wheat yield in Ontario. *Crop Sci.* 41, 19–25. doi: 10.2135/cropsci2001.41119x

Yan, W., and Kang, M. S., (2003). *GGE Biplot Analysis: a Graphical Tool for Breeders, Geneticists, and Agronomists.* Boca Raton, FL: CRC Press.

Young, N. D., and Tanksley, S. D. (1989). Restriction fragment length polymorphism maps and the concept of graphical genotypes. *Theor. Appl. Genet.* 77, 95–101. doi: 10.1007/BF00292322

Check for updates

Temporally and Genetically Discrete Periods of Wheat Sensitivity to High Temperature

Henry M. Barber[1]*, Martin Lukac[1,2], James Simmonds[3], Mikhail A. Semenov[4] and Mike J. Gooding[5]

[1] School of Agriculture, Policy and Development, University of Reading, Reading, UK, [2] Faculty of Forestry and Wood Sciences, Czech University of Life Sciences, Prague, Czechia, [3] Department of Crop Genetics, John Innes Centre, Norwich, UK, [4] Computational and Systems Biology Department, Rothamsted Research, Harpenden, UK, [5] Institute of Biological, Environmental and Rural Sciences, University of Aberystwyth, Aberystwyth, UK

Successive single day transfers of pot-grown wheat to high temperature (35/30°C day/night) replicated controlled environments, from the second node detectable to the milky-ripe growth stages, provides the strongest available evidence that the fertility of wheat can be highly vulnerable to heat stress during two discrete peak periods of susceptibility: early booting [decimal growth stage (GS) 41–45] and early anthesis (GS 61–65). A double Gaussian fitted simultaneously to grain number and weight data from two contrasting elite lines (Renesansa, listed in Serbia, *Ppd-D1a*, *Rht8*; Savannah, listed in UK, *Ppd-D1b*, *Rht-D1b*) identified peak periods of main stem susceptibility centered on 3 (s.e. = 0.82) and 18 (s.e. = 0.55) days (mean daily temperature = 14.3°C) pre-GS 65 for both cultivars. Severity of effect depended on genotype, growth stage and their interaction: grain set relative to that achieved at 20/15°C dropped below 80% for Savannah at booting and Renesansa at anthesis. Savannah was relatively tolerant to heat stress at anthesis. A further experiment including 62 lines of the mapping, doubled-haploid progeny of Renesansa × Savannah found tolerance at anthesis to be associated with *Ppd-D1b*, *Rht-D1b*, and a QTL from Renesansa on chromosome 2A. None of the relevant markers were associated with tolerance during booting. *Rht8* was never associated with heat stress tolerance, a lack of effect confirmed in a further experiment where *Rht8* was included in a comparison of near isogenic lines in a cv. Paragon background. Some compensatory increases in mean grain weight were observed, but only when stress was applied during booting and only where *Ppd-D1a* was absent.

Keywords: heat stress, meiosis, anthesis, Ppd-D1, Rht, wheat

INTRODUCTION

Improving crop resilience to more frequent extreme weather events is required to maintain or improve crop yields across Europe (Semenov et al., 2014). Wheat, a major contributor to human diet and health (Shewry and Hey, 2015), is particularly susceptible to heat stress around meiosis and anthesis (Barnabas et al., 2008). Yield loss due to heat stress at these growth stages is primarily due to disruption of reproductive processes (Saini and Aspinall, 1982; Saini et al., 1983), as evidenced by a reduction in fertility and grain number (Dolferus et al., 2011). Previous reports on heat stress

in wheat usually concern only one of the susceptible timings i.e., meiosis (Saini and Aspinall, 1982; Saini et al., 1984) or anthesis (Tashiro and Wardlaw, 1990; Ferris et al., 1998; Lukac et al., 2012; Pradhan et al., 2012; Steinmeyer et al., 2013; Liu et al., 2016). Fewer studies have attempted to quantify the response to stress at both of these timings: Alghabari et al. (2014) suggest meiosis is the most vulnerable stage, but Prasad and Djanaguiraman (2014) report that it is anthesis that is particularly susceptible. Previous work has often assumed that these growth stages represent two separate, discrete periods of susceptibility but there is currently little evidence to support this. Single experiments on rice and wheat suggest that there may be a period between meiosis and anthesis that is relatively tolerant to heat stress (Satake and Yoshida, 1978; Craufurd et al., 2013), but it is unclear as to the specific growth stages when this tolerance occurs. Genotypic interactions with heat stress timing also require clarification. Although some recent work has compared the heat stress response at anthesis across multiple genotypes (Liu et al., 2016), little work has quantified how genotype influences susceptibility across both stages, even though consecutive exposure of both stages to stress seems likely to occur in field conditions (Wardlaw et al., 1989).

Here, we investigate firstly whether periods of vulnerability to heat stress during reproductive phases can truly be differentiated temporally, in association with growth stage development. Secondly we investigate whether the effect of genotype on heat stress vulnerability interacts with timing of stress. We pay particular attention to the effects of three alleles reported to influence heat stress tolerance and have adaptive significance in wheat grown in European regions with different frequencies and severities of heat stress, namely *Rht8*, *Ppd-D1a*, and *Rht-D1b* (Worland, 1996; Worland et al., 1998; Rebetzke et al., 2007; Gasperini et al., 2012; Alghabari et al., 2014; Barber et al., 2015; Kowalski et al., 2016; Jones et al., 2017). We also assess associations with the 1BL/1RS translocation (Schlegel and Korzun, 1997) which introduced a number of race-specific disease resistance genes (Snape et al., 2007). The translocation has also been variously associated with increased above ground biomass, spikelet fertility, delayed senescence, and drought tolerance (Villareal et al., 1998; Rajaram, 2001), but there is apparently little information with regards to its influence on heat stress tolerance.

This paper describes the use of 1-day transfers of pot-grown wheat to replicated controlled environments to identify and characterize any periods of heat susceptibility during external growth stages extending from the second node detectable growth stage (GS 32; Zadoks et al., 1974) to the grain milky-ripe stage (GS 77) and hence encompassing meiosis and anthesis (Barber et al., 2015). An initial study compared the Southern European wheat Renesansa (*Ppd-D1a*, *Rht-D1a*, *Rht8*) to the UK-adapted wheat Savannah (*Ppd-D1b*, *Rht-D1b*, *1BL/1RS*). Once susceptible growth stages were identified, further experiments compared the heat stress responses of near isogenic lines (NILs) of a Paragon background varying for presence and absence of *Rht8*, and also the responses of a mapping population of 62 doubled haploid progeny of Renesansa × Savannah, at appropriate timings.

MATERIALS AND METHODS
Plant Material
Savannah has a high yield potential in North West Europe with low bread making quality and was recommended in the UK in 1998. Renesansa, a Serbian winter wheat listed in 1995, has high yield potential and high bread making quality in southern Europe. Sixty-two lines were selected from a recombinant doubled haploid (DH) population of Savannah × Renesansa based on their alleles at Ppd-D1, Rht-D1, 1BL/1RS, and Rht8 (Xgwm261; Simmonds et al., 2006; Snape et al., 2007). NILs varying for the presence and absence of *Rht8*, though both remaining sensitive to photoperiod were developed in a Paragon background (Kowalski et al., 2016). Paragon is a photoperiod sensitive spring wheat that can be also sown in autumn and was first listed in the UK in 1999 with good bread making quality.

Growing Conditions and Post-harvest Analysis
Plants used in these experiments were grown in pots (180 mm diameter) at the Plant Environment Laboratory at the University of Reading, UK (51 27′ N latitude, 00 56′ W longitude). Each pot contained 2.8 kg of growing media comprising 4:2:4:1 of vermiculite: sand: gravel: compost mixed with Osmocote slow release granules (2 kg m^{-3}) containing a ratio of 15:11:13:2 of N:P$_2$O$_5$:K$_2$O:MgO. Seven seeds were sown per pot; thinned to four plants per pot at the two leaf stage. The pots were maintained outside in prevailing conditions (**Table 1**) under a protective net cage in four randomized blocks with guard pots of wheat placed around the perimeter of experimental blocks. Fungicide was applied as and when required. Pots were watered up to twice daily by an automatic drip irrigation system to maintain field capacity. All treatments consisted of transfers to Saxil growth cabinets, which began between 10:20 and 11:20 h (BST) and remained there for 24 h (16 h day, night time between 22:00 and 06:00 h) before being returned outside to their original randomized block position. Average daily temperature during the treatment period was 14.3°C in 2013/14 and 13.5°C in 2014/15. Two temperature regimes were used in all experiments, day/night temperatures of 20/15 for the control treatment and 35/30°C for the heat stress treatment. Pots were irrigated to field capacity before transfer, but

TABLE 1 | Outside temperatures under which plants were grown in the 2013/14 season.

Month (2013/14)	Mean of daily minima (°C)	Mean of daily maxima (°C)	Average mean temperature (°C)
December	1.9	9.7	5.8
January	2.7	9.4	6.1
February	3.4	9.8	6.6
March	2.9	13.4	8.1
April	5.1	15.1	10.1
May	7.8	17.1	12.5
June	10.5	21.5	16.0
July	12.4	25.0	18.7

were not irrigated whilst in the cabinets. Eight growth cabinets were used which allowed the two temperature treatments to be replicated for the four blocks. On the day of transfer main stems in each pot were tagged and assessed for growth stage (GS, Zadoks et al., 1974). Pots were weighed immediately before and after transfer to monitor water loss. Main stems and tillers were harvested separately after physiological maturity (GS 89) and dried (48 h at 80°C). Ears and spikelets per ear were counted, after which grain was threshed from ears, then re-dried, weighed, and counted by a Kirby Lester K18 tablet counter.

Experiment 1

Experiment 1, sown on the 16th December 2013, comprised a complete factorial of: the two DH parent winter wheat cultivars, Savannah and Renesansa; day of transfer to Saxil growth cabinets (31 separate timings between May 2nd and June 13th 2014); and the two temperature regimes within growth cabinets. Confounding effects associated with temperature included water loss. The mean weight of pots on entry was 3.40 kg, whilst mean weights of pots on withdrawal were 3.19 and 2.98 kg (SED = 0.016) for the 20/15 and 35/30°C treatments, respectively. More detailed studies on the water relations within this growing medium and system suggests that this degree of water loss would equate to 78 and 56% field capacity (FC; oven dry = 0% FC; Gooding et al., 2003), respectively, and that a FC of <70% maintained for 14 days during grain filling was required to reduce grain yield. A further confounded environmental variate was mean relative humidity [73% for 20/15°C and 47% for 35/30°C (SED = 4.4)] whilst in the cabinets.

Experiment 2

Also sown on the 16th December 2013, the treatment structure comprised a complete factorial design of: three genotypes [Paragon, *Rht8* NIL, and Tall NIL (Kowalski et al., 2016)]; day of transfer to Saxil growth cabinets (5 separate days between 19th May and 10th June 2014) and the two temperature regimes within growth cabinets.

Experiment 3

Experiment 3 was sown on 3rd December 2014. The treatment structure comprised a complete factorial of 62 DH Lines, three growth stages at transfer to Saxil growth cabinets, and two temperature regimes within growth cabinets. The three timings targeted specific stages of growth: early booting (GS 39–41); mid booting (GS 43–45); and early anthesis (GS 63–65). Due to variable rates of development within a 24 h period, and differential rates of progression, not all lines were transferred within target. Nonetheless, GS at transfer was always recorded.

Statistical Analysis

The primary statistical approach was an appropriate factorial analysis of variance (ANOVA) with a blocking structure of Block/Cabinet/Pot (GenStat 14th edn., VSN International Ltd.). For Experiments 1 and 2, polynomial regressions were fitted across day of transfer to growth cabinet using orthogonal polynomial contrasts in the ANOVA i.e., treatment structure was pol (Day; *n*) * Temperature * Genotype, where *n* was the

maximum level of polynomial to be fitted. Where quartic effects or deviations from them were significant in Experiment 1, fits were compared with the double Gaussian model (Equation 1) on an r^2 adj basis. The maximal double Gaussian model permits the estimation of two "bell-shaped" curves:

$$Relative\ Effect\ (\%) = 100 + b(2\pi s_1^2)^{-0.5}\, e^{-(t-m)^2/2s_1^2}$$
$$+\ c(2\pi s_2^2)^{-0.5}\, e^{-(t-n)^2/2s_2^2} \qquad (1)$$

Where: *Relative Effect* is the result at 35°C (day temperature) expressed as a percentage of that achieved at 20°C; *b* and *c* are the size of the two peaks; *m* and *n* are when, in time *t*, they are centered; and s_1 and s_2 are the Gaussian shape factors (standard deviation) for the two peaks. This double Gaussian approach has previously been used to detect other phenologically-dependent responses in wheat time series data sets (Lu et al., 2014). The FITNONLINEAR routine in GENSTAT 14 was used to compare regressions and allow a parsimonious approach to the inclusion of various parameters in the model fits. Additionally, the routine allowed simultaneous fits to different response variates (weighted for the inverses of their variances). Here, it was used to investigate potential compensation in mean grain weights at the time when grain numbers were reduced by heat stress.

Experiment 3 was analyzed by ANOVA with a treatment structure of Genotype × Target Growth Stage × Temperature. A regression analysis was conducted in an attempt to control the effects of varying growth stages within the target GS cohorts. Main and interacting effects of *Rht-D1b*, *Rht8*, *Ppd-D1a*, and *1BL/1RS* were tested for their significance in the model (*P* < 0.05). In addition, after correcting for the linear effect of GS within target GS cohort, a QTL analysis was conducted from the effects of the high temperature treatment on individual lines within each target GS. A framework genetic map was constructed from 93 lines of the population as previously described by Snape et al. (2007), containing 107 single sequence repeat (SSR) markers and perfect markers for Ppd-D1, Rht-D1, and 1BL/1RS. Linkage map construction was performed using JoinMap® 3.0 (Kyazma BV) with default settings. Linkage groups were determined using a Divergent log-of-odds (LOD) threshold of 3.0 and genetic distances were computed using the Kosambi regression. The genetic map consisted of 25 linkage groups with 45 unlinked markers. QTL Cartographer 2.5 (North Carolina State University) was used for QTL detection using single marker analysis and composite interval mapping (CIM). Estimates of the additive effects and percentage of total variation for identified QTL were calculated using the multiple interval mapping (MIM) function.

RESULTS

Experiment 1

Grain yield per pot indicated a three factor interaction between day of transfer, temperature and cultivar (*P* = 0.002; deviation from quartic *P* = 0.007; **Figures 1A,B**). Most of the interaction was due to changes in grain number per pot (*P* < 0.001 for the three factor interaction; deviation from quartic *P* < 0.001), with some modification through partial compensatory increases in
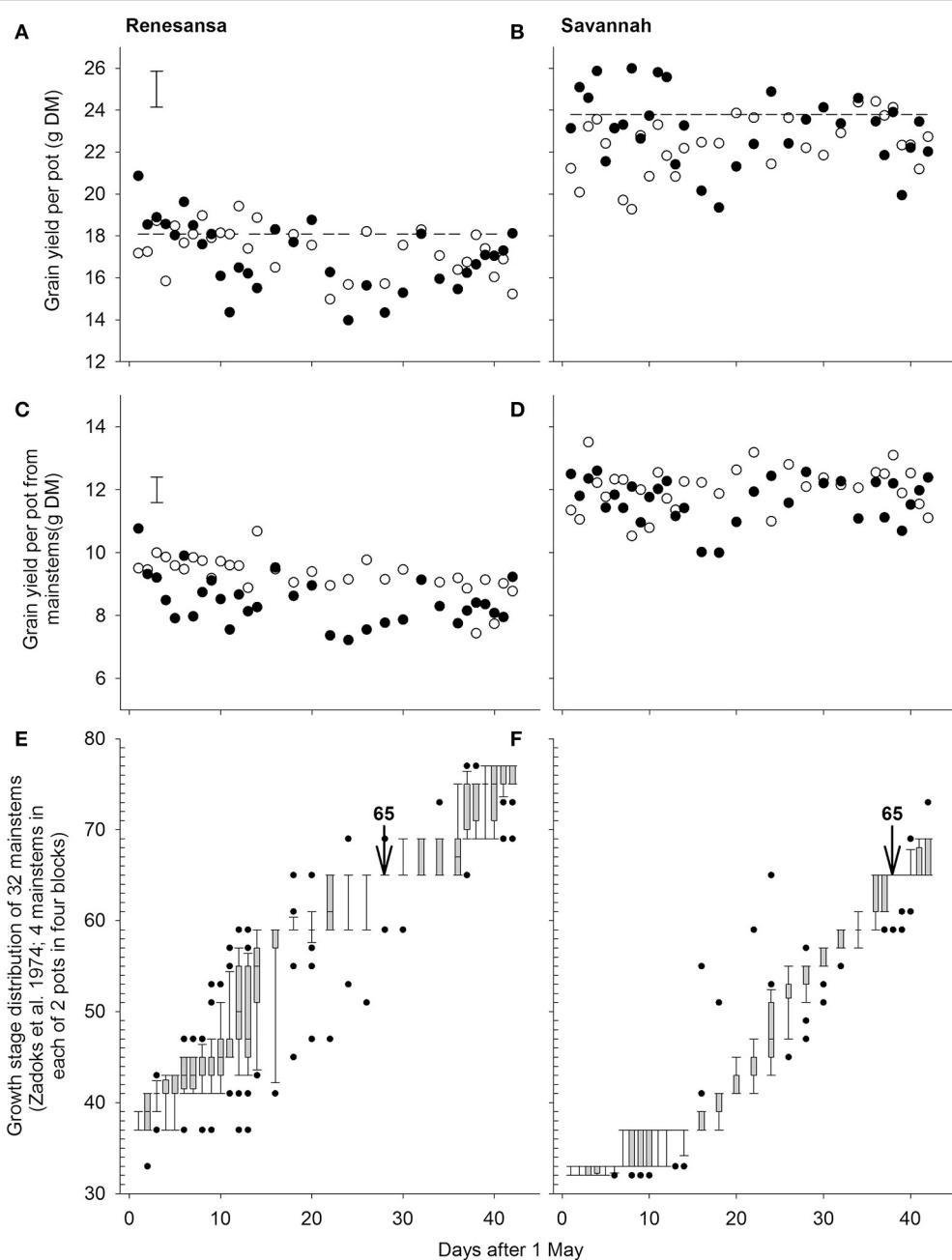
FIGURE 1 | Effects of wheat cultivar and successive 1-day transfers to controlled environment cabinets at 20/15 (○) and 35/30°C (●) day/night temperature (16 h day) on grain yield per pot from all stems (A,B) or only mainstems **(C,D)**. Panels **(E,F)** give the growth stage distributions of the mainstems at the time of transfer in to the cabinets (boxes are limited by 25 and 75%, whiskers by 10 and 90%; points are outliers beyond 10 and 90%, and the line within the box is the median where appropriate). S.E.D. (358 d.f.) in **(A,C)** is for comparing temperatures within day and cultivar for both cultivars. Arrows in **(E,F)** denote the assumed timing of growth stage (GS) 65 (Zadoks et al., 1974). Dashed lines in **(A,B)** are the mean yields from eight pots per cultivar left outside.

mean grain weight, particularly after some of the earlier transfers (e.g., $P < 0.001$ for cubic.Day $\times$ Cultivar). There were no ($P > 0.05$) main, or interacting effects, of temperature on ear number per pot (mean for Renesansa and Savannah = 9.2 and 9.5, respectively; S.E.D. = 0.12; 345 d.f.) or spikelet number per ear (Renesansa = 20.3, Savannah = 20.0; S.E.D. = 0.09).

With regards to timing of susceptibility to heat stress, the grain yields from the main stems provided better clarity than the yields from the whole plot, presumably because of the broader spectrum of the growth stages deriving from the tillers (Jones et al., 2017) and as growth stage assessments focussed primarily on main stems. On the main stems, yields of Renesansa

appeared to be repeatedly compromised by day transfers to the higher temperature from 6 to 12 May, and again from 22 to 30 May (**Figure 1C**). In Savannah there was a significant period of susceptibility from the 17 to 21 May, and possibly a second period from 4 to 9 June (**Figure 1D**). Variation in growth stage amongst mainstems appeared to be greater for Renesansa (**Figure 1E**) than for Savannah (**Figure 1F**). Nonetheless, on average, for much of the period of transfers, the growth stage development of Savannah appeared to be about 10 days later than that for Renesansa. This difference could be identified with accuracy at mid anthesis as over 80% of mainstems were scored as at GS 65 on 28 May for Renesansa and on 7 June for Savannah.

When Day of transfer was expressed as relative to GS 65, there was strong evidence for two peak timings of susceptibility, but there was no evidence that timing of the peaks for susceptibility varied for the two cultivars, or that the standard deviation of the two peaks varied (Gaussian *s*). With regards to grain numbers on the mainstem (**Table 2**; **Figure 2**), a first peak was centered about 18 days before GS 65 when 50% of Renesansa mainstems were at GS 43–45, and 50% of Savannah mainstems were at GS 41–43 (**Figure 1**). Both cultivars appeared comparatively tolerant of the heat stress during late booting and ear emergence. A second period of susceptibility, however, was detected during late ear emergence and early phases of anthesis, centered on 3 days before GS 65 (**Table 2**; **Figure 2**), when most of the ears would have been at GS 61. Grain set in Renesansa appeared equally susceptible to the heat stress during booting and anthesis (**Table 2**; **Figure 2**). Grain set in Savannah was significantly more susceptible during booting than at anthesis, but the only time when grain set was significantly compensated by increased mean grain weight was at the earlier timing (**Table 2**; **Figure 2**). There was no statistical evidence for compensation for grain set failure through mean grain weight by Renesansa during either period of susceptibility.

Experiment 2

There was a significant interaction between the time of transfer and temperature on mainstem grain number ($P = 0.005$ for Temperature × quadratic Day). As in Experiment 1, a significant reduction in grain numbers from the main stems resulted from a day transfer to 35/30°C rather than 20/15°C, 18 days before mid anthesis (GS 65; **Figure 3**), whilst the plants were in the early to mid-stages of booting (c. GS 43). There were smaller reductions in grain numbers following heat stress during late ear-emergence and early anthesis, commensurate with the effects on grain numbers of Savannah at similar timings in Experiment 1. Plants appeared tolerant of the higher temperature at the start of booting (c. GS 40) and by mid anthesis (GS 65). There was no statistical evidence in Experiment 2 that reductions in grain numbers were mitigated by increases in mean grain weight; neither was there any evidence that *Rht8* influenced tolerance to heat stress during booting or anthesis ($P = 0.997$ for Temperature × Day × Genotype on mainstem grain numbers).

Experiment 3

Within the doubled haploid population, when using the "target" growth stages for transfer as a fixed effect there was a very highly significant interaction ($P < 0.001$) between temperature, growth stage, and DH line for grain number. When making

TABLE 2 | Parameter values for simultaneous double Gaussian fit (Figure 2) to the effects of increasing day temperature from 20 to 35°C over successive single days for grain yield components on main stems of two cultivars of winter wheat.

			Estimate	S.e.
Gaussian shape factor (s, days)			3.71	0.416
Peak position (days relative to GS 65)		Peak 1	−18.2	0.55
		Peak 2	−3.0	0.82
Grain number	Renesansa	Peak 1	−359	66.7
		Peak 2	−491	92.1
	Savannah	Peak 1	−555	92.4
		Peak 2	−231	77.6
Mean grain weight (mg)	Renesansa	Peak 1	17.5	8.5
		Peak 2	2.3	11.8
	Savannah	Peak 1	45.3	12.0
		Peak 2	12.2	10.2

some allowance for actual growth stages within target stress timings, there was evidence of increasing susceptibility from GS 37 to 41 (**Figure 4D**) and from GS 59 to 65 (**Figure 4F**). There was wide variation in susceptibility of lines within the doubled-haploid population, particularly at the mid-booting growth stage (**Figures 4B,E**). None of this variation was significantly associated with the markers for *Rht8* or the 1BL/1RS translocation. At anthesis, however, main effect associations with both *Rht-D1b* ($P < 0.001$) and *Ppd-D1a* ($P = 0.006$) were significant. *Rht* (tall) and *Ppd-D1a* were associated with increased susceptibility during anthesis (**Figure 4F**). The QTL analysis confirmed the protective nature of the Savannah alleles (*Rht-D1b* and *Ppd-D1b*), but in addition identified a further, and stronger protective QTL from Renesansa on chromosome 2A (**Table 3**). None of these alleles could be detected as being protective against heat stress applied during booting. There was however, a weak protective QTL from Renesansa for heat applied during early booting on 2B (nearest marker = Xgwm120; LOD = 1.85; additive effect = −3.75).

In addition to effects on fertility, there was a significant three factor interaction on mean grain weight ($P = 0.032$). Increased mean grain weight at the higher temperature during the early stages of booting (**Figure 4A**) occurred in the lines not marked for *Ppd-D1a*, and was most evident in lines containing *Rht-D1b*. As anthesis progressed, the higher temperature caused progressively greater reduction in the mean grain weights of lines containing *Ppd-D1a* (**Figure 4C**).

DISCUSSION

This study clarifies the effect of heat stress on wheat yield during reproductive development, as well as the influence of growth stage and potentially adaptive genotypic effects. We have identified two discrete periods at which grain set in wheat is susceptible to high temperature: the first in early to mid-booting presumably commensurate with susceptible meiotic stages (Barber et al., 2015) and the second during the early phases of anthesis. We have demonstrated that genotypic effects
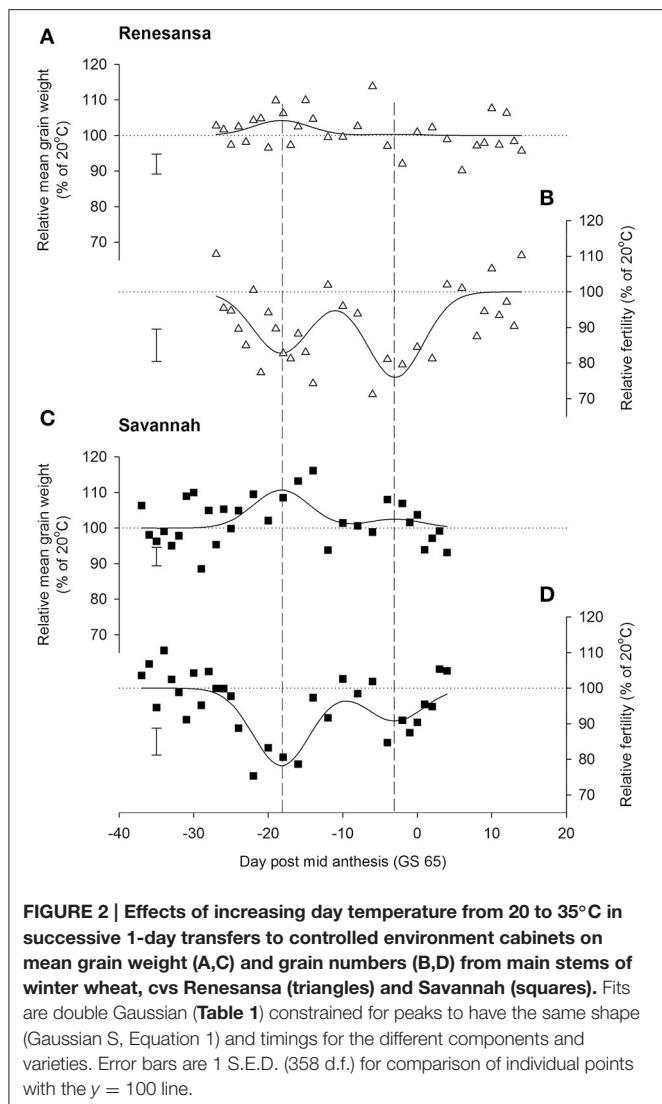
FIGURE 2 | Effects of increasing day temperature from 20 to 35°C in successive 1-day transfers to controlled environment cabinets on mean grain weight (A,C) and grain numbers (B,D) from main stems of winter wheat, cvs Renesansa (triangles) and Savannah (squares). Fits are double Gaussian (**Table 1**) constrained for peaks to have the same shape (Gaussian S, Equation 1) and timings for the different components and varieties. Error bars are 1 S.E.D. (358 d.f.) for comparison of individual points with the $y = 100$ line.

FIGURE 3 | Effects of increasing day temperature from 20 to 35°C in 1-day transfers to controlled environment cabinets on yield components per pot from main stems of near isogenic lines with (●) and without (○) *Rht8* in a Paragon wheat background. Error bars in **(A,B)** are S.E.D.s for comparing points without (left) and with (right) *Rht8* with the 100% line. Box-whisker plots (**Figure 1** for description) in **(C)** show growth stage distributions of mainstems on day of transfer.

on tolerance to heat stress vary with the particular period of vulnerability.

Reductions in grain number due to heat stress caused by reduced fertility found across all experiments in this study are in agreement with previous work (Saini and Aspinall, 1982; Ferris et al., 1998; Dolferus et al., 2011; Liu et al., 2016). There is some evidence to suggest that grain size can increase and partially compensate for losses caused by abiotic stresses (Semenov et al., 2014), however this is mostly confined to the booting period of susceptibility and was not consistently observed across genotypes. Grain size increases found at booting but not at anthesis support the lack of grain size compensation found by Liu et al. (2016). This variation in compensatory increases in mean grain weight over genotype and growth stage should be accounted for when attempting to improve the response of crop models to abiotic stress (Stratonovitch and Semenov, 2015; Liu et al., 2016). Consistent with previous literature, the peak periods of susceptibility appear to be early to mid-booting (Saini and Aspinall, 1982; Alghabari et al., 2014)

and early flowering (Ferris et al., 1998; Craufurd et al., 2013; Prasad and Djanaguiraman, 2014). There is some evidence to suggest that the period between meiosis and anthesis appears to be relatively tolerant to short durations of heat stress: similar to what has been observed in rice (Satake and Yoshida, 1978, 1981; Craufurd et al., 2013), with indications that this could also be true in wheat (Prasad and Djanaguiraman, 2014). Responses to heat stress are strongly influenced by genotype, as shown by variation within these experiments, especially between Savannah and Renesansa. Genotypic differences, especially at anthesis, as observed here, have been identified previously (Stone and Nicolas, 1994; Alghabari et al., 2014; Lobell et al., 2015; Liu et al., 2016). This suggests that there is potential for identifying heat tolerant traits within the current genetic diversity of wheat, which will be crucial for crop production
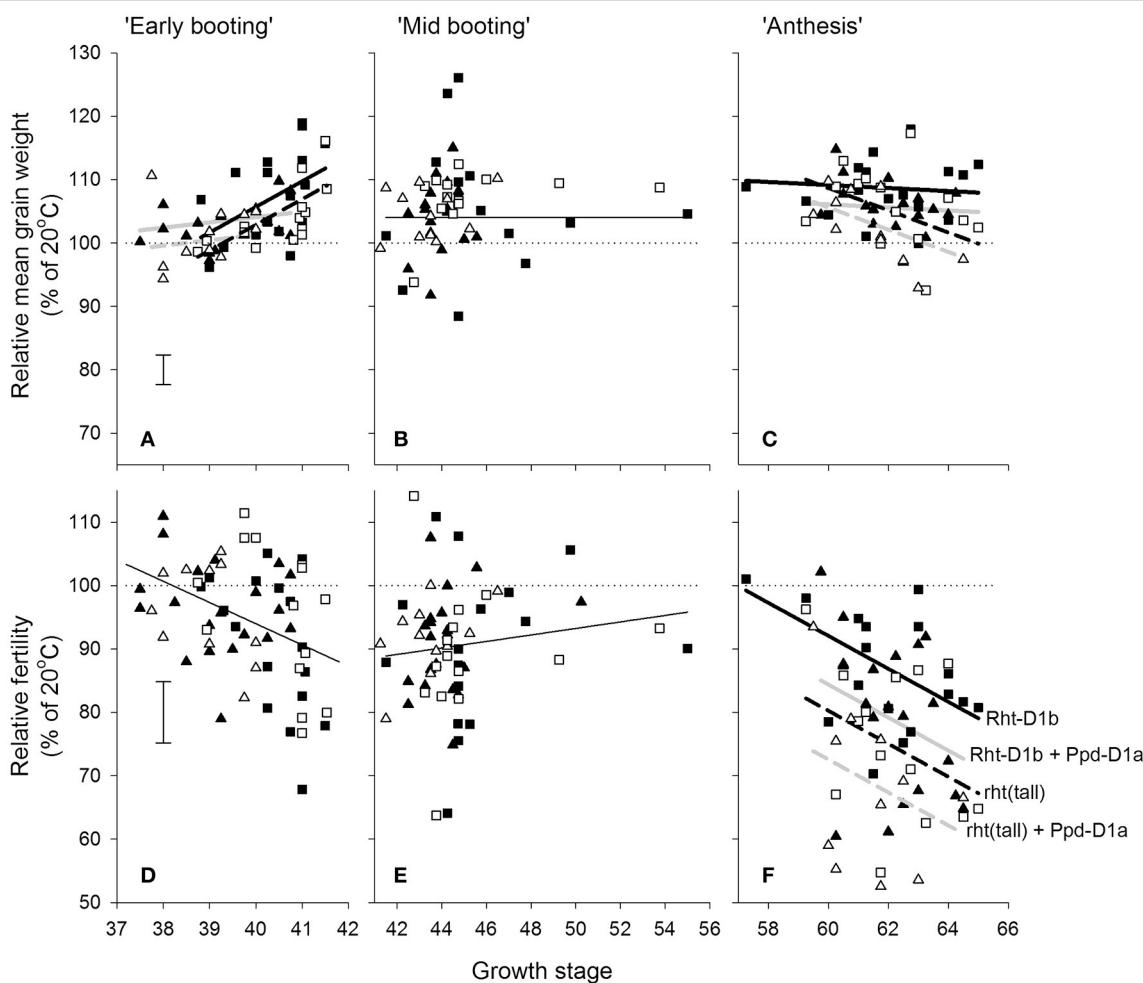
FIGURE 4 | Effects of increasing day temperature from 20 to 35°C in 1-day transfers to controlled environment cabinets and growth stage (A,D = early booting; **B,E** = mid booting; **C,F** = Anthesis) on mean grain weight **(A–C)** and grain numbers **(D–F)** from main stems of the doubled haploid progeny of Savannah × Renesansa marked for with (solid symbols) and without (open) *Rht-D1b* and with (triangles) and without (squares) *Ppd-D1a*. Error bars are S.E.D.s for comparing any point with the 100% line. In **A,C,F** lines are fits corresponding to markers as described in **(F)**: with (solid) and without (dashed) *Rht-D1b*; and with (light line) and without (heavy line) *Ppd-D1a*.

in future climates (Godfray et al., 2010; Semenov et al., 2014).

It is necessary to acknowledge the possible confounding effects between heat stress tolerance and water deficit (Barnabas et al., 2008; Alghabari et al., 2014) in these experiments. However, the deficits below FC reported here at the end of pot transfer, and the durations over which significant deficits could have occurred, are considered to be relatively minor compared with the results from experiments with longer periods of stress (Gooding et al., 2003; Alghabari et al., 2014). Nonetheless, booting is known to be a period particularly susceptible to drought (Barber et al., 2015) and future work on identifying tolerant traits to abiotic stresses will require consideration of the combination of drought and heat stress.

There has previously been some suggestion that the semi dwarfing allele *Rht8*, commonly found in southern European genotypes of wheat (Worland, 1996; Gasperini et al., 2012), could also increase tolerance to heat and drought stress compared to

TABLE 3 | Quantitative trait loci for relative fertility (%) in response to heat stress during anthesis (grain numbers following 1 day transfer to 35°C as a percentage of that achieved at 20°C).

Chromosome	Closest marker	LOD	Additive effect	Source of protecting allele	Effect (%)
2A	*Xgwm448*	7.02	−7.1971	Renesansa	38.1
2D	*Ppd-D1*	2.11	3.7296	Savannah	7.1
4D	*Rht-D1*	3.77	5.2518	Savannah	16.7

other semi dwarfing alleles (Alghabari et al., 2014). However, our study found no effect of *Rht8* on susceptibility to heat stress. This suggests that even in future climates, *Rht8* would not be of benefit to northern European genotypes due to its lower yield in comparison to other semi dwarfing alleles (Rebetzke et al., 2007). Furthermore, *Ppd-D1a*, to which *Rht8* is closely linked (Gasperini et al., 2012) was shown to increase susceptibility to

heat stress. Photoperiod insensitivity caused by the allele *Ppd-D1a*, a mechanism used to avoid abiotic stress (Gomez et al., 2014), is widely considered to be a beneficial trait in future climates due to reducing thermal time to senescence (Barber et al., 2015), thereby avoiding late season heat and drought stress. It was also suggested by Jones et al. (2017) that the increase in flowering duration associated with *Ppd-D1a* would add further resilience by increasing diversity of flowering timing within a field. However, the increase in susceptibility to heat stress associated with this allele, as well as lower overall grain yield in non-stressed seasons (Addisu et al., 2010) casts doubt over the benefits that *Ppd-D1a* might bring under future northern European climates. Although the introduction of *Rht-D1b* in to Northern European wheats has increased yield through increased harvest index and reduced lodging in fertile conditions (Flintham et al., 1997), it has also been associated with some negative traits, including decreases in fertility (Law et al., 1981). Preliminary work by Law and Worland (1985) suggested that the decrease in GA sensitivity caused by *Rht-D1b* increases susceptibility to heat stress. This is supported by later work in other cereals, such as barley, which shows that reducing sensitivity to GA increases susceptibility to heat stress (Vettakkorumakankav et al., 1999; summary provided by Maestri et al., 2002). However, our study shows evidence to the contrary. Here, *Rht-D1b* was associated with greater tolerance of high temperatures at anthesis than the other alleles associated with stature. In particular, the tall allele at the *Rht-D1* locus was associated with susceptibility to heat stress at anthesis. This contrasts with the effects of *Rht-D1* dwarfing alleles in some, but not all, backgrounds reported by Alghabari et al. (2014). We have found no genetic explanation for the poor performance of the Northern European genotype at booting. However, this can likely be attributed to the lack of selection pressure previously on breeding programmes for this trait.

With respect to the QTL analyses, others have also found regions on chromosomes on 2A and 2B to be associated with differential responses to heat stress (Mason et al., 2010; Talukder et al., 2014). Given the strength of the protective effect associated with the QTL on 2A further investigation is warranted for alleles in the relevant region from Renesansa. What is very clear from this study is that alleles and QTL detected as being associated with

heat stress tolerance is highly dependent on the precise growth stage of the plant when excessive heat is experienced.

CONCLUSIONS

In conclusion, this paper provides the strongest existing evidence that the key phases susceptible to heat stress at booting and anthesis in wheat are discrete and that genotypes vary with regards to the most susceptible growth stage. Periods of susceptibility are repeatedly observed during GS 41–45 and again from GS 61–65. In the prevailing conditions (mean daily temperature 14.3°C) periods of peak susceptibility could be separated by 15 days. We found no evidence that the southern European semi dwarfing allele *Rht8* adds tolerance to heat stress within NILs or a DH population. In contrast, the north European allele *Rht-D1b* was associated with increased tolerance to heat stress at anthesis. The photoperiod insensitivity allele *Ppd-D1a* was also found to be linked to increased susceptibility to heat stress.

AUTHOR CONTRIBUTIONS

HB, MG, and MS contributed to experimental design, HB and MG conducted analysis on the data with assistance from ML on interpretation of the data, whilst JS conducted QTL and genetic analysis. HB and MG drafted the work with revisions from MS, ML, and JS. HB, MG, MS, ML, and JS approve of the final version of the manuscript and all agree to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

REFERENCES

Addisu, M., Snape, J. W., Simmonds, J. R., and Gooding, M. J. (2010). Effects of reduced height (Rht) and photoperiod insensitivity (Ppd) alleles on yield of wheat in contrasting production systems. *Euphytica* 172, 169–181. doi: 10.1007/s10681-009-0025-2

Alghabari, F., Lukac, M., Jones, H. E., and Gooding, M. J. (2014). Effect of Rht alleles on the tolerance of wheat grain set to high temperature and drought stress during booting and anthesis. *J. Agron. Crop Sci.* 200, 36–45. doi: 10.1111/jac.12038

Barber, H. M., Carney, J., Alghabari, F., and Gooding, M. J. (2015). Decimal growth stages for precision wheat production in changing environments? *Ann. Appl. Biol.* 166, 355–371. doi: 10.1111/aab.12207

Barnabas, B., Jaeger, K., and Feher, A. (2008). The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell Environ.* 31, 11–38. doi: 10.1111/j.1365-3040.2007.01727.x

Craufurd, P. Q., Vadez, V., Jagadish, S. V. K., Prasad, P. V. V., and Zaman-Allah, M. (2013). Crop science experiments designed to inform crop modeling. *Agric. For. Meteorol.* 170, 8–18. doi: 10.1016/j.agrformet.2011.09.003

Dolferus, R., Ji, X., and Richards, R. A. (2011). Abiotic stress and control of grain number in cereals. *Plant Sci.* 181, 331–341. doi: 10.1016/j.plantsci.2011.05.015

Ferris, R., Ellis, R. H., Wheeler, T. R., and Hadley, P. (1998). Effect of high temperature stress at anthesis on grain yield and biomass of field-grown crops of wheat. *Ann. Bot.* 82, 631–639. doi: 10.1006/anbo.1998.0740

Flintham, J. E., Borner, A., Worland, A. J., and Gale, M. D. (1997). Optimizing wheat grain yield: effects of Rht (giberellin-insensitive) dwarfing genes. *J. Agri. Sci. Camb.* 128, 11–25. doi: 10.1017/S0021859696003942

Gasperini, D., Greenland, A., Hedden, P., Dreos, R., Harwood, W., and Griffiths, S. (2012). Genetic and physiological analysis of Rht8 in bread wheat: an alternative source of semi-dwarfism with a reduced sensitivity to brassinosteroids. *J. Exp. Bot.* 63, 4419–4436. doi: 10.1093/jxb/ers138

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., et al. (2010). Food security: the challenge of feeding 9 billion people. *Science* 327, 812–818. doi: 10.1126/science.1185383

Gomez, D., Vanzetti, L., Helguera, M., Lombardo, L., Fraschina, J., and Miralles, D. J. (2014). Effect of Vrn-1, Ppd-1 genes and earliness *per se* on heading time in Argentinean bread wheat cultivars. *Field Crop. Res.* 158, 73–81. doi: 10.1016/j.fcr.2013.12.023

Gooding, M. J., Ellis, R. H., Shewry, P. R., and Schofield, J. D. (2003). Effects of restricted water availability and increased temperature on the grain filling, drying and quality of winter wheat. *J. Cereal Sci.* 37, 295–309. doi: 10.1006/jcrs.2002.0501

Jones, H. E., Lukac, M., Brak, B., Martinez-Eixarch, M., Alhomedhi, A., Gooding, M. J., et al. (2017). Photoperiod sensitivity affects flowering duration in wheat. *J. Agri. Sci.* 155, 32–43. doi: 10.1017/S0021859616000125

Kowalski, A. M., Gooding, M., Ferrante, A., Slafer, G. A., Orford, S., Gasperini, D., et al. (2016). Agronomic assessment of the wheat semi-dwarfing gene *Rht8* in contrasting nitrogen treatments and water regimes. *Field Crops. Res.* 191, 150–160. doi: 10.1016/j.fcr.2016.02.026

Law, C. N., Snape, J. W., and Worland, A. J. (1981). *Reduced Fertility of Wheat Associated with Rht3*. Plant Breeding Institute Annual Report 1980, Plant Breeding Institute, Cambridge.

Law, C. N., and Worland, A. J. (1985). *An Effect of Temperature on the Fertility of Wheats Containing the Dwarfing Genes Rht1, Rht2 and Rht3*. Cambridge: Plant Breeding Institute.

Liu, B., Asseng, S., Liu, L., Tang, L., Cao, W., and Zhu, Y. (2016). Testing the responses of four wheat crop models to heat stress at anthesis and grain filling. *Glob. Change Biol.* 22, 1890–1903. doi: 10.1111/gcb.13212

Lobell, D. B., Hammer, G. L., Chenu, K., Zheng, B., Mclean, G., and Chapman, S. C. (2015). The shifting influence of drought and heat stress for crops in northeast Australia. *Glob. Change Biol.* 21, 4115–4127. doi: 10.1111/gcb.13022

Lu, L. L., Wang, C. Z., Guo, H. D., and Li, Q. T. (2014). Detecting winter wheat phenology with SPOT-VEGETATION data in the North China Plain. *Geocarto Int.* 29, 244–255. doi: 10.1080/10106049.2012.760004

Lukac, M., Gooding, M. J., Griffiths, S., and Jones, H. E. (2012). Asynchronous flowering and within-plant flowering diversity in wheat and the implications for crop resilience to heat. *Ann. Bot.* 109, 843–850. doi: 10.1093/aob/mcr308

Maestri, E., Klueva, N., Perrotta, C., Gulli, M., Nguyen, H. T., and Marmiroli, N.,(2002). Molecular genetics of heat tolerance and heat shock proteins in cereals. *Plant Mol. Biol.* 48, 667–681. doi: 10.1023/A:1014826730024

Mason, R. E., Mondal, S., Beecher, F. W., Pacheco, A., Jampala, B., Ibrahim, A. M. H., et al. (2010). QTL associated with heat susceptibility index in wheat (*Triticum aestivum* L.) under short-term reproductive stage heat stress. *Euphytica* 174, 423–436. doi: 10.1007/s10681-010-0151-x

Pradhan, G. P., Prasad, P. V. V., Fritz, A. K., Kirkham, M. B., and Gill, B. S. (2012). High temperature tolerance in aegilops species and its potential transfer to wheat. *Crop Sci.* 52, 292–304. doi: 10.2135/cropsci2011.04.0186

Prasad, P. V. V., and Djanaguiraman, M. (2014). Response of floret fertility and individual grain weight of wheat to high temperature stress: sensitive stages and thresholds for temperature and duration. *Funct. Plant Biol.* 41, 1261–1269. doi: 10.1071/FP14061

Rajaram, S. (2001). Prospects and promise of wheat breeding in the 21st century. *Euphytica* 119, 3–15. doi: 10.1023/A:1017538304429

Rebetzke, G. J., Richards, R. A., Fettell, N. A., Long, M., Condon, A. G., Forrester, R. I., et al. (2007). Genotypic increases in coleoptile length improves stand establishment, vigour and grain yield of deep-sown wheat. *Field Crop. Res.* 100, 10–23. doi: 10.1016/j.fcr.2006.05.001

Saini, H. S., and Aspinall, D. (1982). Abnormal sporogenesis in wheat (*Triticum-Aestivum* L) induced by short periods of high-temperature. *Ann. Bot.* 49, 835–846.

Saini, H. S., Sedgley, M., and Aspinall, D. (1983). Effect of heat-stress during floral development on pollen-tube growth and ovary anatomy in wheat (*Triticum-Aestivum*-L). *Aust. J. Plant Physiol.* 10, 137–144. doi: 10.1071/PP9830137

Saini, H. S., Sedgley, M., and Aspinall, D. (1984). Developmental anatomy in wheat of male-sterility induced by heat-stress, water deficit or abscisic-acid. *Aust. J. Plant Physiol.* 11, 243–253. doi: 10.1071/PP9840243

Satake, T., and Yoshida, H. (1978). High temperature-induced sterility in indica rices at flowering. *Jpn. J. Crop Sci.* 47, 6–17. doi: 10.1626/jcs.47.6

Satake, T., and Yoshida, H. (1981). High temperature stress in rice. *IRRI Res. Pap. Ser.* 67, 1–15.

Schlegel, R., and Korzun, V. (1997). About the origin of 1RS.1BL wheat-rye chromosome translocations from Germany. *Plant Breed.* 116, 537–540. doi: 10.1111/j.1439-0523.1997.tb02186.x

Semenov, M. A., Stratonovitch, P., Alghabari, F., and Gooding, M. J. (2014). Adapting wheat in Europe for climate change. *J. Cereal Sci.* 59, 245–256. doi: 10.1016/j.jcs.2014.01.006

Shewry, P. R., and Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food Energy Secur.* 4, 178–202. doi: 10.1002/fes3.64

Simmonds, J. R., Leverington-Waite, M., Wang, Y., Greenland, A., and Snape, J. W. (2006). *Discovering QTL Controlling Yield and Yield Components in Wheat., John Innes Centre*. Norwich: European Cereals Genetics Cooperative.

Snape, J. W., Foulkes, M. J., Simmonds, J., Leverington, M., Fish, L. J., Wang, Y., et al. (2007). Dissecting gene × environmental effects on wheat yields via QTL and physiological analysis. *Euphytica* 154, 401–408. doi: 10.1007/s10681-006-9208-2

Steinmeyer, F. T., Lukac, M., Reynolds, M. P., and Jones, H. E. (2013). Quantifying the relationship between temperature regulation in the ear and floret development stage in wheat (*Triticum aestivum* L.) under heat and drought stress. *Funct. Plant Biol.* 40, 700–707. doi: 10.1071/FP12362

Stone, P. J., and Nicolas, M. E. (1994). Wheat cultivars vary widely in their responses of grain yield and quality to short periods of post-anthesis heat stess. *Aust. J. Plant Physiol.* 21, 887–900. doi: 10.1071/PP9940887

Stratonovitch, P., and Semenov, M. A. (2015). Heat tolerance around flowering in wheat identified as a key trait for increased yield potential in Europe under climate change. *J. Exp. Bot.* 66, 3599–3609. doi: 10.1093/jxb/erv070

Talukder, S. K., Babar, M. A., Vijayalakshmi, K., Poland, J., Prasad, P. V. V., Bowden, R., et al. (2014). Mapping QTL for the traits associated with heat tolerance in wheat (*Triticum aestivum*.L.). *BMC Genet.* 15:97 doi: 10.1186/s12863-014-0097-4

Tashiro, T., and Wardlaw, I. F. (1990). The response to high-temperature shock and humidity changes prior to and during the early stages of grain development in wheat. *Aust. J. Plant Physiol.* 17, 551–561. doi: 10.1071/PP9900551

Vettakkorumakankav, N. N., Falk, D., Saxena, P., and Fletcher, R. A. (1999). A crucial role for gibberellins in stress protection of plants. *Plant Cell Physiol.* 40, 542–548. doi: 10.1093/oxfordjournals.pcp.a029575

Villareal, R. L., Banuelos, O., Mujeeb-Kazi, A., and Rajaram, S. (1998). Agronomic performance of chromosomes 1B and T1BL.1RS near-isolines in the spring bread wheat Seri M82. *Euphytica* 103, 195–202. doi: 10.1023/A:1018392002909

Wardlaw, I. F., Dawson, I. A., Munibi, P., and Fewster, R. (1989). The tolerance of wheat to high temperatures during reproductive growth. 1 survey procedures and general response patterns. *Aust. J. Agric. Res.* 1, 1–13. doi: 10.1071/AR9890001

Worland, A. J. (1996). The influence of flowering time genes on environmental adaptability in European wheats. *Euphytica* 89, 49–57. doi: 10.1007/BF00015718

Worland, A. J., Borner, A., Korzun, V., Li, W. M., Petrovic, S., and Sayers, E. J. (1998). The influence of photoperiod genes on the adaptability of European winter wheats (Reprinted from Wheat: prospects for global improvement, 1998). *Euphytica* 100, 385–394. doi: 10.1023/A:1018327700985

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). Decimal code for growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Determining Phenological Patterns Associated with the Onset of Senescence in a Wheat MAGIC Mapping Population

Anyela V. Camargo[1]*, Richard Mott[2], Keith A. Gardner[3], Ian J. Mackay[3], Fiona Corke[1], John H. Doonan[1], Jan T. Kim[4] and Alison R. Bentley[3]*

[1] National Plant Phenomics Centre, Institute of Biological Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, UK, [2] UCL Genetics Institute, University College London, UK, [3] The John Bingham Laboratory, National Institute of Agricultural Botany, Cambridge, UK, [4] The Pirbright Institute, Surrey, UK

The appropriate timing of developmental transitions is critical for adapting many crops to their local climatic conditions. Therefore, understanding the genetic basis of different aspects of phenology could be useful in highlighting mechanisms underpinning adaptation, with implications in breeding for climate change. For bread wheat (*Triticum aestivum*), the transition from vegetative to reproductive growth, the start and rate of leaf senescence and the relative timing of different stages of flowering and grain filling all contribute to plant performance. In this study we screened under Smart house conditions a large, multi-founder "NIAB elite MAGIC" wheat population, to evaluate the genetic elements that influence the timing of developmental stages in European elite varieties. This panel of recombinant inbred lines was derived from eight parents that are or recently have been grown commercially in the UK and Northern Europe. We undertook a detailed temporal phenotypic analysis under Smart house conditions of the population and its parents, to try to identify known or novel Quantitative Trait Loci associated with variation in the timing of key phenological stages in senescence. This analysis resulted in the detection of QTL interactions with novel traits such the time between "half of ear emergence above flag leaf ligule" and the onset of senescence at the flag leaf as well as traits associated with plant morphology such as stem height. In addition, strong correlations between several traits and the onset of senescence of the flag leaf were identified. This work establishes the value of systematically phenotyping genetically unstructured populations to reveal the genetic architecture underlying morphological variation in commercial wheat.

Keywords: wheat, senescence, data science, phenology, phenotyping, MAGIC

INTRODUCTION

Wheat is a pillar of global food security, providing 20% of protein and calories consumed worldwide and up to 50% in developing countries. It is the main food staple in Central Asia, West Asia and North Africa, which have the world's highest per capita wheat consumption (Valluru et al., 2015). Global wheat production is at risk due to climate change, population growth, changing food

preferences and the plant health challenges associated with its widespread cultivation. In order to maintain optimal production and profitability, wheat producers and processors must prepare for and adapt to these challenges. The current emphasis on food security has focused research attention on two avenues to improve wheat yield (Valluru et al., 2015): (1) increasing photosynthetic capacity and efficiency (Reynolds et al., 2009); and (2) increasing partitioning of assimilates to the developing spike and grain.

The timing of key developmental transitions is critical for many crops, but is particularly important in the temperate small grain cereals. For example, the transition from vegetative to reproductive growth can have major effects on biomass accumulation and harvest index that profoundly affect either the locales in which a variety can be profitably grown, or its ultimate use. Thus, crops destined for grain production should transition early, relative to the length of the growing season, to allow ripening, avoid stress, and achieve a high harvest index of grain to total biomass. Forage, biofuel or dual purpose crops could usefully transition later to allow greater total biomass accumulation, but this has to be tempered with the likelihood of deleterious stress/weather events. Flowering time, therefore, has been a key selection target since the beginning of domestication (Izawa, 2007), initially inadvertently but since modern breeding began, very directly.

Understanding the extent and basis of other aspects of phenological variation may be useful in breeding for yield potential and stress adaptation. In wheat, leaves can contribute up to 40% of the nitrogen incorporated by the grains on the fifteenth day after anthesis (Simpson et al., 1983). Therefore, lifespan of the leaves (hence yield) is a trade-off with N remobilization. Delayed leaf senescence (as in the stay-green effect), which maintains active photosynthesis for a longer period, can increase grain yields under certain circumstances (Gregersen et al., 2008). Conversely, accelerated senescence leads to low carbon (C) but high N remobilization, indicating plasticity of C and N remobilization during development, perhaps correlated with senescence. A better understanding the genetic and environmental factors affecting these processes would help optimize C and N remobilization to the actively developing grains under different growth/stress conditions.

While several patterns of senescence have been proposed (Thomas and Howarth, 2000), an ideal senescence phenotype in wheat, and in cereals in general, still needs to be identified (Gregersen et al., 2008), perhaps because of strong and variable environmental effects. In monocarpic crops such as wheat the initiation of senescence typically leads to a massive remobilization of phloem-mobile nutrients from the senescing plant parts to developing sinks, such as seeds or grains (**Figure 1**; Gregersen et al., 2008; Distelfeld et al., 2014). Pathogen infection also interacts with developmental processes in a complex way and symptoms of senescence often accompany the progression of disease, although senescence can also be delayed in response to pathogen infection (Häffner et al., 2015).

A major constraint to progress in breeding for high yield varieties is the access to appropriate and consistent selection environments. The selection environment plays a key role in the
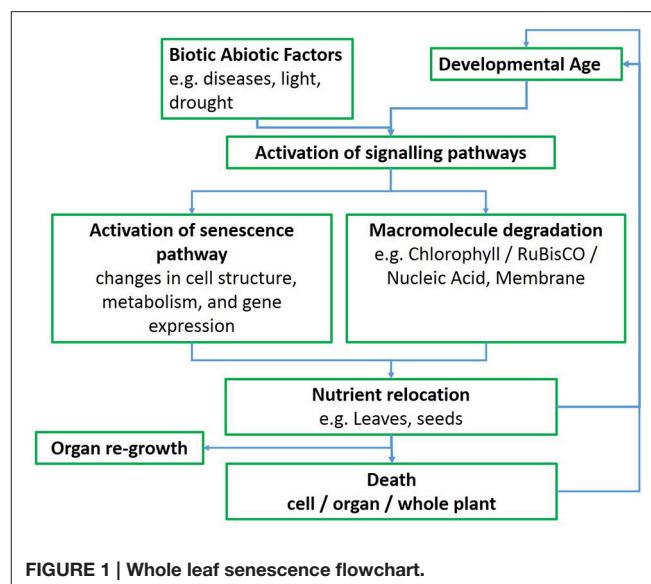
FIGURE 1 | Whole leaf senescence flowchart.

efficiency of the selection process. Since environmental variables are almost impossible to control under field conditions, the identification of specific genetic factors associated to crop yield becomes more challenging (Bentley et al., 2013).

Modern controlled environment (CE) growth control and/or recording of variable environmental parameters such as temperature and watering allow for elimination or reduction of uncontrollable influences. Having control and access to an experiment's environmental parameters allows for reproducibility, and decreases the levels of uncertainty as it is easier to reverse-engineer an experiment in order to identify—or at least justify—the causes of a given phenotype, and minimizes the amount of replication per subject due to the low variability of the environment. Therefore, CE phenotyping offers closely defined conditions compared to the relatively homogeneous but less controllable growing conditions in a field plot.

To understand the genetic control of phenology and the onset of senescence in wheat, we screened a core set of the NIAB elite MAGIC wheat population (Mackay et al., 2014) across time. The eight founders of this MAGIC population were selected in partnership with UK wheat breeders to sample trait variation and germplasm important to current UK breeding programmes (Bentley et al., 2013). MAGIC populations combine high levels of genetic diversity, recombination and homozygosity (Mackay et al., 2014) to create a panel of recombinant inbred lines (RILs). A well-designed MAGIC population captures and immortalizes the variation released by intercrossing, thereby providing a stable well-defined population to be shared across sites and used across years.

In this study, plants from the elite MAGIC core set and its parents were scored throughout their life cycle to capture traits, including decimal growth stages, biomass and plant height. We discuss the use of a subset of MAGIC lines in the Smarthouse as a proof of concept that the combination of MAGIC + Smarthouse phenotyping should be repeated on a grander scale.

MATERIALS AND METHODS

Plant Material

A subset of the NIAB Elite eight-founder MAGIC population described in Mackay et al. (2014) was used for all phenotypic screening. The complete population consists of approximately 1000 recombinant inbred lines (RILs) generated from three cycles of recombination between eight elite United Kingdom wheat varieties (Alchemy, Brompton, Claire, Hereward, Rialto, Robigus, Soissons, Xi-19) followed by five rounds of selfing to derive RILs. Further information about the population, including pedigree, genotype, and existing phenotype data can be found at www.niab.com/MAGIC.

The core set used in this study was selected to represent all funnels of the 210 8-way crosses within the population. Two funnels were not represented due to limited seed availability making a total of 208 RILs in the core set.

Glasshouse Cultivation

Plants were grown between mid-January 2015 and mid April 2015 in The National Plant Phenomics Centre facilities in Aberystwyth, UK. The eight parents of the MAGIC population and four additional elite varieties (Avalon, Santiago, Cadenza, and Zircon) were grown with the 208 RILs (see Table S1) under well watered conditions, with two replicates per genotype. Two seeds were sown in 8×8 cm pots of Levington F2 compost. After germination (approximately on the 30/10/2014) the seedlings were thinned to one per pot and transferred to a controlled environment room for vernalization (5°C, 16 h daylength) for 9 weeks. Following vernalization plants were transferred to $15 \times 15 \times 20$ cm pots of M2 compost. Field capacity and dry matter content of the compost was determined. Plants were transferred to the growth chamber where each pot was placed into a cart on a conveyor system. Pots were weighed and watered automatically to 75% gravimetric water content daily. Growth conditions were 14 h daylength using 600W sodium lamps to supplement (350 μM/m²/sec) natural lighting, with the temperature settings of 18°C (day) and 15°C (night). Plant hygiene was monitored by visual inspection throughout the experiment, with an appropriate prophylactic and responsive spraying regime. Once the ears started to ripen, plants were removed from the system and allowed to finish ripening naturally with reduced watering. Reduced watering only occurred after all plants had passed Flag leaf senescence.

Phenotyping

Plants were manually scored for developmental stages according to the Zadoks scale (Zadoks et al., 1974) three times per week, and scored as days after sowing (DAS) when plants reached growth stage 39 (GS39; flag leaf fully emerged), GS55 (ear 50% emerged), GS65 (50% anthesis), and the onset of flag leaf senescence (FLS). At the end of the experiment, plants were harvested and above ground biomass (PW), tiller number (TN), plant height (PH), stem height (SH), top internode length (TIL), first/second/third ear length (FEL, SEL, TEL), first/other ear weight (FEW/OEW), and first flag leaf length (FFLL) were scored. The number of days between GS39 and GS55 (d1), GS55 and GS65 (d2) and GS55

and FLS (d3) were also determined. A multiple linear regression model (MLRM) was fitted to identify predictors of FLS among all the traits used in the analysis. A list of traits, abbreviations and Crop ontology terms (Shrestha et al., 2012) are provided in **Table 1**.

Genotyping

The lines were genotyped using the Illumina Infinium iSelect 80,000 SNP wheat array ("80K array," http://www.illumina.com/), described in Wang et al. (2014). 20,639 SNP markers were scorable and polymorphic, of which 18,601 were successfully mapped in the MAGIC population (Mackay et al., 2014; Gardner et al., 2016). Linkage map generated by mpMap is reported in Gardner et al. (2016).

Plant Stress

The young plants showed symptoms that included chlorosis and necrosis (Figure S1). The chlorotic symptoms consisted of yellow areas surrounding lesions on the leaf blades. The necrotic symptoms comprised brown spots, lens-shaped lesions, surrounded by yellow borders. Although symptoms were controlled by routine spraying (Priori Xtra, Syngenta), we speculated that it constituted an undiagnosed disease (possibly Septoria) and the degree of infection was scored manually using the seedling infection type (IT) score shown in Table S1. Plant visual stress symptoms were scored first at GS31–39 and for the second time around GS70–GS80, and the average was calculated (SM). These qualitative IT scores were converted to a numerical scale for statistical analysis.

Statistical and Quantitative Trait Locus Analysis

Statistical analyses were performed in the R environment using Core Team (2013). Quantitative Trait Locus (QTL) analysis was performed using the R package HAPPY for multi-parental populations analyses (Mott et al., 2000). The genetic analysis of multi-parental populations requires a haplotype-based approach because single marker association or interval mapping can fail to detect a QTL if the causative alleles are not dispersed among the founders with the same strain distribution pattern as the linked markers (Mott et al., 2000).

QTL Mapping

HAPPY's analysis is essentially two stage; ancestral haplotype reconstruction using dynamic programming, followed by QTL testing by linear regression:

- Assume that at a QTL, a pair of chromosomes originating from the progenitor strains, labeled s, t contribute an unknown amount T_{st} to the phenotype. In the special case where the contribution from each chromosome is additive at the locus then $T_{st} = T_s + T_t$, say
- A test for a QTL is equivalent to testing for differences between the T's.
- A dynamic-programming algorithm is used to compute the probability F_{iLst} that a given individual i has the ancestral

TABLE 1 | Plant trait descriptions.

No.	Abbreviation in paper	Trait description (wheat ontology)	Unit	Wheat ontological abbreviation	Trait ID (TO or CO)
1	GS39	Growth/developmental stages based on zadoks decimal code: Flag leaf fully emerged	DAS*	GrwStg Zadok 39	CO_321:0000016
2	GS55	Growth/developmental stages based on zadoks decimal code: Ear 50% emerged	DAS	GrwStg Zadok 55	CO_321:0000016
3	GS65	Growth/developmental stages based on zadoks decimal code: Anthesis time	DAS	GrwStg Zadok 65	CO_321:0000016
4	FLS	Onset of flag leaf senescence. FLS is scored on the day after sowing when greater than 2.5 cm at the tip the primary flag leaf was senescent (color changed from green to yellow/brown)	DAS	tFleafSen	TO:0000249
5	PW	All above-ground biomass at maturity	g/plant		CO_321:0001431
6	TC	Tiller number	Tiller/plant		CO_321:0000190
7	SH	Stem height (it is stem height to the base of the ear for primary tiller)	cm		CO_321:0000060
8	TIL	Top internode length	cm		TO:0000145
9	FEL	First (primary) ear length	cm	SpkLng	CO_321:0000056 TO:0000431
10	SEL	Second ear length	cm	SpkLng	CO_321:0000056 TO:0000431
11	TEL	Third ear length	cm	SpkLng	CO_321:0000056 TO:0000431
12	FEW	First (primary) ear weight	g		n/a
13	OEW	Other ear weight	g		n/a
14	TEW	Total ear weight (FEW + OEW)	g		n/a
15	FFLL	First (primary) flag leaf length	cm	FLFLG	TO:0002757
16	d1	Number of days between GS39 and GS55	Days		n/a
17	d2	Number of days between GS55 and GS65	Days		n/a
18	d3	Number of days between GS55 and FLS	Days		n/a
19	SM	Plant stress score mean	Observational score		n/a
20	HI	Harvest Index	Ratio	HI	TO:0000128

TO, Trait Ontology (Liang et al., 2008).
CO, Crop Ontology (Shrestha et al., 2012).
*DAS, Days after sowing.
Wheat ontology (Leo Valette, Bioversity, France, Personal Communication).
Unit is the metric of thetrait. Trait id is the TO or CO reference id.

alleles s, t at locus labeled L, conditional upon all the genotype data for the individual. Then the expected phenotype is

$$y = \sum_{st} T_{st} F_{iLst},$$

and the T's are estimated by a linear regression of the observed phenotypes on these expected values across all individuals, followed by an analysis of variance to test whether the progenitor estimates differ significantly.

- The method's power depends on the ability to distinguish ancestral haplotypes across the interval.
- All inference is based on regression of the phenotypes on the probabilities of descent from the founder loci, F_{nst}.

The models are presented here in the linear model framework (i.e., least-squares estimation, with ANOVA F-tests).

For an additive QTL, the parameters are the strain effect sizes; for a full interaction model there is a parameter for every possible strain combination. Then the one-QTL model is $E(y) = X_L t_L$.

There are $S(S-1)/2 + S$ parameters (where S is the number of strains) to be estimated in a full model allowing for interactions between the alleles within the locus, and $S-1$ parameters in an additive model. For the full model, the i, j'th element of the design matrix X is related to the strain probabilities thus:

$$X_{Lij} = F_{iLst},$$

where

$$(s, t) = in(s + S(t-1)), t + S(s-1)$$

and for the additive model

$$X_{Lij} = \sum_s F_{iLsj}$$

We used an additive model, where the contribution of each allele at the locus are assumed to act additively.

Furthermore, when mapping QTLs in structured populations the evidence for the existence of a QTL has to be considered

in the context of other QTLs, which might explain some of the same component of variation. Population structure can produce long-range correlations between genotypes and hence ghost QTL, although the LD analysis suggests that the MAGIC population is relatively immune to this phenomenon. Although the MAGIC population is relatively unstructured, and therefore can be analyzed one locus at a time, in order to ensure the evidence for a given QTL was not confounded with that for others, statistical significance was assessed based on permuting the phenotypes (1000 times) between individuals, repeating the model fit, and finding the top-scoring marker interval. The empirical distribution of the max $-\log P$ values was then used to assess statistical significance. This technique is useful for non-normally distributed phenotypes and for estimating region-wide significance levels. We used $-\log P = 4$ as a threshold in the multiple QTL modeling to test for association and FDR = 0.05 to identify significantly differential markers. The dashed line in QTL plots corresponds to an FDR rate of 0.05 and is calculated using the qvalue package (Storey and Tibshirani, 2003). The p-value corresponding to a q-value of 0.05 is determined by interpolation. When there are no q-values less than 0.05, the dashed line is omitted.

In addition, a multiple linear regression model (MLRM) was fitted to identify predictors of FLS among all the traits used in the analysis. We selected FLS because the trait is used as indicator of crop yield and biomass accumulation (Gan, 2014).

Principal Component Analyses (PCA) over normalized trait data was carried out to identify patterns between traits and genotypes. Biplots were used to show information on samples in a graphical manner (Kempton, 1984). PCA of marker data was carried out separately to test for population structure.

QTL Validation

We used the R package mpMap to confirm QTL mapping results (Huang and George, 2011) and to analyse the effect of including marker covariates. QTL analysis is performed using interval mapping, then selected marker covariates are included in the linear model in a forward selection process.

Data Processing

Data were pre-processed using standard methods. Data corresponding to one replicate of the MAGIC line MEL 086-1 (note, all MAGIC lines are named with prefix "MEL") and another from one replicate of the elite line Avalon were removed due to seed infection. A small number of outliers (data points with suspicious values) were checked and, where possible, corrected. Missing values (2.19%) were imputed using multivariate imputation by chained equation (MICE) (van Buuren and Groothuis-Oudshoorn, 2011). Briefly, MICE operates under the assumption that given the variables used in the imputation procedure, the missing data are Missing At Random (MAR), which means that the probability that a value is missing depends only on observed values and not on unobserved values (Schafer and Graham, 2002). MICE creates a number of datasets by imputing missing values. That is, one missing value in original dataset is replaced by m plausible imputed values. We set $m = 5$ as the number of imputations. These values take imputation uncertainty into consideration. Statistics of interest are estimated from each dataset and then combined into a final one and replicates were averaged (Zhang, 2016).

RESULTS

In this experiment, RILs, MAGIC parents (illustrated in **Figure 2**) and 4 other elite genotypes were grown to maturity
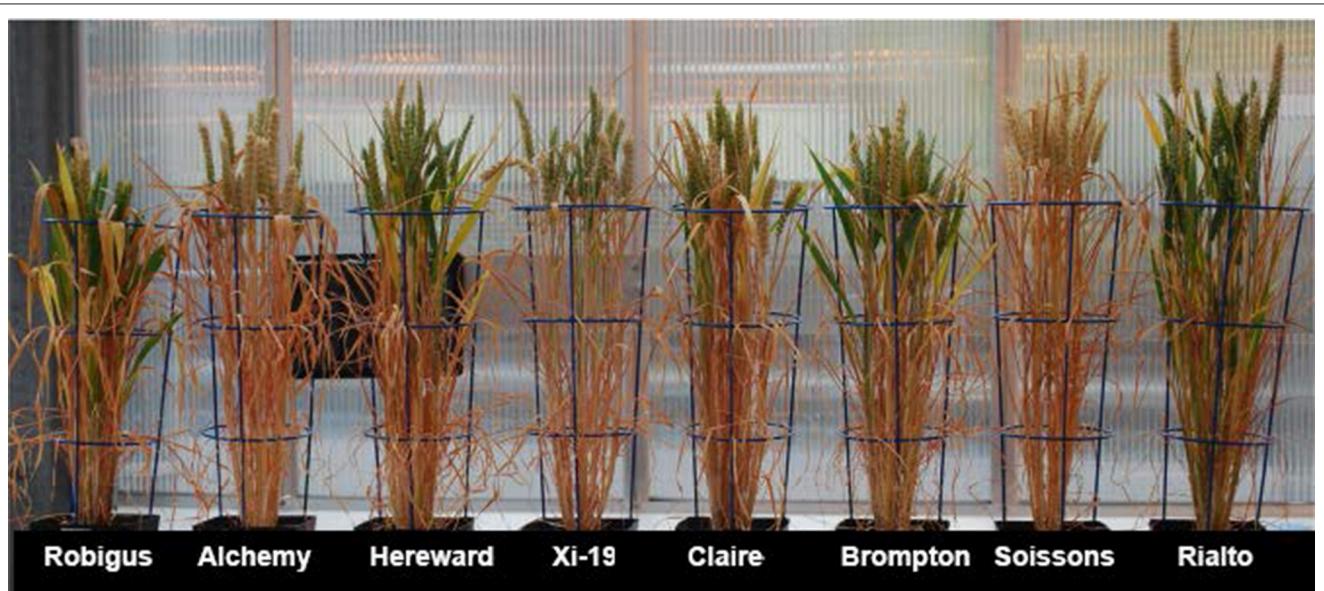
FIGURE 2 | MAGIC parents at senescence.

over a single time span and within a single glasshouse chamber with controlled watering and supplementary lighting and heating. Phenotype data were curated, and missing values (accounting for 2.19% of data points) imputed. Figure S2 shows a comparison between original data (red dots) and imputed data (blue dots), which suggested a high similarity between the two distributions as indicated by the overlapping dots.

Analysis of Traits

The distributions of the traits are shown in Figure S3. Most traits showed similar distributions with the exception of the discrete trait SM, which was skewed (see **Table 1** for trait description). SM is a discrete trait and the skewedness of the plot reflects that most RILs' scores were in the 0–2 range. **Figure 3** shows the frequency distributions for GS55, FLS, SH and d3. Pair-wise correlation analysis between all traits (Figure S3) identified strong correlations between FLS and GS39 (0.79), GS55 (0.73) and GS65 (0.69) and d3 (0.70); between FEL, SEL, and TEL (>0.86) between TIL and SH (0.76), PW and OEW (0.87), and between TEW and PW (0.87).

To determine if there was variation in duration between key developmental stages that was not simply a result of variation in overall developmental progression, we examined the time in days taken to progress from GS39 to GS55 (d1), from GS55 to GS65 (d2) and from GS55 to FLS (d3). Our MLRM identified d3 as a strong predictor ($P < 0.05$) of FLS, indicating that the time lapse between GS55 and FLS is a good candidate to predict FLS (**Figure 4**). In addition to d3, our MLRM identified other important predictors of FLS ($P < 0.05$). For example, the size of the flag leaf on the primary shoot, FFLL, was significantly ($P < 0.05$) associated with the timing of senescence (FLS). To demonstrate this result, dot size in **Figure 4** was used to represent an additional feature in the plots. In the case of **Figure 4**, FFLL (represented by dot size) was longer in RILs that senesced earlier. Among the MAGIC founders, Brompton, Hereward and Rialto senesced after Xi-19 and the two elite controls Zircon and Cadenza. The latter also had the shortest duration between GS39 and GS55. Previously, Mackay et al. (2011) reported Cadenza as the most environmentally sensitive variety detected in 8 years in Recommended List trials showing a linear increase in yield with increasing summer rainfall. This
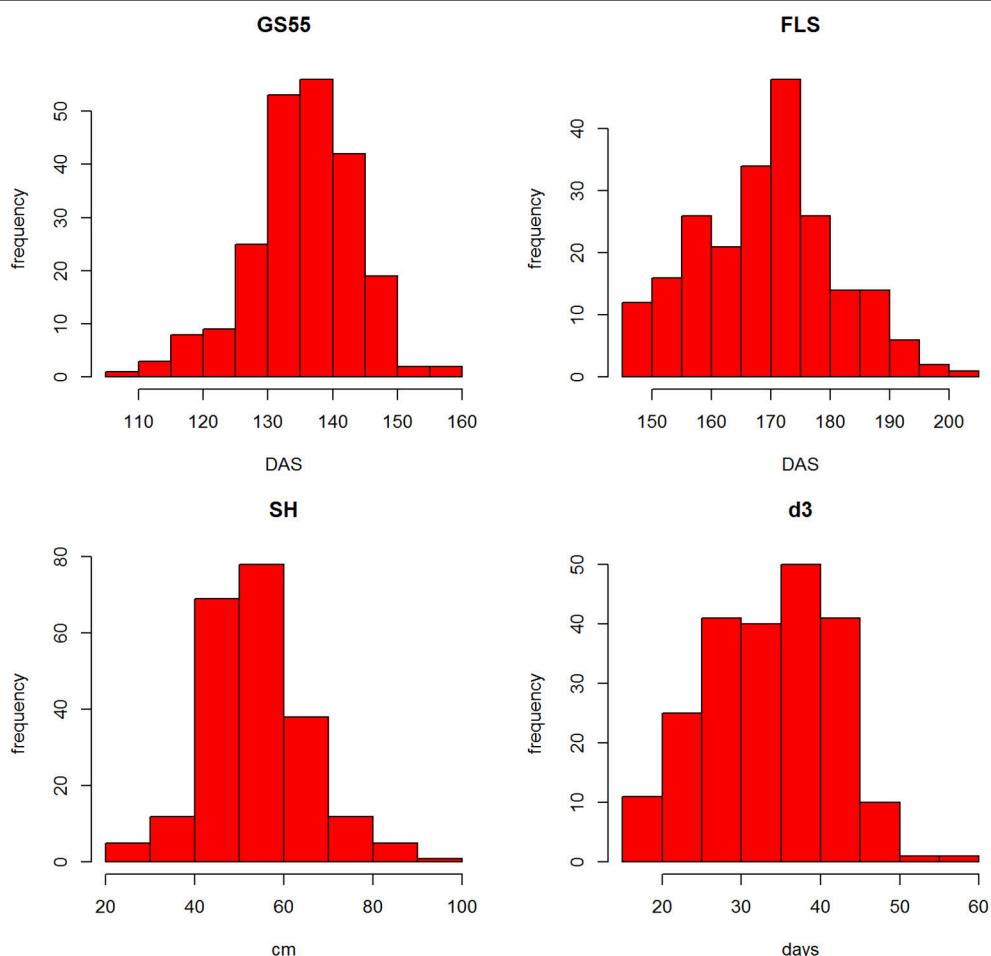
FIGURE 3 | Frequency distributions of GS55, FLS, SH, and d3 from all RILs.
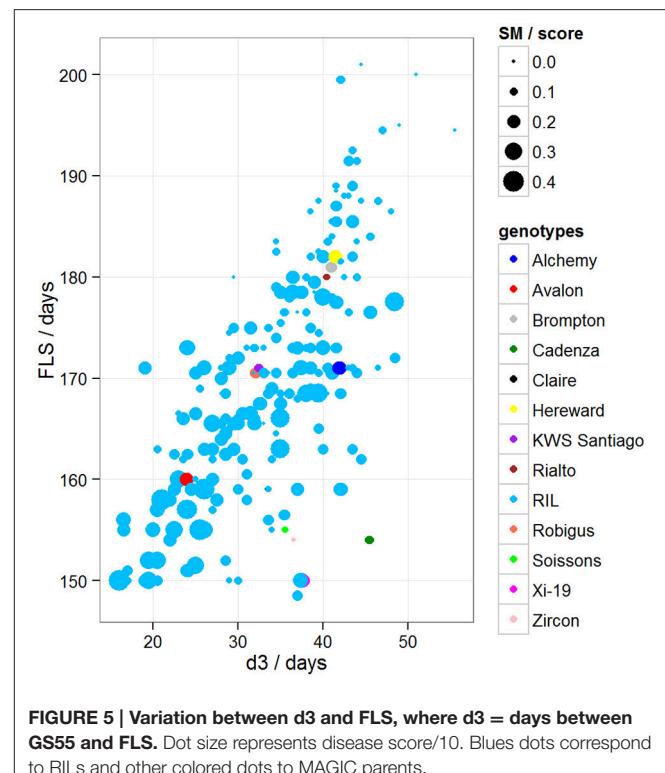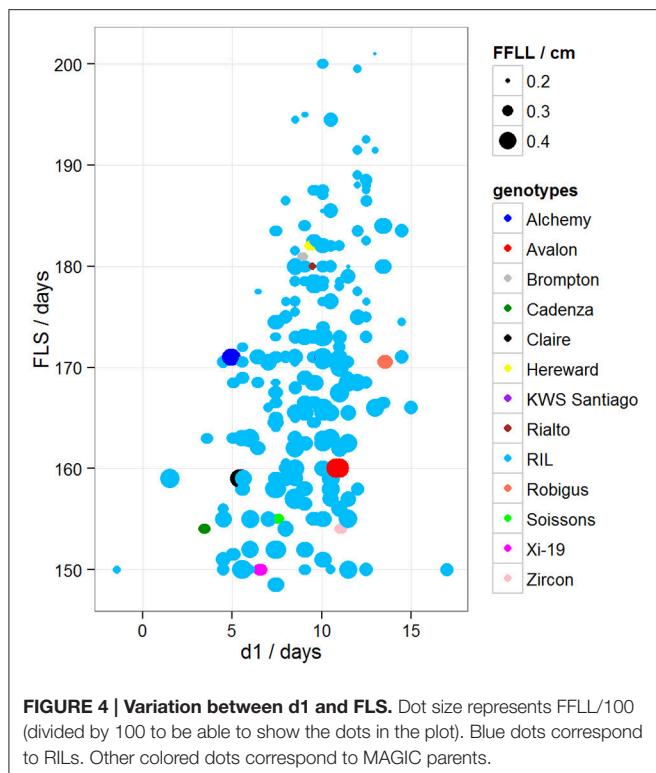
FIGURE 4 | Variation between d1 and FLS. Dot size represents FFLL/100 (divided by 100 to be able to show the dots in the plot). Blue dots correspond to RILs. Other colored dots correspond to MAGIC parents.

FIGURE 5 | Variation between d3 and FLS, where d3 = days between GS55 and FLS. Dot size represents disease score/10. Blues dots correspond to RILs and other colored dots to MAGIC parents.

supports the observations from the trait analysis that progression through different developmental processes e.g., flowering vs. senescence, is controlled independently. **Figure 5** shows d3 in relation to FLS (the contract of d2 to FLS is shown in Figure S5). Plants which senesced earlier took less time between GS55 and GS65 (d2) and between GS55 and FLS (d3).

To further evaluate the relationship between these traits, PCA was conducted over trait data scaled to have unit variance. Results of the analysis are shown on the biplot in **Figure 6**. The plot shows that PC1 and PC2 account for 48% of the total variance of the traits. Also, four clearly defined trait groups (anti clockwise) can be seen in the plot, the first one containing FLS, GS39, GS55 and GS65, d1, d3, and TN, the second group contained SEL, FEL and TEL, the third group contained OEW, PW, TIL, SH, FEW, and FFL and the fourth group contained HI and SM. Since (1) the smaller the angle between the trait vectors, the higher the correlation (2) trait values are smaller toward the middle of the plot and higher toward the edge, we can deduce that SM is negatively correlated to group one which is confirmed by the results from the correlation shown in Figure S4. The same argument could also be used between group one and group three, which is also confirmed by Figure S4.

Looking at the relationship between traits and MAGIC parents, Brompton and Hereward are positively correlated, tending to be slow to FLS (late to senesce) as indicated by their proximity to the FLS vector. In contrast, Soissons and Xi-19 show an opposite effect, rapidly reaching FLS as indicated by their location at the opposite side of FLS. Through this analysis we can also see that, in general, most RILs and parents have a

similar overall phenome, as represented by their location close to the center of the plot. We can also see there are a number of divergent phenotypes, such as the one at the bottom of the plot (**Figure 6**, right panel) which corresponds to MAGIC line MEL 091-1a or the one at the top which correspond to MEL 089-1a. Looking closer, MEL 089-1a is proximal to PW and SH while MEL 091-1a is further away, which indicates these two lines contrast strongly for these particular traits. A picture of both lines taken on 27/04/2015 was added to the plot to facilitate interpretation. The plants show clearly contrasting differences in height and biomass.

Plant Stress Analysis

A low level of chlorotic and necrotic lesions was observed on the leaves early during the growth period. Symptoms (SM) were scored independently by two people at two time points and scores were averaged. When comparing symptom scores against the onset of FLS, we found that the more severely affected plants started senescence earlier than those plants that were mildly affected (**Figure 5**). Our MLRM also identified SM as a predictor of FLS ($P < 0.05$).

Figure 6 confirms the negative correlation between FLS and SM as indicated by its opposite location from FLS. This correlation is consistent with that of *Mycosphaerella graminocola*, where infection induces senescence by manipulating signaling pathways in plants (Mengiste, 2012). However, it should be noted that the precise identity of the putative pathogen could not be confirmed.
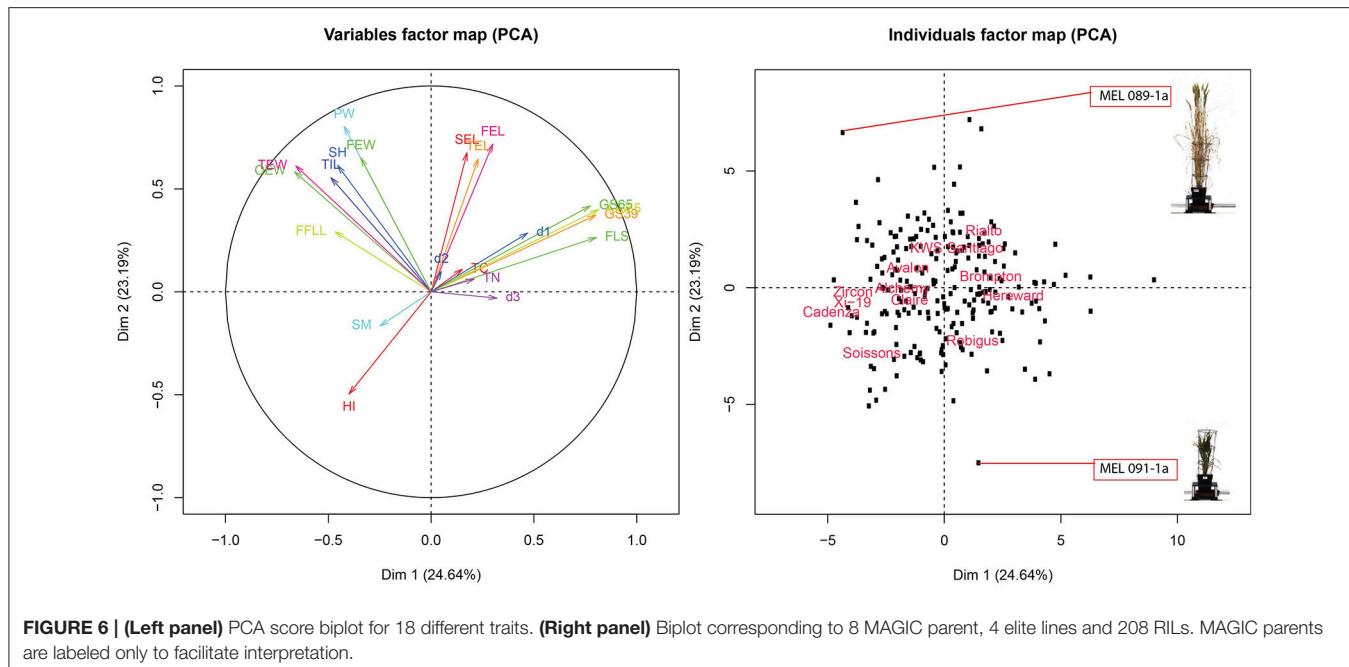
FIGURE 6 | (Left panel) PCA score biplot for 18 different traits. **(Right panel)** Biplot corresponding to 8 MAGIC parent, 4 elite lines and 208 RILs. MAGIC parents are labeled only to facilitate interpretation.

Quantitative Trait Loci

We evaluated whether trait variation could be ascribed to underlying genetic variation. The lines were genotyped using the Illumina Infinium iSelect 80,000 SNP wheat array ("80K array," http://www.illumina.com/), described in Wang et al. (2014). 20,639 SNP markers were scorable and polymorphic, of which 18,601 were successfully mapped in the MAGIC population (Mackay et al., 2014; Gardner et al., 2016); linkage map for this population was produced using mpMap and reported in Gardner et al. (2016). First, we checked for signs of population structure. To do this, the marker based relationship matrix (A) was calculated using the R package rrBLUp (Endelman, 2011), then a PCA analysis by eigenvalue decomposition of A was calculated. Results are shown in Figure S7. This shows that the first PC accounts for less than 4% of the total spectrum. This confirms the expected absence of population structure. Genome mosaics corresponding to MEL 15-2, MEL 091-1a, and MEL 209-1 are shown in Figure S8. These decompose the lines' genomes into mosaics of founder haplotypes. The lines appear to be a random mix of the founders, which indicates an absence of gross population structure.

After confirming the absence of population structure, we used HAPPY (Mott et al., 2000) to test for association between each phenotype and the predicted founder haplotypes at each locus in the genome. We used the P-value threshold $<10^{-4}$ to call QTLs [$-\log P = 4$, corresponding to a false discovery rate (FDR) = 0.05]. This analysis identified loci associated with two phenological traits, GS39 and GS55, and a number of traits such as SEL, SH, TIL, TEW, and HI. For GS39, three significant QTLs were found on chromosome 5A, at 201.36, 212.52, and 224.64 cM, corresponding to the markers BS00009369_51, BS00021942_51

and wsnp_Ex_c5978_10478584, respectively (**Table 2**, **Figure 7**). For GS55, three QTLs were found on chromosome 5A, at 201.36, 216.05 and 227.66 cM, which correspond to BS00009369_51, wsnp_Ex_rep_c66689_65011117, and Excalibur_c7729_144, respectively (**Figure 8**). In both cases, these three close peaks are likely to represent a single QTL. To confirm this hypothesis, we performed composite interval analysis using 5A as covariate and identified a single clear and strong marker on 5A (Figures S11A–E).

For SEL, four QTLs were identified on chromosome 5B at 60.65, 66.38, 90.8, 92.31 cM, corresponding to wsnp_Ex _c6548_11355524, BS00001101_51, wsnp_Ku_c2185_4218722, and RAC875_c19099_434, respectively (**Figure 9**).

For SH, one QTL was identified on chromosome 4D at 32.24 cM which corresponded to the marker RAC875_c1673_193 (**Figure 10**). A QTL for TIL was also identified in the same location (**Figure 11**). QTLs for OEW and PW were also identified on 4D at 26.97 cM corresponding to RAC875_c6922_291 (Figure S10A) and 4D at 32.24 cM corresponding to RAC875_c6922_291 (Figure S10B), respectively. We identified one QTL for TEW on chromosome 4D at 26.97 cM, corresponding to RAC875_c6922_291 (Figure S10D). All of these QTLs co-located with the semi-dwarfing gene *Rht-D1* (Rht2) (Ellis et al., 2002). The Rht (Reduced height) genes *Rht-B1* (Rht1) or *Rht-D1* (Rht2) are present in many high-yielding, semi-dwarf varieties, where they offer simple genetic control of high harvest index and resistance to lodging (Flintham et al., 1997).

For HI, 1 weak peak QTL on chromosome 2D at 55.4 cM was identified which corresponded to RAC875_c6922_291 (Figure S10E). Figure S9 shows a contrast between FLS and HI across all the MAGIC and elite lines. Cadenza and Soissons have some of the highest HIs and the shortest time to senesce. **Table 2** also

TABLE 2 | QTLs mapped for different traits.

Phenotype	Markers	Chr	cM	−logP	h²	GW P-value	Alchemy	Brompton	Claire	Hereward	Rialto	Robigus	Soissons	Xi-19
GS39	BS00009369_51	5A	201.36	8.04	0.35	0	127.34	127.15	125.21	128.42	128.01	125.13	127.07	111.46
GS39	BS00021942_51	5A	212.52	4.44	0.24	0.06	126.75	127.32	125.39	127.76	129.01	125.5	126.79	116.98
GS39	wsnp_Ex_c5978_10478584	5A	224.64	9.14	0.38	0	126.54	127.01	126.27	127.48	128.24	125.8	126.92	112.83
GS55	BS00009369_51	5A	201.36	7.13	0.33	0	136.79	135.89	134.09	138.57	136.96	134.49	136.99	119.18
GS55	wsnp_Ex_c37943_45584325	5A	216.05	4.64	0.25	0.03	135.3	135.57	134.37	138.38	137.99	134.57	136.39	127.6
GS55	Excalibur_c7729_144	5A	227.66	8.32	0.36	0	135.6	135.43	135.9	137.54	137.82	135.01	136.59	120.31
HI	Kukri_c27309_590	2D	55.4	4.97	0.01	0.01	0.55	0.55	0.55	0.55	0.55	0.55	0.6	0.55
OEW	RAC875_c6922_291	4D	26.97	4.09	0.21	0.06	37.4	37.14	37.17	36.3	36.81	42.6	39.2	35.57
PW	RAC875_c1673_193	4D	32.24	4.79	0.24	0.02	70.96	71.2	71.48	71.33	68.74	88.5	82.52	71.26
SEL	wsnp_Ex_c6548_11355524	5B	60.65	3.99	0.22	0.09	11.2	11.18	11.21	11.18	11.16	11.18	10.45	10.02
SEL	BS00001101_51	5B	66.38	4.01	0.22	0.09	11.12	11.19	11.19	11.13	11.17	11.33	10.36	10
SEL	wsnp_Ku_c2185_4218722	5B	90.8	4.15	0.23	0.07	11.08	11.16	11.1	11.08	11.14	11.48	10.36	10.2
SEL	RAC875_c19099_434	5B	92.31	3.97	0.22	0.09	11.05	11.12	11.05	11.1	11.02	11.27	10.49	10.3
SH	RAC875_c1673_193	4D	32.24	9.52	0.38	0	49.88	51.63	49.92	50.13	49.98	63.89	60.82	47.34
TEW	RAC875_c6922_291	4D	26.97	4.02	0.085	0.08	41.14	40.41	40.93	40.15	40.66	46.71	43.58	39.47
TIL	RAC875_c1673_193	4D	32.24	7.09	0.31	0	27.78	28.2	27.8	27.59	27.64	32.11	32.4	26.25

cM is the marker position. −logP is −log10 at the QTL peak; h² is the fraction of variance accounted for by QTL, after removing covariates Chr is the chromosome. P-value is the genome wise P-value for the QTL based on permutations. Alchemy, Brompton, Claire, Hereward, Rialto, Robigus, Soissons.

FIGURE 7 | QTLs markers for the trait GS39. Dashed lines indicate threshold (FDR < 0.05, logP4).

FIGURE 8 | QTLs markers for the trait GS55. Dashed lines indicate threshold (FDR < 0.05, logP4).
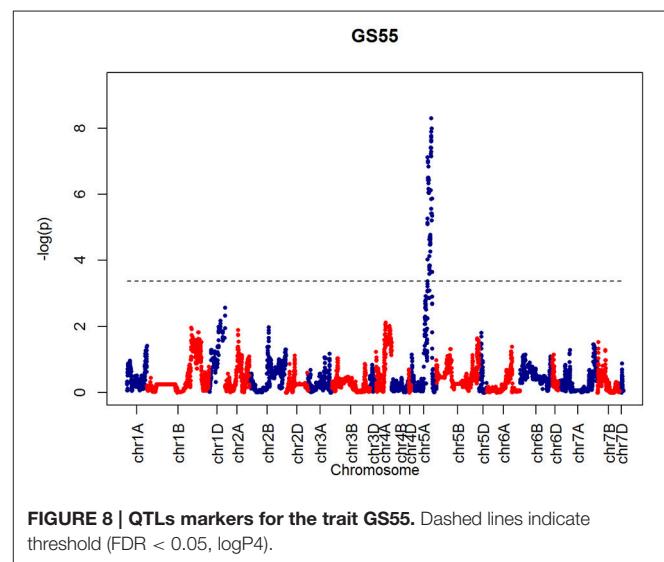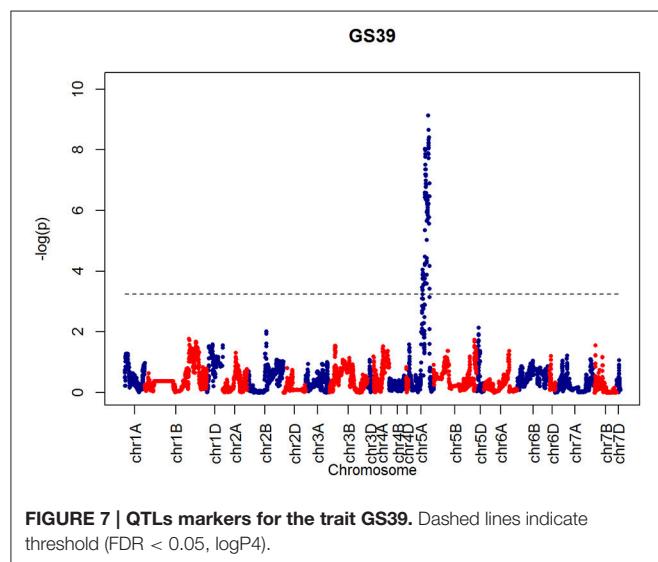
shows that Soissons has the highest contribution (0.6) to that particular marker.

DISCUSSION

This study screened a core set of lines derived from the NIAB wheat MAGIC population under Smarthouse conditions as strategy to understand the physical and genetic relationship between different phenological traits. This strategy resulted in the detection of QTL interactions with novel traits suggesting that the methodology should be taken further in the future.

Pair-wise correlation between all traits identified high positive (≥0.69) correlations between GS39, GS55, GS65 and senescence at the flag leaf (FLS), and between TEW and PW. There was also a negative correlation between FLS and length of the first flag leaf (FFLL), indicating that the shorter the flag leaf the more delayed was the start of senescence. Short flag leaves provide less nutrient assimilation therefore the plant has to compensate by either living longer, or producing a large number of tillers. Consistent with this idea, PC analysis also indicated a correlation between these traits. Suggestions that delayed leaf senescence leads to increased yield have been thrown into doubt (Borrill et al., 2015) but it may contribute under certain conditions. It will be interesting to see whether the correlation between leaf length and senescence is maintained under other environmental conditions. Genetically unstructured populations such as the MAGIC collection will be ideal to test whether experimental manipulation of the
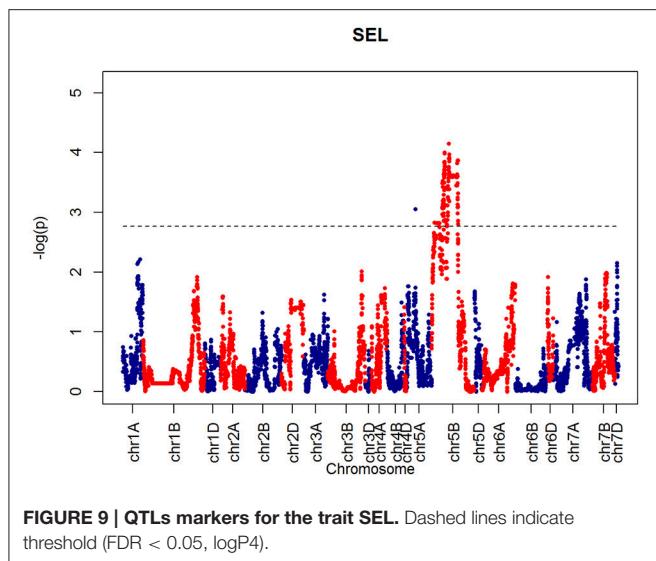
FIGURE 9 | QTLs markers for the trait SEL. Dashed lines indicate threshold (FDR < 0.05, logP4).
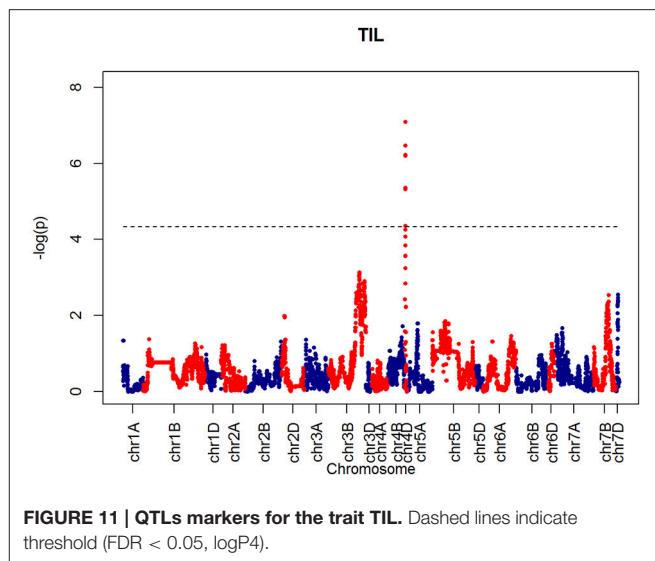
FIGURE 11 | QTLs markers for the trait TIL. Dashed lines indicate threshold (FDR < 0.05, logP4).
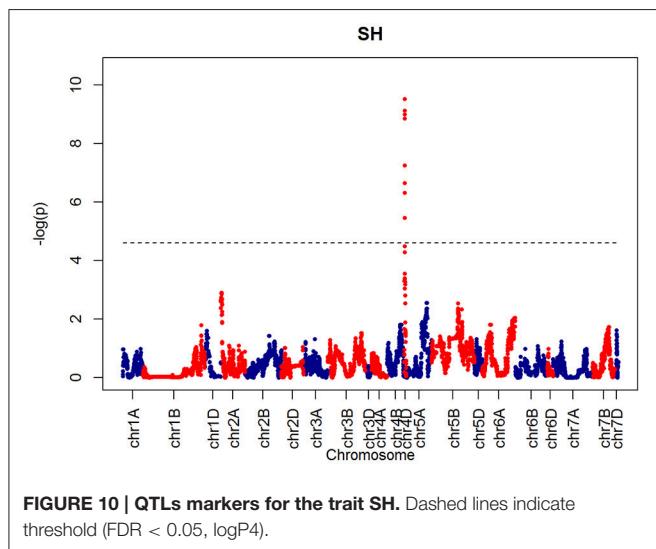
FIGURE 10 | QTLs markers for the trait SH. Dashed lines indicate threshold (FDR < 0.05, logP4).

size of the sink (grain mass) can further modulate flag leaf senescence.

In order to see if there was any clustering of individuals according to phenotype, a projection plot of the MAGIC lines onto the first two PCs was generated (Figure S6). The map shows lines grouped around 3 clusters of traits where Group 1 contained Cadenza, Zircon, Xi-19, Claire, Alchemy, Soissons, and Robigus; Group 2 contained Brompton and Hereward and Group 3 contained Avalon, showed similar trait profiles. Group 1 was early to senescence and group 2 later. Group 3 contained smaller plants as indicated by their opposite location to the FEL, SEL, and TEL vectors.

FLS was also negatively correlated with disease resistance (SM), indicating that highly susceptible plants were more likely to trigger senescence early. Support for this perspective may come from the observation that Xi-19 had the earliest FLS of all the parents and controls (Figure S9) and the joint highest disease score. In the field, Xi-19 flowers considerably later than

the earliest flowering parent, Soissons, which carries the Ppd-$D1a$ allele for early flowering (Scarth et al., 1985).

After confirming absence of population structure with PCA of the kinship matrix, QTL mapping identified a very strong marker ($-$logP > 8.00) on chromosome 5A associated with GS39 and GS55, which we believe is likely to correspond to the vernalization gene VRN-$A1$. This gene plays an important role in the vernalization process in diploid (_Triticum monococcum_) and polyploid wheat (_Triticum aestivum_) (Loukoianov et al., 2005; Kiss et al., 2014). However, using HAPPY, no significant QTL was found on chromosome 2D, the location of the Ppd-$D1$ locus, for GS39 or GS55. In the field, the presence of the Ppd-$D1$ allele in Soissons results in this line flowering 7–14 days earlier than the other MAGIC founder lines. These contrasting results between field and CER for Ppd-$D1$ and Vrn-$A1$ associated QTL suggest that the plants in this experiment might have experienced reduced vernalization as a result of a lack of cold treatment. However, this idea was discarded because inspection of CER records did not show any temperature discrepancies during vernalization. Another possibility is that plants displayed disease-like symptoms at an early stage. We also noticed that Cadenza, one of the 4 elites and a genotype that does not need vernalization, was one of the first to senesce, had the shortest duration between GS39 and GS55 and was more disease susceptible than the similarly early-flowering Soissons (**Figures 4**, **5**). The fact that Cadenza has no vernalization requirement, might suggest that indeed plants we not fully vernalized or encountered a de-vernalizing effect. Whatever the cause, the results of the experiment appear to have been strongly affected by a vernalization issue. This may explain some of the "anomalous" behavior of Cadenza, Xi-19 and Zircon, all of which do not require vernalization.

Further insight is provided by QTL validation analyses using mpMap. With no covariates included in the QTL model, the mpMap interval mapping approach produces very similar results to HAPPY. However, many more QTLs ($-$logP > 10) are detected using a model with 10 covariates in mpMap, as can

be seen in Figure S11. For GS39 and GS55, it can be seen that although the 5A QTL is still the highest peak, the *Ppd-D1* marker is significant and detected as the third (GS39) or 2nd (GS55) highest QTL. For FLS, 5A is also the most prominent QTL, but there is no evidence for a QTL around the *Ppd-D1* locus. This supports the observations from the trait analysis that progression through different developmental processes e.g., flowering vs. senescence, is controlled independently. Furthermore, QTL detected using mpMap for the length of the interval (d3) from flowering to senescence (Figure S11C) show a distinct pattern from both GS55 and FLS, although some loci are in common (e.g., 4D). As well as phenological traits, markers associated to morphological traits were also identified. For example, a strong marker on chromosome 4D was associated with shoot height (SH) and shoot number (TIL), as well as ear weight (OEW) and above ground biomass (PW). In all these cases, the QTL interval includes the *Rht-D1* (*Rht2*) locus. The *Rht-D1b* allele at this locus causes a semi-dwarfing phenotype in wheat, is strongly correlated with a reduction in height and several other morphological traits and is segregating in the MAGIC population (dwarfing alleles are present in all parent lines except Robigus and Soissons). Interestingly, this locus also shows up in the highest QTL interval for the d3 developmental interval in the mpMap covariate analysis (Figure S11D), suggesting that there may be independent effect on timing of progression through development. The *Rht-D1b* allele has a premature stop codon resulting in reduced sensitivity to gibberellic acid, which has been associated with reduced plant height and earlier heading date (Wilhelm et al., 2013). Our analysis indicates there are differential effects on the duration of other developmental processes not directly related to height or flowering *per se*.

In addition to FLS, GS55, and GS39, the multiple covariate analysis also identified a strong ($-$logP $>$ 18) peak on chromosome 7B (Figure S11E). Chromosome 7 has been previously associated to Septoria leaf blotch in an analysis of wheat-barley disomic addition lines. The highest level of resistance to infection by *S. tritici* was found in the *H. vulgare* chromosome addition line 7 followed by 4 and 6 (Rubiales et al., 2001).

Another interesting result is related to HI, for which a weak peak QTL was identified on the 2D chromosome, the location of *Ppd-D1*. In the field, the presence of the *Ppd-D1* allele in Soissons results in flowering 7–14 days earlier than the other MAGIC founder lines. In our analysis, Soissons have

some of the highest HIs and the shortest time to senesce but it also has the highest contribution (0.6) to that particular maker.

This study provides the first systematic phenological characterization of a wheat MAGIC population under controlled environment conditions. With careful developmental staging and end of life measurements of the MAGIC core set we were able to identify previously detected QTL loci on chromosomes 5A and 4D associated with the onset of senescence at the flag leaf. This powerful multi-founder population captures much of the genetic variation present in elite cultivars and a more detailed knowledge of fine-scale developmental and physiological patterns can be exploited for fine-tuning wheat's response to the environment. We have shown that the combination of MAGIC + Smarthouse can help extend the current understanding of developmental plasticity in elite wheat varieties with potential application for responding to the adaptation challenges facing agriculture in a changing climate.

AUTHOR CONTRIBUTIONS

Conceived and designed the study: AB and AC; analyzed the data: AC; assisted with QTL analysis: RM, KG, and IM; provided genetic data: KG and IM; provided scoring data: FC; wrote the paper: AC; provided comments and corrected the manuscript: All authors.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01540

REFERENCES

Bentley, A., Jensen, E., Mackay, I., Hönicka, H., Fladung, M., Hori, K., et al. (2013). *Genomics and Breeding for Climate-Resilient Crops, Vol 2*. Berlin; Heidelberg: Springer-Verlag.

Borrill, P., Fahy, B., Smith, A. M., and Uauy, C. (2015). Wheat grain filling is limited by grain filling capacity rather than the duration of flag leaf photosynthesis: a case study using NAM RNAi plants. *PLoS ONE* 10:e0134947. doi: 10.1371/journal.pone.0134947

Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: http://www.R-project.org/

Distelfeld, A., Avni, R., and Fischer, A. M. (2014). Senescence, nutrient remobilization, and yield in wheat and barley. *J. Exp. Bot.* 65, 3783–3798. doi: 10.1093/jxb/ert477

Ellis, M., Spielmeyer, W., Gale, K., Rebetzke, G., and Richards, R. (2002). "Perfect" markers for the Rht-B1b and Rht-D1b dwarfing genes in wheat. *Theor. Appl. Genet.* 105, 1038–1042. doi: 10.1007/s00122-002-1048-4

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024

Flintham, J. E., Börner, A., Worland, A. J., and Gale, M. D. (1997). Optimizing wheat grain yield: effects of Rht (gibberellin-insensitive) dwarfing genes. *J. Agric. Sci.* 128, 11–25. doi: 10.1017/S0021859696003942

Gan, S.-S. (2014). Leaf senescence as an important target for improving crop production. *Adv. Crop Sci. Tech.* 2:e116. doi: 10.4172/2329-8863.1000e116

Gardner, K. A., Wittern, L. M., and Mackay, I. J. (2016). A highly recombined, high-density, eight-founder wheat MAGIC map reveals extensive segregation distortion and genomic locations of introgression segments. *Plant Biotechnol. J.* 14, 1406–1417. doi: 10.1111/pbi.12504

Gregersen, P. L., Holm, P. B., and Krupinska, K. (2008). Leaf senescence and nutrient remobilisation in barley and wheat. *Plant Biol.* 10, 37–49. doi: 10.1111/j.1438-8677.2008.00114.x

Häffner, E., Konietzki, S., and Diederichsen, E. (2015). Keeping control: the role of senescence and development in plant pathogenesis and defense. *Plants* 4:449. doi: 10.3390/plants4030449

Huang, B. E., and George, A. W. (2011). R/mpMap: a computational platform for the genetic analysis of multiparent recombinant inbred lines. *Bioinformatics* 27, 727–729. doi: 10.1093/bioinformatics/btq719

Izawa, T. (2007). Adaptation of flowering-time by natural and artificial selection in Arabidopsis and rice. *J. Exp. Bot.* 58, 3091–3097. doi: 10.1093/jxb/erm159

Kempton, R. A. (1984). The use of biplots in interpreting variety by environment interactions. *J. Agric. Sci.* 103, 123–135. doi: 10.1017/s0021859600043392

Kiss, T., Balla, K., Veisz, O., Láng, L., Bedõ, Z., Griffiths, S., et al. (2014). Allele frequencies in the VRN-A1, VRN-B1 and VRN-D1 vernalisation response and PPD-B1 and PPD-D1 photoperiod sensitivity genes, and their effects on heading in a diverse set of wheat cultivars (*Triticum aestivum* L.). *Mol. Breed.* 34, 297–310. doi: 10.1007/s11032-014-0034-2

Liang, C., Jaiswal, P., Hebbard, C., Avraham, S., Buckler, E. S., Casstevens, T., et al. (2008). Gramene: a growing plant comparative genomics resource. *Nucleic Acids Res.* 36, D947–D953. doi: 10.1093/nar/gkm968

Loukoianov, A., Yan, L., Blechl, A., Sanchez, A., and Dubcovsky, J. (2005). Regulation of VRN-1 vernalisation genes in normal and transgenic polyploid wheat. *Plant Physiol.* 138, 2364–2373. doi: 10.1104/pp.105.064287

Mackay, I., Horwell, A., Garner, J., White, J., McKee, J., and Philpott, H. (2011). Reanalyses of the historical series of UK variety trials to quantify the contributions of genetic and environmental factors to trends and variability in yield over time. *Theor. Appl. Genet.* 122, 225–238. doi: 10.1007/s00122-010-1438-y

Mackay, I. J., Bansept-Basler, P., Barber, T., Bentley, A. R., Cockram, J., Gosman, N., et al. (2014). An eight-parent multiparent advanced generation inter-cross population for winter-sown wheat: creation, properties, and validation. *G3* 4, 1603–1610. doi: 10.1534/g3.114.012963

Mengiste, T. (2012). Plant immunity to necrotrophs. *Ann. Rev. Phytopathol.* 50, 267–294. doi: 10.1146/annurev-phyto-081211-172955

Mott, R., Talbot, C. J., Turri, M. G., Collins, A. C., and Flint, J. (2000). A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc. Natl. Acad. Sci. U.S.A.* 97, 12649–12654. doi: 10.1073/pnas.230304397

Reynolds, M., Foulkes, M. J., Slafer, G. A., Berry, P., Parry, M. A. J., Snape, J. W., et al. (2009). Raising yield potential in wheat. *J. Exp. Bot.* 60, 1899–1918. doi: 10.1093/jxb/erp016

Rubiales, D., Moral, A., and Martín, A. (2001). Chromosome location of resistance to septoria leaf blotch and common bunt in wheat-barley addition lines. *Euphytica* 122, 369–372. doi: 10.1023/A:1012952819255

Scarth, R., Kirby, E. J. M., and Law, C. N. (1985). Effects of the photoperiod genes Ppd1 and Ppd2 on growth and development of the shoot apex in wheat. *Ann. Bot.* 55, 351–359.

Schafer, J. L., and Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychol. Methods* 7, 147–177. doi: 10.1037/1082-989X.7.2.147

Shrestha, R., Matteis, L., Skofic, M., Portugal, A., McLaren, G., Hyman, G., et al. (2012). Bridging the phenotypic and genetic data useful for integrated breeding through a data annotation using the Crop Ontology developed by the crop communities of practice. *Front. Physiol.* 3:326. doi: 10.3389/fphys.2012.00326

Simpson, R. J., Lambers, H., and Dalling, M. J. (1983). Nitrogen redistribution during grain growth in wheat (*Triticum aestivum* L.): IV. Development of a quantitative model of the translocation of nitrogen to the grain. *Plant Physiol.* 71, 7–14. doi: 10.1104/pp.71.1.7

Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9440–9445. doi: 10.1073/pnas.1530509100

Thomas, H., and Howarth, C. J. (2000). Five ways to stay green. *J. Exp. Bot.* 51, 329–337. doi: 10.1093/jexbot/51.suppl_1.329

Valluru, R., Reynolds, M. P., and Lafarge, T. (2015). Food security through translational biology between wheat and rice. *Food Energy Security* 4, 203–218. doi: 10.1002/fes3.71

van Buuren, S., and Groothuis-Oudshoorn, K. (2011). mice: multivariate imputation by chained equations in R. *J. Stat. Softw.* 45:67.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

Wilhelm, E. P., Boulton, M. I., Al-Kaff, N., Balfourier, F., Bordes, J., Greenland, A. J., et al. (2013). Rht-1 and Ppd-D1 associations with height, GA sensitivity, and days to heading in a worldwide bread wheat collection. *Theor. Appl. Genet.* 126, 2233–2243. doi: 10.1007/s00122-013-2130-9

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zhang, Z. (2016). Multiple imputation with multivariate imputation by chained equation (MICE) package. *Ann. Trans. Med.* 4:30.

Four Tomato FLOWERING LOCUS T-Like Proteins Act Antagonistically to Regulate Floral Initiation

Kai Cao[1,2], Lirong Cui[1], Xiaoting Zhou[1], Lin Ye[1], Zhirong Zou[1]* and Shulin Deng[2]*

[1] State Key Laboratory of Crop Stress Biology for Arid Areas, Horticulture College, Northwest A&F University, Yangling, China,
[2] Laboratory of Plant Molecular Biology, Rockefeller University, New York, NY, USA

The transition from vegetative growth to floral meristems in higher plants is regulated through the integration of internal cues and environmental signals. We were interested to examine the molecular mechanism of flowering in the day-neutral plant tomato (*Solanum lycopersicum* L.) and the effect of environmental conditions on tomato flowering. Analysis of the tomato genome uncovered 13 PEBP (phosphatidylethanolamine-binding protein) genes, and found six of them were *FT*-like genes which named as *SlSP3D*, *SlSP6A*, *SlSP5G*, *SlSP5G1*, *SlSP5G2,* and *SlSP5G3*. Six FT-like genes were analyzed to clarify their functional roles in flowering using transgenic and expression analyses. We found that SlSP5G, SlSP5G2, and SlSP5G3 proteins were floral inhibitors whereas only SlSP3D/SFT (*SINGLE FLOWER TRUSS*) was a floral inducer. *SlSP5G* was expressed at higher levels in long day (LD) conditions compared to short day (SD) conditions while *SlSP5G2* and *SlSP5G3* showed the opposite expression patterns. The silencing of *SlSP5G* by VIGS (Virus induced gene silencing) resulted in tomato plants that flowered early under LD conditions and the silencing of *SlSP5G2* and *SlSP5G3* led to early flowering under SD conditions. The higher expression levels of *SlSP5G* under LD conditions were not seen in *phyB1* mutants, and the expression levels of *SlSP5G2* and *SlSP5G3* were increased in *phyB1* mutants under both SD and LD conditions compared to wild type plants. These data suggest that *SlSP5G*, *SlSP5G2,* and *SlSP5G3* are controlled by photoperiod, and the different expression patterns of *FT*-like genes under different photoperiod may contribute to tomato being a day neutral plant. In addition, PHYB1 mediate the expression of *SlSP5G*, *SlSP5G2,* and *SlSP5G3* to regulate flowering in tomato.

Keywords: tomato, floral repressor, floral activator, PEBP protein, FT-like genes, phytochromes

INTRODUCTION

In flowering plant, the timing of the transition from vegetative to reproductive phase is a major event in the plant life cycle. Both physiological and genetic studies have revealed the complexity of mechanisms that tightly control switch from vegetative to reproductive growth in the apical meristem (Bernier et al., 1993; Shalit et al., 2009). The phosphatidylethanolamine-binding proteins (PEBPs), found in both angiosperms and gymnosperms, have evolved to become both activators and repressors of flowering and they can be classified into three clades (Gyllenstrand et al., 2007; Karlgren et al., 2011). An example of this functional diversification is seen in the six PEBP family members of Arabidopsis. FLOWERING LOCUS T (FT) and TWIN SISTER OF FT (TSF), which

belong to the FT-like clade, function as flowering activators, TERMINAL FLOWER 1 (TFL1), BROTHER OF FT AND TFL1 (BFT), and ARABIDOPSIS THALIANA CENTRORADIALIS (ATC), which classify to the TFL1-like clade, are usually flowering repressors, and MOTHER OF FT AND TFL1(MFT), which defines the MFT-like clade, is predominantly a floral promoter (Karlgren et al., 2011).

In Arabidopsis, a long-day plant, FT is expressed in leaf phloem companion cells. This protein which triggers floral development in the shoot apical meristem (SAM) under long day (LD) conditions is a major output of the photoperiod pathway and controls floral transition in response to the changes in day length (Kardailsky et al., 1999; Kobayashi et al., 1999). CONSTANS (CO) encodes a zinc finger protein and promotes flowering under LD conditions (Putterill et al., 1995). In LD conditions, FT is activated by CO (Samach et al., 2000), and the FT protein then interacts with a novel endoplasmic reticulum membrane protein called FT-INTERACTING PROTEIN 1 (FTIP1; Liu et al., 2012). Following the interaction FT is transported from the companion cells to the sieve elements and entered the SAM by mass flow, where it associates with the basic leucine zipper domain (bZIP) transcription factor FD to activate downstream targets such as SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1) and the floral meristem identity gene APETALA 1 (AP1; Abe et al., 2005; Wigge et al., 2005; Corbesier et al., 2007). Also a PEBP family protein TSF probably acts in a similar way to FT (Yamaguchi et al., 2005). In SD conditions, Arabidopsis flowering is controlled by a gibberellin pathway, which promotes flowering through the activation of the flower meristem identity gene LEAFY (LFY) with no involvement of any PEBP family proteins (Moon et al., 2003). In SD plant rice, Hd3a, a FT homolog promotes flowering under SD conditions (Komiya et al., 2008, 2009). In the day-neutral plant tomato, the homolog of FT, SlSP3D/SFT (SINGLE-FLOWER TRSS), has been shown to encode the mobile florigen signal and promote tomato flowering (Molinero-Rosales et al., 2004).

Although almost all FT-like proteins act as floral activators an antagonistically functional switch has occurred in *Beta vulgaris* (sugar beet) and *Nicotiana tabacum* (tobacco) because of gene duplication event(s) generating other paralog(s). In sugar beet, BvFT1 protein acts as an inhibitor in floral development whereas another FT-like protein BvFT2 works as a promoter (Pin et al., 2010). Substitutions of specific amino acids can convert BvFT1 to a floral inducer and BvFT2 into a floral repressor (Pin et al., 2010). In tobacco, four FT-like proteins, NtFT1, NtFT2, and NtFT3 proteins are floral inhibitors whereas only NtFT4 is a floral inducer (Harig et al., 2012). These data suggest that some FT-like proteins, which are evolutionarily more related to FT than to TFL1/CEN, have evolved into flowering repressors.

Phytochromes are primary photosensory receptors that perceive red and far-red light of higher plants. These photochromic proteins exist in two photo-interconvertible isomeric forms: the red light absorbing form and the far-red light absorbing form (Hughes and Lamparter, 1999). Arabidopsis has five phytochrome genes, *PHYA* to *PHYE*, which encode the apoproteins of PHYA to PHYE, respectively (Quail et al., 1995).

PHYB plays an inhibitory role in floral initiation in Arabidopsis; the *phyB* mutant flowered earlier than WT in both LD and SD conditions, but the early-flowering phenotype of the *phyB* mutant is more pronounced in SD than in LD conditions (Goto et al., 1991; Mockler et al., 1999). *phyB* mutations of the LD pea plant (Weller and Reid, 1993), SD plant sorghum (Childs et al., 1997), and rice (Izawa et al., 2002) showed early-flowering and decreased photoperiodic sensitivities. *PHYB* delays flowering by suppressing the expression of *FT* in Arabidopsis (Endo et al., 2005) and *Hd3a* in rice (Izawa et al., 2002). Tomato contains five phytochrome genes, named *PHYA, PHYB1, PHYB2, PHYE,* and *PHYF* (Hauser et al., 1997). The tomato *PHYB1* is mainly involved in the de-etiolation response of seedlings, unfolding of the hypocotyl hook, cotyledon expansion, hypocotyl elongation, and anthocyanin accumulation (Kerckhoffs et al., 1997; Weller et al., 2000). However, the function of phytochromes in tomato flowering have not yet been reported.

Tomato is a photoperiod-insensitive, perennial in its native habit. The flowering time of tomato is measured by the number of leaves in the initial segment, which is rather stable under various environmental conditions (Kinet, 1977). Here, we performed expression and transgenic studies to clarify the functional roles of four expressed *FT*-like genes in tomato. One of the *FT*-like genes has already been identified by Molinero-Rosales et al. (2004), whereas the other three genes have not been studied. Here, we demonstrate the functional differentiation between these genes in controlling flowering through overexpression in Arabidopsis and VIGS-mediated knocking down in tomato. Our data suggest that among four expressed *FT*-like proteins, three of them act as floral repressors and only one of them function as a floral promoter. We also showed the expression profiles of tomato *FT*-like genes under LD and SD conditions in tomato wild-type (WT) and *phy* mutants. The evolution of antagonistic *FT*-like paralogs may be a common strategy in *Solanaceous* plants to fine-tune floral development in response to internal and environmental cues.

MATERIALS AND METHODS

Plant Material and Growth Conditions

We used cv. MoneyMaker (*Solanum lycopersicum* L.) wild type (WT) as control in this study, and *phyA, phyB1, phyB2,* and *phyB1B2* mutants in the MoneyMaker background were provided by the Tomato Genetic Resource Center (Department of Vegetable Crops, University of California, Davis) and their TGR accession numbers were LA4356, LA4357, LA4358, and LA4364, respectively. Tomato seeds were soaked in 50% bleach for 30 min. After the treatment, seeds were rinsed thoroughly in running water, then sown directly on a germination paper and incubated at 25°C. After germination, seedings were sowed onto commercial substrate and grown in a growth chamber under LD (16 h of light/8 h of dark) conditions or SD (8 h of light/16 h of dark) at $300 \, \mu mol \, m^{-2} \, s^{-1}$ and 25°C (both day and night).

To study the spatial expression patterns of *FT*-like genes, we extracted total RNA from leaf, apex, stem, flower, and root tissues, pooled from three 7-week-old plants. For diurnal changes in the expression of *FT*-like genes, leaves were harvested every 4 h for

24 h (0, 4, 8, 12, 16, 20, and 24 h), pooled from 3 third leaves of 5-week old plants. To study the effect of photoperiod on the expression of these genes, 5-week old uniform plantlets were transferred from LD conditions to SD conditions and reversely. Three different leaves at the same level were harvested 1, 2, and 3 day after the transfer.

Phylogenetic Analysis

Tomato protein sequences of the PEBP family members were downloaded from https://solgenomics.net/, *Arabidopsis thaliana* PEBP family members were downloaded from https://www.arabidopsis.org/, tobacco FT-like proteins reported by Harig et al. (2012) were download from https://solgenomics.net/, and sugar beet FT-like proteins reported by Pin et al. (2010) were download from http://www.ncbi.nlm.nih.gov/. Protein sequences were aligned using the maximum-likelihood method implemented in ClustalW software (Thompson et al., 1994). An N-J tree was produced from the results of 1000 bootstrap replicates using the ClustalW program.

Gene Expression Studies

Total RNA was extracted using an RNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. cDNA synthesis was performed by using the SuperscriptIII First strand synthesis system (Invitrogen) following the manufacturer's instructions. Real-time PCR was performed using SYBR Premix Ex Taq (TAKARA) in a Biorad CFX96 realtime PCR system. *ACTIN* was used as an internal control. The primers used were listed in **Supplementary Table S1**. Real-time quantitative PCR was repeated with three biological replicates, and each sample was assayed in triplicate by PCR.

Plasmid Constructs and Plant Transformation

The ORFs of *SlSP3D*(*Solyc03g063100*), *SlSP5G*(*Solyc05g053850*), *SlSP5G2*(*Solyc11g008640*), *SlSP5G3*(*Solyc11g008650*) were amplified by PCR, cloned in *pENTR/3C* vector (Invitrogen) and then transferred into *pBCO-DC* by recombination (Jang et al., 2007) using LR Clonase enzyme (Invitrogen). The resultant plasmid was used to transform *A. thaliana* (Col-0) plants by the *Agrobacterium tumefaciens* strain GV3101-mediated floral dip method (cite the original ref as well; Zhang et al., 2006). Transformed plants were selected on 0.8% agar media containing Murashige and Skoog salts, 0.5 g/L MES, and 10 g/L Sucrose and containing 10 µg/L basta. Arabidopsis plants were grown in a growth chamber under LD conditions at a light intensity of 100 µmol m^{-2} s^{-1} at 20°C (day and night).

Virus-Induced Gene Silencing (VIGS) in Tomato Plants

pTRV1 (*pYL192*) and *pTRV2* (*pYL156*) vectors had been described in Liu et al. (2002). The *pYL170 TRV2* vector was derived by cloning a *PstI*-blunt-*DraIII* fragment of *pYL156* into *EcoRI*-blunt-*DraIII*-cut *pCAMBIA3301*. This vector was identical to *pYL156*, except for a plant selection marker. To generate *pTRV2-SlPDS*, *pTRV2-SlSP5G*, *pTRV2SlSP5G2,* and *pTRV2SlSP5G3*, a cDNA fragment was PCR amplified using a tomato ecotype MoneyMaker cDNA library and primers were described in **Supplementary Table S1**. The resulting PCR products were cloned into *EcoRI-BamH*I-cut *pTRV2* (*PYL170*).

One-week-old tomato seedlings were used for the VIGS assay, *pTRV1* and *pTRV2* or its derivatives were introduced into *A. tumefaciens* strain GV3101 and the Agrobacterial strains mixed. A 5-mL culture was grown for 16 h at 28°C in 50 mg/L gentamycin and 50 mg/L kanamycin. The next day, the culture was inoculated into 30 mL of Luria-Bertani medium containing antibiotics, 10 mM MES, and 20 mM acetosyringone. The culture was grown 16 h in a 28°C shaker (200 r.p.m). *A. tumefaciens* cells were harvested and resuspended in infiltration media (10 mM MgCl$_2$, 10 mM MES, and 200 mM acetosyringone), adjusted to an OD600 of 1.5, and left at room temperature for 3–4 h. Agroinfiltration was performed with a needleless 1-mL syringe into two tomato cotyledons (Velásquez et al., 2009).

RESULTS

Identification and Phylogenetic Classification of Tomato FT-Like Genes

To identify FT-like proteins encoded by the tomato genome, the amino acid sequence of Arabidopsis FT protein was used to perform a BLAST survey against the tomato whole-genome database (https://solgenomics.net/). A total of 13 predicted PEBP genes were identified and annotated. In a previous study, the plant PEBP family could be classified into three main clades, described as FT-like, TFL1-like, and MFT-like (Chardon and Damerval, 2005). To evaluate the evolutionary relationships among the tomato, tobacco, sugar beet, and Arabidopsis FT-like proteins, specific and combined phylogenetic analysis based on their amino acid sequence were performed. We created a maximum-likelihood tree from an alignment of the 13 tomato PEBP proteins, the Arabidopsis 6 PEBP proteins, the tobacco FT-like (NtFT1-NtFT4) proteins and the sugar beet BvFT1 and BvFT2 proteins. **Figure 1** shows that there were six *FT*-like genes, five *TFL1*-like genes, and two *MFT*-like genes in the tomato PEBP family. PEBP family proteins contained two key motifs which are a putative ligand-binding pocket and an external loop. Protein sequence alignment also revealed a change of an amino acid residue from Tyr in tomato FT-like proteins to His in TFL1-like at the entrance of the binding pocket (**Supplementary Figure S1**), This amino acid residue in part determines the functional difference between FT and TFL1 in Arabidopsis (Hanzawa et al., 2005). Another amino acid residue was changed from Gln in tomato FT-like proteins to Asp in TFL1-like at the external loop encoded by the fourth exon (**Supplementary Figure S1**). This was another critical residue for the functional difference between FT and TFL1 in Arabidopsis (Ahn et al., 2006). These results suggest that *SlSP3D, SlSP6A, SlSP5G, SlSP5G1, SlSP5G2,* and *SlSP5G3* are *FT*-like genes.

Nucleotide sequence comparisons between genomic and predicted CDS allowed the identification of the exon-intron structures of tomato PEBP genes. Tomato PEBP genes showed conserved genomic organization and the exons were placed in

FIGURE 1 | Phylogenetic tree of PEBP proteins from tomato, sugar beet, tobacco, and Arabidopsis. Sequences were aligned with ClustalX and the results are displayed graphically using TreeView. The tomato PEBP genes are grouped into three major clades: the FT-like, TFL1-like, and MFT-like clades. AtFT, *Arabidopsis thaliana* FLOWERING LOCUS T; AtTSF, *A. thaliana* TWIN SISTER OF FT; AtTFL1, *A. thaliana* TERMINAL FLOWER 1; AtATC, *A. thaliana* ARABIDOPSIS THALIANA CENTRORADIALIS; AtMFT, *A. thaliana* MOTHER OF FT AND TFL1; BvFT1-2, *Beta vulgaris* FLOWERING LOCUS T 1-2; NtFT1-4, *Nicotiana tabacum* FLOWERING LOCUS T 1-4.

identical positions relative to the amino acid sequence of the Arabidopsis PEBP genes family, except for *SlSP5G1* and *SlSP5G3* (**Supplementary Figure S2**). The length of exons was quite conserved compared among tomato FT-related genes themselves and with *Arabidopsis* FT-related genes, but the introns differed in length. For the FT-like genes exon-intron structures, *SlSP6A* and *SlSP5G1* were truncated by a premature stop codon in their last exon and there was only one 222 exon without intron for *SlSP5G3*. In sugar beet and tobacco, FT-like protein could be further divided into floral promoters and floral repressors.

Expression Pattern of FT-Like Genes in Different Organs Under LD and SD Conditions

To investigate the roles of the six tomato *FT*-like genes in flowering, we first monitored their expression levels in different

organs. We isolated total RNA from the leaf, cotyledon, apex, stem, flower, and root tissues of 7-week-old tomato plants growth under LD and SD conditions. We compared the expression levels of *SlSP3D, SlSP6A, SlSP5G, SlSP5G1, SlSP5G2,* and *SlSP5G3* with those of the housekeeping gene *ACTIN* by qRT-PCR. No expression was detected for *SlSP6A* and *SlSP5G1* in all tissues. Considering there are premature stop codons in their last exons (Carmel-Goren et al., 2003; Consortium, 2012), these two genes probably do not encode functional proteins and are in fact pseudogenes. *SlSP3D, SlSP5G, SlSP5G2,* and *SlSP5G3* are mainly expressed in leaf and cotyledon under both LD and SD conditions (**Figures 2A,B**). A much higher expression level of *SlSP5G* was observed under LD conditions compared to SD conditions and *SlSP5G2* and *SlSP5G3* displayed an opposite expression pattern (**Figures 2A,B**). Under SD conditions, the number of leaves on the tomato main stem at flowering was eight on average, while this number increased to nine under LD conditions (**Figure 5A**).
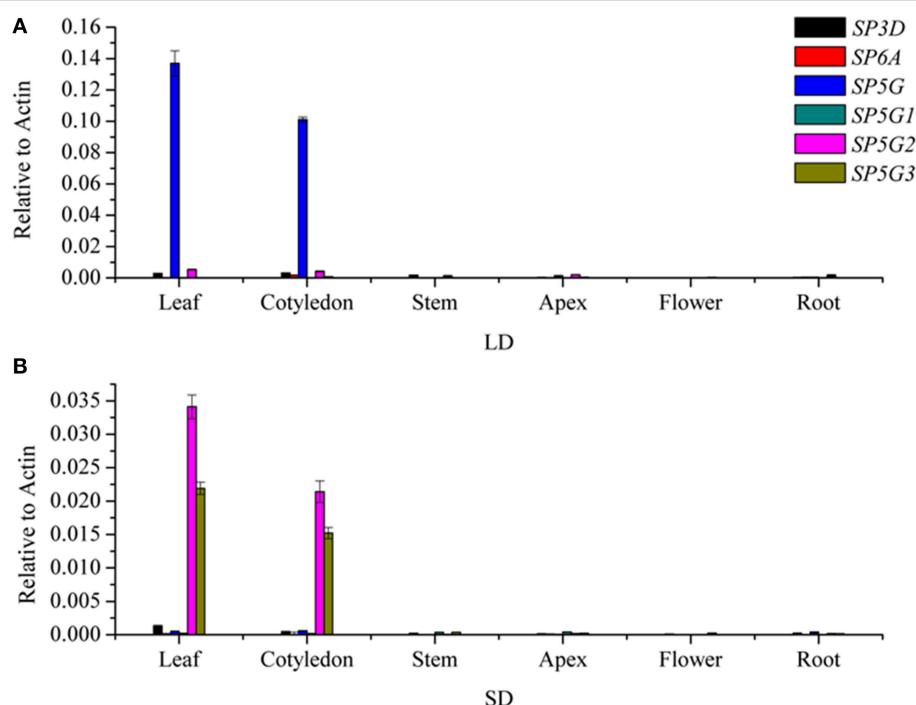
FIGURE 2 | Transcription analyses of tomato *FT*-like genes in different organs. *FT*-like genes expression analyzed by qRT-PCR using RNA extracted from leaf, cotyledon, stem, apex, flower, and root in 7-week old tomato plants grown under LD condition **(A)** and SD condition **(B)**. Expression of the tomato *ACTIN* gene was used as a reference. All data are showed as mean ±SE of three independent pools of extracts. Three technical replicates were performed for each extract.

Diurnal Rhythmic Expression Patterns of FT-Like Genes

To investigate the relationships among the four expressed FT-like genes, *SlSP3D, SlSP5G, SlSP5G2,* and *SlSP5G3*, we examined their diurnal expression patterns using the third leaves of 5-week-old seedlings. We performed qRT-PCR analyses using RNA from tomato plants grown in a LD diurnal cycle or a SD diurnal cycle. The expression of *SlSP3D* peaked at 4 h after dawn under LD conditions (**Figure 3A**) confirming previous results (Shalit et al., 2009). Our results also revealed that *SlSP5G* was transcribed at dawn and its expression peaked at the end of the day under LD conditions (**Figure 3B**). Under SD conditions, *SlSP5G* was constantly expressed at a lower level compared with its expression under LD conditions (**Figure 3B**). The expression pattern of *SlSP5G2* was different from that of *SlSP5G*, which showed a higher expression level under SD conditions, with expression peaking after 4 h of light under SD conditions (**Figure 3C**). On the other hand, *SlSP5G3* showed nearly the same diurnal oscillation pattern as *SlSP5G2*, and it peaked at 4 h after light under SD conditions (**Figure 3D**).

Tomato FT-Like Genes Have Antagonistic Functions in Floral Development in Transgenic Arabidopsis Plants

According to previous studies, *SlSP6A* and *SlSP5G1* were not expressed in tomato plants (Abelenda et al., 2014; **Figures 2A,B**) and consistent with this result we failed to clone *SlSP6A* and *SlSP5G1* from our tomato cDNA library. To investigate the functions of other *FT*-like genes in tomato flowering, *SlSP3D, SlSP5G, SlSP5G2,* and *SlSP5G3* were transferred into Arabidopsis plants under the control of a cauliflower mosaic virus (CaMV) *35S* promoter.

Overexpressing *SlSP3D* led to early flowering in transgenic Arabidopsis (**Figure 4B**) cofirming previous report that *SlSP3D* was a flowering promoter (Molinero-Rosales et al., 2004). Overexpression of *SlSP5G, SlSP5G2,* and *SlSP5G3* delayed flowering in transgenic Arabidopsis plants compared to wild-type controls (**Figures 4A,C–E**). The number of rosette leaves before flowering was seven in Col-0 under LD conditions. However, this number decreased to four in *SlSP3D* overexpressing plants (line 1), increased to 9.5 in *SlSP5G3* overexpressing plants (line 1), 12.5 in *SlSP5G* overexpressing plants (line 2) and 15.5 in *SlSP5G2* (line 2) overexpressing plants under LD conditions (**Figure 4F**). There were four overexpressing *SlSP3D, SlSP5G, SlSP5G2,* and *SlSP5G3* lines, respectively, and the number of rosette leaves before flowering in the other overexpressing *SlSP3D, SlSP5G, SlSP5G2,* and *SlSP5G3* lines were shown in **Supplementary Figure S3**. These results indicate that *SlSP3D* is a floral promoter, and *SlSP5G, SlSP5G2,* and *SlSP5G3* are floral repressors.

The Effect of Photoperiod on the Expression of SlSP3D, SlSP5G, SlSP5G2, and SlSP5G3 Genes

Figure 5B shows that *SlSP5G* expression increased under LD conditions, while *SlSP5G2* and *SlSP5G3* expression increased under SD conditions. These results suggested that *SlSP5G,*
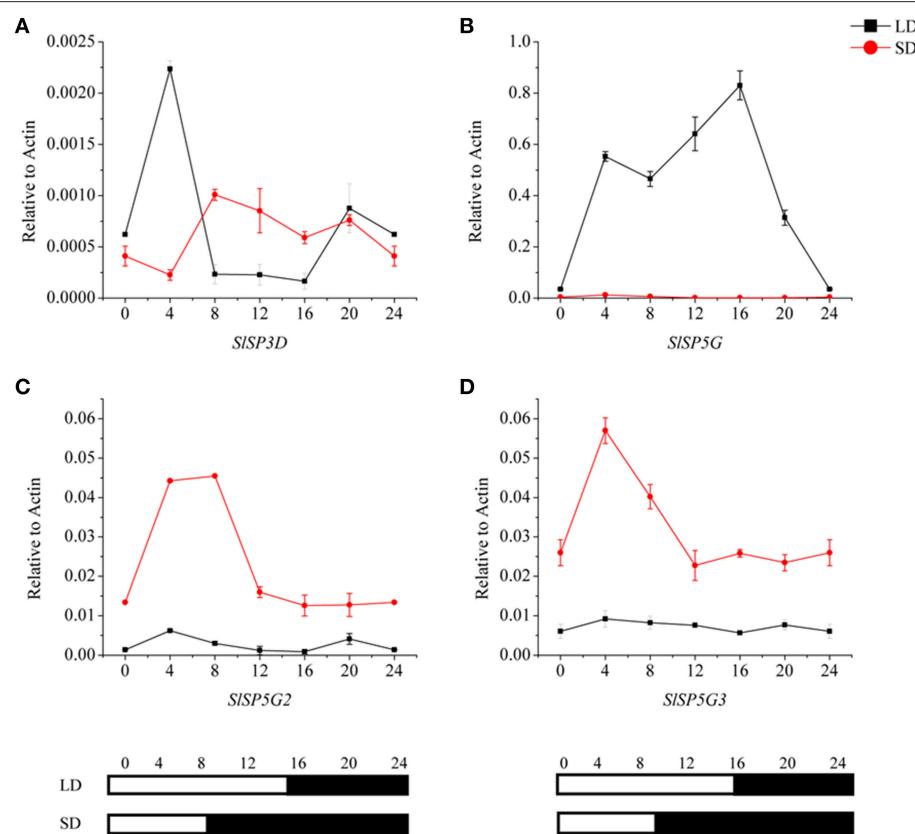
FIGURE 3 | Diurnal expression patterns of tomato *FT*-like genes, *SlSP3D* (A), *SlSP5G* (B), *SlSP5G2* (C), and *SlSP5G3* (D) under LD and SD conditions.
The black line represents LD condition and the red line represents SD condition. Leaves were harvested from plants at 4-h intervals throughout a light cycle. The
vertical axis shows relative mRNA levels of *FT*-like genes to *ACTIN* expression levels. All data are showed as mean ±SE of three independent pools of extracts. Three
technical replicates were performed for each extract. White and black bars at the bottom indicate light and dark periods, respectively.

SlSP5G2, and *SlSP5G3* were targets of photoperiodic regulation.
Therefore, we determined the expression levels of *SlSP3D,
SlSP5G, SlSP5G2,* and *SlSP5G3* in tomato plants grown under LD
conditions for 4 weeks and then transferred to SD conditions for
3 days, and vise-versa. There was no change of *SlSP3D* expression
when tomato plants were transferred from LD conditions to SD
conditions or from SD conditions to LD conditions (**Figure 5C**).
Downregulation of *SlSP5G* and upregulation of *SlSP5G2* and
SlSP5G3 were apparent after tomato plants were transferred from
LD conditions to SD conditions (**Figures 5D–F**). With tomato
plants transferred from SD conditions to LD conditions, we
found a directly increase of *SlSP5G* expression and a decrease of
SlSP5G2 and *SlSP5G3* expression after only one LD photoperiod
(**Figures 5D–F**). These results indicated that *SlSP5G, SlSP5G2,*
and *SlSP5G3* are directly regulated by day length.

Silencing of the Tomato SlSP5G, SlSP5G2, and SlSP5G3 Genes Using TRV-VIGS Vector

To study the function of *SlSP5G, SlSP5G2,* and *SlSP5G3* in
tomato flowering under LD and SD conditions, we constructed
a TRV-VIGS vector to suppress the expression of the endogenous

SlSP5G, SlSP5G2, and *SlSP5G3*. A mixture of Agrobacterium
cultures containing *pTRV1* and *pTRV2,* carrying tomato
SlSP5G (pTRV2-SlSP5G), SlSP5G2 (pTRV2-SlSP5G2), or *SlSP5G3
(pTRV2-SlSP5G3),* were infiltrated into the cotyledon of 1-week-
old tomato plants. We also used TRV-VIGS vector to suppress
the expression of the endogenous phytoene desaturase gene
(PDS) as a control. Tomato plants infected with *pTRV-SlPDS*
developed a photo-bleached phenotype in the upper leaves 10
days post-agro-infiltration (**Supplementary Figure S4**). Under
LD conditions, the number of leaves on tomato main stem
upon flowering was nine on average. However, this number
was reduced to seven when the tomato plants were infected
with *pTRV1/pTRV2-SlSP5G* (**Figure 6A**). Sixteen out of twenty
tomato plants showed early flowering after infiltration with
pTRV1/pTRV2-SlSP5G compared with tomato plants infiltrated
with *pTRV1/pTRV2.* We also extracted RNA from leaves of early
flowering tomato plants to confirm that *SlSP5G* was indeed
silenced by qRT-PCR. The primers that anneal to the *SlSP5G*
gene outside the region targeted for silencing were used. In early
flowering tomato plants infiltrated with *pTRV2-SlSP5G, SlSP5G*
expression was reduced significantly compared with the *TRV*
infected controls (**Figure 6B**). The results suggest that SlSP5G is
a flowering repressor under LD conditions.
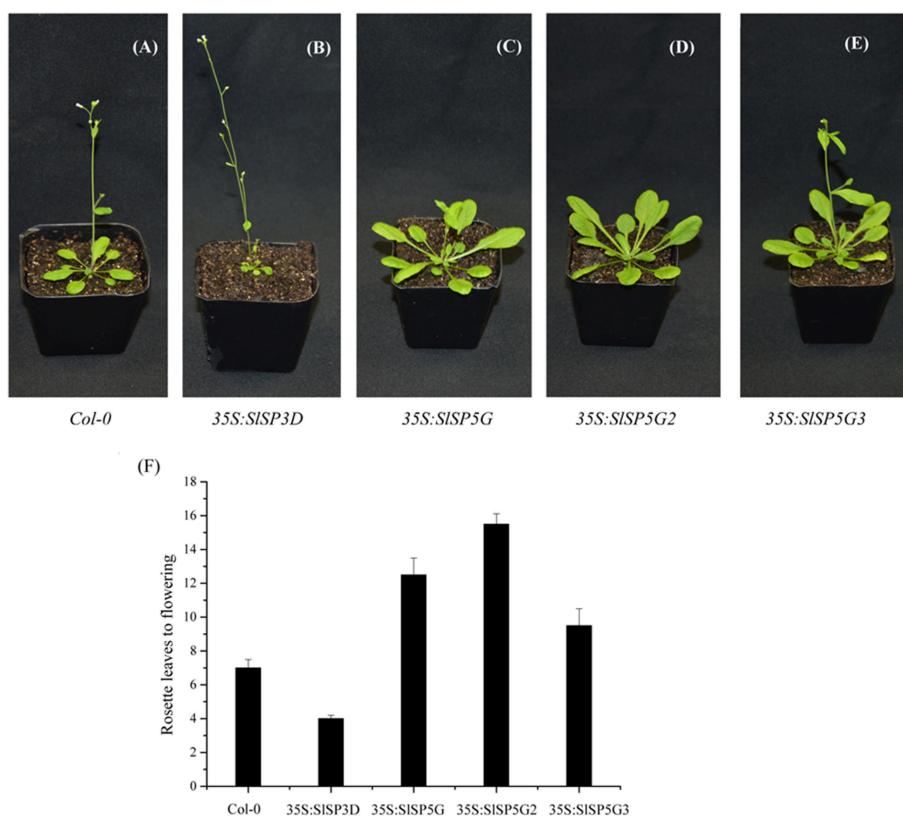
FIGURE 4 | Overexpression of *SlSP3D* **in transgenic Arabidopsis promotes flowering, but overexpression of** *SlSP5G, SlSP5G2,* **and** *SlSP5G3* **in transgenic Arabidopsis causes a delay in flowering. (A–E)** Wild type Col-0 **(A)** or representative transgenic Arabidopsis line overexpressing *SlSP3D* **(B)**, *SlSP5G* **(C)**, *SlSP5G2* **(D)**, and *SlSP5G3* **(E)** were grown under LD conditions. **(F)** The rosette leaves produced from SAM prior to flowering under LD conditions. The rosette leaves data are showed as mean ±SE of eight plants in each overexpression line. qRT-PCR data are showed as mean ±SE of three independent pools of extracts. Three technical replicates were performed for each extract.

Under SD conditions, the number of leaves on tomato main stem at flowering was eight on average; however, this number was reduced to 6.5 in tomato plants infected with *pTRV1/pTRV2-SlSP5G2* and *pTRV1/pTRV2-SLSP5G3* (**Figure 6C**). Fourteen out of twenty tomato plants showed slight early flowering after infiltration with *pTRV1/pTRV2-SlSP5G2* and *pTRV1/pTRV2-SLSP5G3*. RT-PCR also confirmed the decreased expression of *SlSP5G2* and *SlSP5G3* in infiltrated tomato plants (**Figure 6D**). These data suggest that *SlSP5G2* and *SlSP5G3* are factors that control tomato flowering under SD conditions.

Effects of Phytochrome B1 on the Expression of SlSP3D, SlSP5G, SlSP5G2, and SlSP5G3 Genes

As phytochromes are very important photoreceptors mediating flowering both in LD plants and SD plants (Izawa et al., 2002; Endo et al., 2005). We examined whether these photoreceptors have an effect on tomato flowering. We determined the number of leaves at flowering and the expression of the four expressed *FT*-like genes, *SlSP3D, SlSP5G, SlSP5G2,* and

SlSP5G3 in seedlings of *phyA, phyB1, phyB2,* and *phyB1B2* tomato mutants. We found that the number of leaves at flowering in *phyA* and *phyB2* mutants was the same as that in WT under both LD and SD conditions. However, the number of leaves at flowering in *phyB1* and *phyB1B2* mutants was 6 and 6.5, respectively, under LD conditions and this number was 8.5 on average under SD conditions (**Figures 7A–E**). In *phyB1* and *phyB1B2* mutants, there were a constant low expression level of *SlSP5G* under both LD and SD conditions (**Figures 7H,J**). However, there was a stable high expression level of *SlSP5G* in WT, *phyA* and *phyB2* mutants under LD conditions. Under SD conditions, there was a higher expression of *SlSP5G2* and *SlSP5G3* mRNA in *phyB1* and *phyB1B2* mutants compared to WT (**Figures 7F,H,J**). No difference was detected between *phyA, phyB2* mutants and WT on the expression levels of *SlSP5G, SlSP5G2,* and *SlSP5G3* under both LD and SD conditions (**Figures 7F,G,I**). Together, these results clearly demonstrate that PHYB1 has significant influence on the expression of *SlSP5G, SlSP5G2,* and *SlSP5G3*, and on the flowering time of tomato plants under both LD and SD conditions.
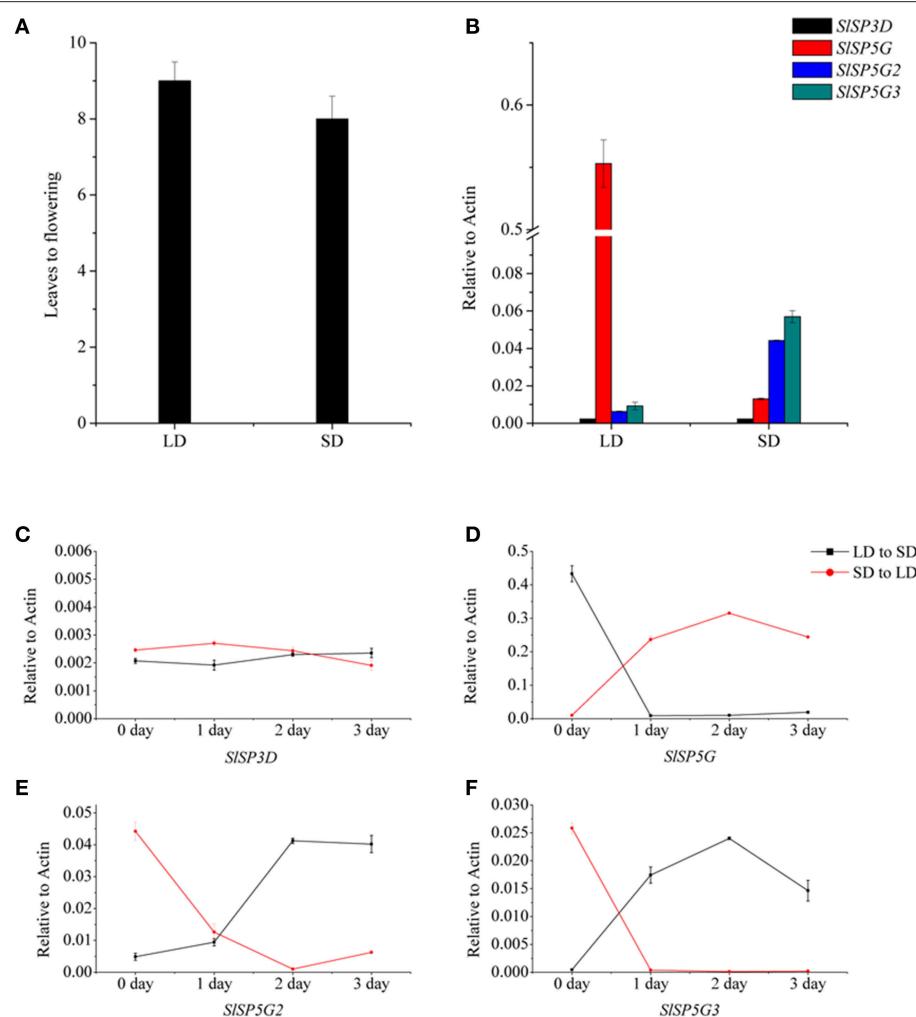
FIGURE 5 | The number of leaves at flowering of tomato plants under different photoperiod and transcription analysis of tomato *FT*-like genes by qRT-PCR in plants exposed to LD or SD conditions. (A) The number of leaves at flowering of tomato plants under LD or SD conditions. Data showed are mean ±SE (*n* = 10 plants). **(B)** The expression of *SlSP3D*, *SlSP5G*, *SlSP5G2*, and *SlSP5G3* under LD and SD conditions. **(C–F)** Expression patterns of *SlSP3D* **(C)**, *SlSP5G* **(D)**, *SlSP5G2* **(E)**, and *SlSP5G3* **(F)** of tomato plants transferred from SD to LD condition, and vise-versa. All data are expressed as means ±SE of three independent pools of extracts. Three technical replicates were performed for each extract.

DISCUSSION

FT-Like Genes Act Antagonistically to Regulate Floral Initiation in Tomato

Plant PEBP family proteins are divided into three major clades, with the FT-like and MFT-like clades primarily acting to promote and the TFL1-like clade primarily acting to repress floral development. In this study, we queried the complete tomato genome sequences and identified 13 PEBP genes, six of which belong to the FT-like clade, five are classified in the TFL-like clade and two are MFT-like clade. In the six FT-like clade, *SlSP3D, SlSP6A, SlSP5G, SlSP5G1, SlSP5G2,* and *SlSP5G3*, two, *SlSP6A* and *SlSP5G1* were not expressed in tomato plants. It has already been demonstrated that *SlSP3D/SFT*, the tomato ortholog of *FT*, induces flowering in day-neutral tomato and *sft* mutants show late flowering phenotype (Molinero-Rosales

et al., 2004; Lifschitz et al., 2006). Here, we show that transgenic Arabidopsis plants possessing *SlSP3D* displayed much earlier flowering phenotype compared to control plants. FT-like proteins that promote flowering have been identified in many species such as *Populus* spp. (poplar; Böhlenius et al., 2006), *Malus domestica* (apple; Hättasch et al., 2008), *B. vulgaris* (sugar beet; Pin et al., 2010), *Solanum tuberosum* (potato; Navarro et al., 2011), *N. tabacum* (tobacco; Harig et al., 2012), and *Oryza sativa* (rice; Kojima et al., 2002).

Based on phylogenetic data *SlSP5G, SlSP5G2,* and *SlSP5G3* have been postulated to be orthologous to FT-like genes (Abelenda et al., 2014). However, overexpression of *SlSP5G, SlSP5G-2,* or *SlSP5G-3* in Arabidopsis resulted in late flowering phenotype compared control plants. In sugar beet and tobacco, FT-like genes can act as flowering promoters and repressors. The two sugar beet *FT*-like genes, *BvFT1* and *BvFT2* differ in
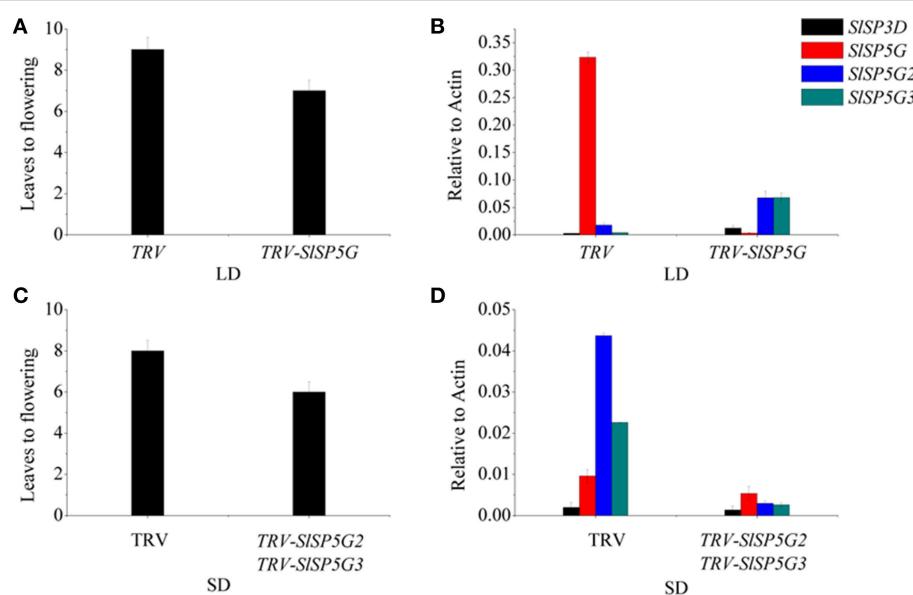
FIGURE 6 | Flowering time and transcript analysis of VIGS of Sl*SP5G* or Sl*SP5G2* and *SlSP5G3* tomato plants. (A) The number of leaves at flowering of tomato plants when infected by *TRV-SlSP5G* vectors under LD conditions. Data are mean ±SE (*n* = 10 plants). **(B)** The expression of *SlSP3D*, *SlSP5G*, *SlSP5G2,* and *SlSP5G3* of tomato plants infected by *TRV-SP5G* vectors under LD conditions. **(C)** The number of leaves at flowering of tomato plants when infected by *TRV-SlSP5G2* and *SlSP5G3* vectors under SD conditions. **(D)** The expression of *SlSP3D*, *SlSP5G*, *SlSP5G2,* and *SlSP5G3* of tomato plants infected by *TRV-SP5G2* and *SlSP5G3* vectors under SD conditions. All data are expressed as means ±SE of three independent pools of extracts. Three technical replicates were performed for each extract.

three amino acid residues within the critical region encoded by the fourth exon (**Supplementary Figure S1**). Try (134), Gly (137), and Trp (138) are the most important three amino acids of the external loop for BvFT2 protein. Substitution of these three amino acid residues in BvFT2 was sufficient to convert it into a repressor (Pin et al., 2010). The change of BvFT1 Asn (138) into Try, Gln (141) into Gly and Gln (142) into Trp could completely revert its repressing function to promoting function in flowering. Four FT-like proteins have been reported in tobacco. NtFT4 is a flowering activator and the amino acid residues at the three conserved positions matched those of Arabidopsis FT and BvFT2 whereas the corresponding positions in the floral repressors NtFT1-3 were not conserved (Harig et al., 2012). Through protein sequence alignment, we also found that in the tomato SlSP3D the amino acid residues at the three critical positions were Tyr (133), Gly (136), and Trp (137) and these matched to those of the other floral activators, such as FT in Arabidopsis, BvFT2 in sugar beet and NtFT4 in tobacco (**Supplementary Figure S1**). However, the amino acid residues of SlSP5G in these three conserved positions were the same as those found in the floral repressors like NtFT1-3 in tobacco. The corresponding positions of SlSP5G2 and SlSP5G3 in these positions were not conserved compared with other floral activators and repressors (**Supplementary Figure S1**). These results suggest that SlSP5G, SlSP5G2, and SlSP5G3 were initially promoters of flowering but these mutations within the external loop converted its function to flowering repression. The three amino acids are critical for the activator vs. repressor function.

The Expression Profiles of *FT*-Like Genes is Influenced by Photoperiod

The expression of *FT*-like genes in many species is regulated in a photoperiod-dependent manner (Samach et al., 2000; Kojima et al., 2002). Termination and flowering in cultivated tomato are not sensitive to day length, but flower initiation occurs earlier and inflorescence development far better in SD conditions than in LD conditions (Kinet, 1977). All four tomato *FT*-like genes were expressed exclusively in leaf tissue (**Figure 2**), which was the same as tobacco *FT*-like genes (Harig et al., 2012). In tobacco, *NtFT1, NtFT2,* and *NtFT4* showed higher expression levels under SD conditions than under LD conditions (Harig et al., 2012). In sugar beet, the floral repressor *BvFT1* was expressed at high levels when plants were grown in SD or in non-vernalized biennials plants that were not competent to flower (Pin et al., 2010). We also found that *SlSP5G* mRNA expression was up-regulated under LD conditions, while *SlSP5G2* and *SlSP5G3* mRNA increased under SD conditions (**Figure 5B**). The expression of *SlSP3D* was similar under both LD and SD conditions (**Figure 5B**). Although tomato is day-neutral with respect to flowering, the expression of the *SlSP5G, SlSP5G2,* and *SlSP5G3* identified here seem to be photoperiod dependent. *SlSP5G* most likely controls tomato flowering under LD conditions while *SlSP5G2* and *SlSP5G3* seem to regulate flowering under SD conditions. Tomato plants have an adaptive mechanism to adjust flowering according to photoperiod using a combination of different *FT*-like genes.

In this study, the silencing of *SlSP5G* by *TRV-VIGS* vector under LD conditions resulted in early flowering of tomato
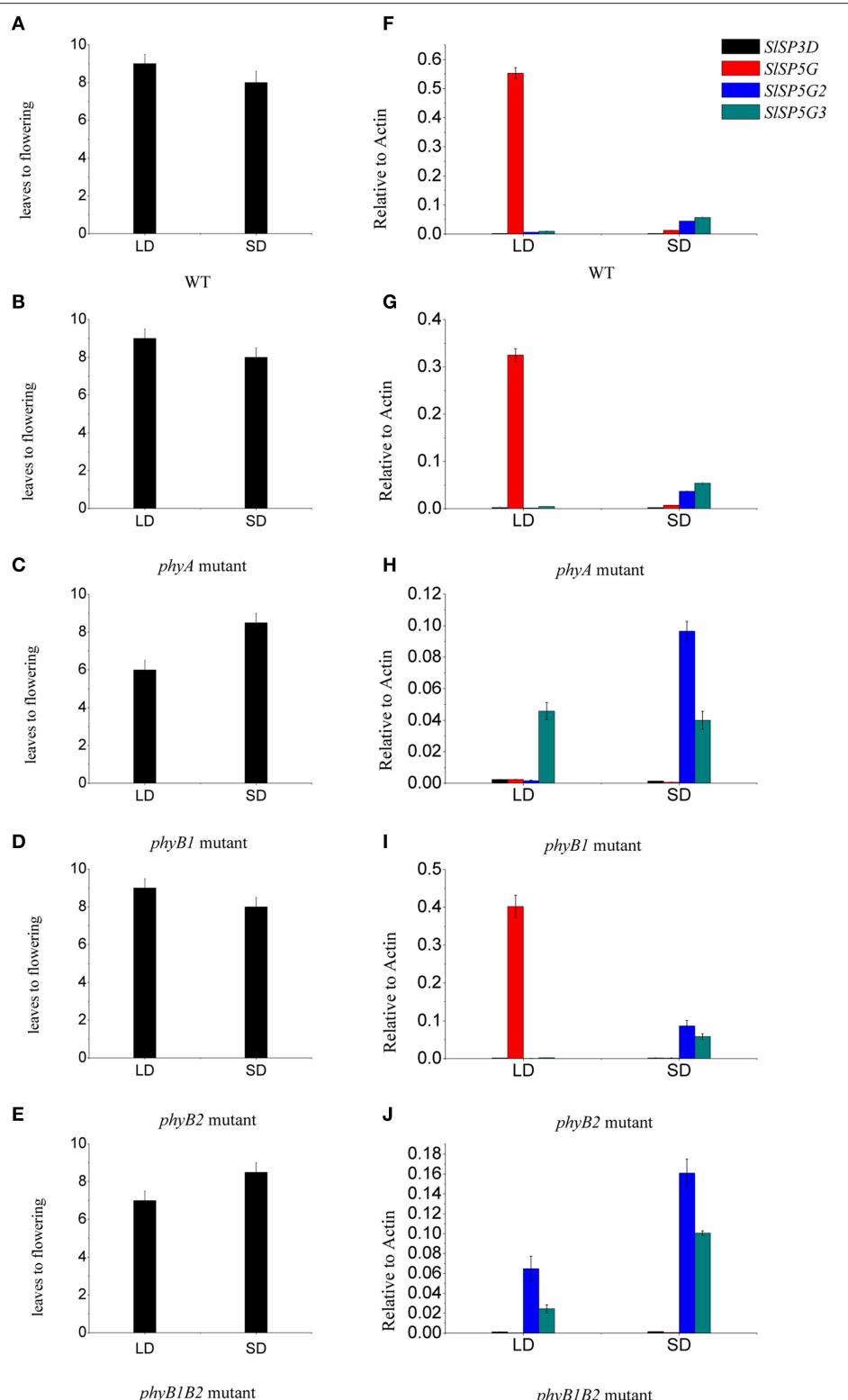
FIGURE 7 | Phytochrome B1 is responsible for the expression of *SlSP5G* and influences the flowering of tomato plants. (A–E) The number of leaves at flowering of tomato WT **(A)**, *phyA* **(B)**, *phyB1* **(C)**, *phyB2* **(D)**, and *phyB1B2* **(E)** mutants grown under LD and SD conditions. Data are mean ±SE (*n* = 10 plants). **(F–J)** Expression levels of *SlSP3D*, *SlSP5G*, *SlSP5G2*, and *SlSP5G3* of tomato WT **(F)**, *phyA* **(G)**, *phyB1* **(H)**, *phyB2* **(I)**, and *phyB1B2* **(J)** mutants that were grown in LD or SD conditions. Data are expressed as means ±SE of three independent pools of extracts. Three technical replicates were performed for each extract.

plants, and the silencing of *SlSP5G2* and *SlSP5G3* by *TRV-VIGS* vector under SD conditions led to early flowering of tomato plants. These results also showed that SlSP5G, SlSP5G2, and SlSP5G3 were floral repressors. SELF PRUNING (SP) is a homolog of TFL1-like gene and SP protein functions as an anti-terminator, maintaining vegetative growth (Pnueli et al., 1998). Mutant *sp* plants form progressively shorter sympodial units, until the shoots terminate in two successive inflorescences. In many species the ratio of floral activators and repressors, e.g., local ratios of SFT/SP3D (FT-like) and SP (TFL1-like), has been proposed to regulate local growth termination equilibria in all meristems of the tomato shoot system (Shalit et al., 2009; McGarry and Ayre, 2012). The three tomato FT-like floral repressors appear to have taken on the role usually played by TFL1 homologs in most other plants. Additional research is required to classify how FT-like floral activators and repressors and SP set the timing of the developmental switch from vegetative to reproductive growth. Both SFT/SP3D and SP of tomato bind to 14-3-3 and bZIP (SPGB, a homolog of FD) proteins in yeast, but each protein also has its own specific binding proteins (Pnueli et al., 2001). In Arabidopsis, FT protein is first transferred into the sieve elements and then subsequently transported by mass flow to the apex, where it interacts with FD to promote flowering (Abe et al., 2005; Wigge et al., 2005; Corbesier et al., 2007). In tomato, SlSP5G, SlSP5G2, and SlSP5G3 maybe like SlSP3D/SFT and they may interact with SPGB to control tomato flowering.

Phytochrome B1 Regulates FT-Like Genes

Phytochromes are photochromic proteins that regulate light responses under different light conditions (quantity, quality, and timing). Our data showed that, in the tomato *phyB1* mutant, the expression of *SlSP5G* under both LD and SD conditions was very low. The expression of *SlSP5G2* and *SlSP5G3* was always in a fairly high level, compared with WT under both LD and SD conditions. Based on the results we obtained in tomato plants, we found that the PHYB1 could promote the expression of *SlSP5G* under LD conditions but suppress the expression of *SlSP5G2* and *SlSP5G3* under both LD and SD conditions. It has been shown that PHYB has a general inhibitory effect on flowering in both LD plants and SD Plants (Lin, 2000; Yanovsky and Kay, 2003). An inhibitory effect of PHYB on FT expression has been shown in Arabidopsis (Valverde et al., 2004; Endo et al., 2005). In rice, the *phyB* mutation abolishes the night break effect on flowering and *Hd3a* mRNA, and PHYB suppresses the expression of *Hd3a* (Izawa et al., 2002; Ishikawa et al., 2005).

Phytochromes need to interact with the circadian clock to regulate flowering time in different day-lengths, but the molecular details of such interactions remains unclear (Valverde et al., 2004; Song et al., 2012). *phyB* mutations of the SD plant sorghum and the LD plant Arabidopsis both caused an early flowering phenotype; tomato *phyB*1 mutant also has an early flowering phenotype under LD conditions. One interpretation of this observation is that PHYB action may suppress floral initiation regardless of photoperiods, but the signal transduction or plant's responsiveness to PHYB signaling is gated by the action of the circadian clock, resulting in different day-length responses in the flowering time of different plants.
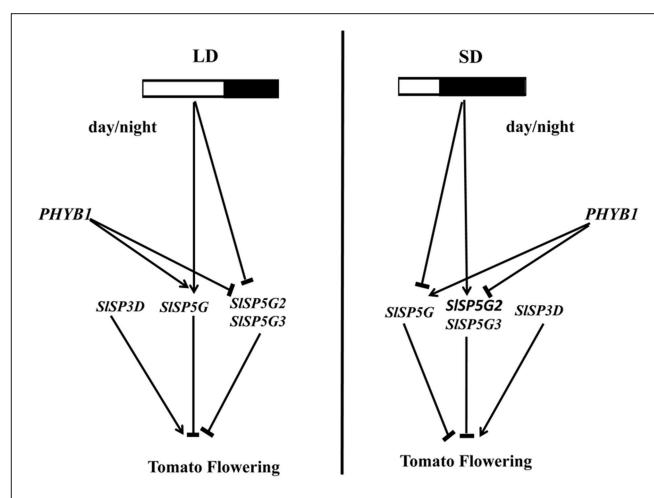
FIGURE 8 | Model of the photoperiod effect on flowering in tomato. The expression of FT-like genes was regulated by photoperiod and mediated by phytochrome B1. In LD conditions, the expression of *SlSP5G* was induced, and the expression of *SlSP5G2* and *SlSP5G3* were inhibited. In SD conditions, the expression of *SlSP5G* was inhibited, and the expression of *SlSP5G2* and *SlSP5G3* were induced. The different expression pattern of tomato FT-like genes under different photoperiod may contribute tomato being a day neutral plant. Phytochrome B1 could promote the expression of *SlSP5G,* and inhibit the expression of *SlSP5G2* and *SlSP5G3*.

Based on the results we obtained in this study, we propose a model to explain the photoperiod effect on tomato flowering (**Figure 8**). This model is consistent with all of the results we obtained in our studies and suggests that the expression pattern of these *FT*-like genes is regulated by photoperiod and mediated by *PHYB1*. In addition, four tomato *FT*-like genes reveal they act antagonistically to regulate floral initiation. Understanding the molecular mechanism of flowering in a day-neutral plant has important implications for agriculture. Further studies are required to integrate the knowledge obtained from model species like LD plant Arabidopsis and SD plant rice to provide further insight on the mechanisms regulating flowering in day-neutral plant.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: SD and ZZ. Performed the experiments: KC. Analyzed the data: LC and KC. Contributed reagents/materials/analysis tools: XZ. Amplify the seed: LY. Wrote the paper: KC.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.01213

Supplementary Table S1 | Sequences of primers used in this study for plasmid construction, quantitative RT-PCR and VIGS.

Supplementary Figure S1 | Multiple amino acid alignment of PEBP domain of Arabidopsis, sugar beet, tobacco, and tomato PEBP family proteins. The vertical arrowhead indicates the crucial amino acid change responsible for the difference between FT-like and TFL-like functions identified by Hanzawa et al. (2005). Amino acid residues conserved in FT-like proteins that

promote flowering are shaded in red, identified by Pin et al. (2010) and Harig et al. (2012).

Supplementary Figure S2 | The exon-intron structures of tomato PEBP genes resembles that of AtFT. Boxed areas depict the exons and lines represent introns. Numbers represent exon and intron lengths (bp).

Supplementary Figure S3 | The number of rosette leaves before flowering in the other _FT_-like genes overexpression lines. All data are showed as mean ±SE of eight plants in each overexpression lines.

Supplementary Figure S4 | Silencing of PDS control gene causes photobleaching in tomato plans. Photographs were taken 4 weeks after silencing. **(A)** Tomato plant infected by _TRV-SlPDS_ vectors. **(B)** Tomato plant infected by empty _TRV_ vectors.

REFERENCES

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., et al. (2005). FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. _Science_ 309, 1052–1056. doi: 10.1126/science.1115983

Abelenda, J. A., Navarro, C., and Prat, S. (2014). Flowering and tuberization: a tale of two nightshades. _Trends Plant Sci._ 19, 115–122. doi: 10.1016/j.tplants.2013.09.010

Ahn, J. H., Miller, D., Winter, V. J., Banfield, M. J., Lee, J. H., Yoo, S. Y., et al. (2006). A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. _EMBO J._ 25, 605–614. doi: 10.1038/sj.emboj.7600950

Böhlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A. M., Jansson, S., Strauss, S. H., et al. (2006). CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. _Science_ 312, 1040–1043. doi: 10.1126/science.1126038

Bernier, G., Havelange, A. E., Houssa, C., Petitjean, A., and Lejeune, P. (1993). Physiological signals that induce flowering. _Plant Cell_ 5, 1147.

Carmel-Goren, L., Liu, Y. S., Lifschitz, E., and Zamir, D. (2003). The SELF-PRUNING gene family in tomato. _Plant Mol. Biol._ 52, 1215–1222. doi: 10.1023/B:PLAN.0000004333.96451.11

Chardon, F., and Damerval, C. (2005). Phylogenomic analysis of the PEBP gene family in cereals. _J. Mol. Evol._ 61, 579–590. doi: 10.1007/s00239-004-0179-4

Childs, K. L., Miller, F. R., Cordonnier-Pratt, M.-M., Pratt, L. H., Morgan, P. W., and Mullet, J. E. (1997). The sorghum photoperiod sensitivity gene, Ma3, encodes a phytochrome B. _Plant Physiol._ 113, 611–619. doi: 10.1104/pp.113.2.611

Consortium, T. G. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. _Nature_ 485, 635–641. doi: 10.1038/nature11119

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., et al. (2007). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. _Science_ 316, 1030–1033. doi: 10.1126/science.1141752

Endo, M., Nakamura, S., Araki, T., Mochizuki, N., and Nagatani, A. (2005). Phytochrome B in the mesophyll delays flowering by suppressing FLOWERING LOCUS T expression in Arabidopsis vascular bundles. _Plant Cell_ 17, 1941–1952. doi: 10.1105/tpc.105.032342

Goto, N., Kumagai, T., and Koornneef, M. (1991). Flowering responses to light−breaks in photomorphogenic mutants of _Arabidopsis thaliana_, a long−day plant. _Physiol. Plant._ 83, 209–215. doi: 10.1111/j.1399-3054.1991.tb02144.x

Gyllenstrand, N., Clapham, D., Källman, T., and Lagercrantz, U. (2007). A Norway spruce FLOWERING LOCUS T homolog is implicated in control of growth rhythm in conifers. _Plant Physiol._ 144, 248–257. doi: 10.1104/pp.107.095802

Hanzawa, Y., Money, T., and Bradley, D. (2005). A single amino acid converts a repressor to an activator of flowering. _Proc. Natl. Acad. Sci. U.S.A._ 102, 7748–7753. doi: 10.1073/pnas.0500932102

Harig, L., Beinecke, F. A., Oltmanns, J., Muth, J., Müller, O., Rüping, B., et al. (2012). Proteins from the FLOWERING LOCUS T−like subclade of the PEBP family act antagonistically to regulate floral initiation in tobacco. _Plant J._ 72, 908–921. doi: 10.1111/j.1365-313X.2012.05125.x

Hättasch, C., Flachowsky, H., Kapturska, D., and Hanke, M.-V. (2008). Isolation of flowering genes and seasonal changes in their transcript levels related to

flower induction and initiation in apple (_Malus domestica_). _Tree Physiol._ 28, 1459–1466. doi: 10.1093/treephys/28.10.1459

Hauser, B. A., Pratt, L. H., and Cordonnier-Pratt, M.-M. (1997). Absolute quantification of five phytochrome transcripts in seedlings and mature plants of tomato (_Solanum lycopersicum_ L.). _Planta_ 201, 379–387. doi: 10.1007/s004250050080

Hughes, J., and Lamparter, T. (1999). Prokaryotes and phytochrome. The connection to chromophores and signaling. _Plant Physiol._ 121, 1059–1068. doi: 10.1104/pp.121.4.1059

Ishikawa, R., Tamaki, S., Yokoi, S., Inagaki, N., Shinomura, T., Takano, M., et al. (2005). Suppression of the floral activator Hd3a is the principal cause of the night break effect in rice. _Plant Cell_ 17, 3326–3336. doi: 10.1105/tpc.105.037028

Izawa, T., Oikawa, T., Sugiyama, N., Tanisaka, T., Yano, M., and Shimamoto, K. (2002). Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. _Genes Dev._ 16, 2006–2020. doi: 10.1101/gad.999202

Jang, I. C., Yang, S. W., Yang, J. Y., and Chua, N. H. (2007). Independent and interdependent functions of LAF1 and HFR1 in phytochrome A signaling. _Genes Dev._ 21, 2100–2111. doi: 10.1101/gad.1568207

Kardailsky, I., Shukla, V. K., Ahn, J. H., Dagenais, N., Christensen, S. K., Nguyen, J. T., et al. (1999). Activation tagging of the floral inducer FT. _Science_ 286, 1962–1965. doi: 10.1126/science.286.5446.1962

Karlgren, A., Gyllenstrand, N., Källman, T., Sundström, J. F., Moore, D., Lascoux, M., et al. (2011). Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. _Plant Physiol._ 156, 1967–1977. doi: 10.1104/pp.111.176206

Kerckhoffs, L., Schreuder, M., Tuinen, A. V., Koornneef, M., and Kendrick, R. (1997). Phytochrome control of anthocyanin biosynthesis in tomato seedlings: analysis using photomorphogenic mutants. _Photochem. Photobiol._ 65, 374–381. doi: 10.1111/j.1751-1097.1997.tb08573.x

Kinet, J. (1977). Effect of light conditions on the development of the inflorescence in tomato. _Sci. Hortic._ 6, 15–26. doi: 10.1016/0304-4238(77)90074-7

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M., and Araki, T. (1999). A pair of related genes with antagonistic roles in mediating flowering signals. _Science_ 286, 1960–1962. doi: 10.1126/science.286.5446.1960

Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araki, T., et al. (2002). Hd3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. _Plant Cell Physiol._ 43, 1096–1105. doi: 10.1093/pcp/pcf156

Komiya, R., Ikegami, A., Tamaki, S., Yokoi, S., and Shimamoto, K. (2008). Hd3a and RFT1 are essential for flowering in rice. _Development_ 135, 767–774. doi: 10.1242/dev.008631

Komiya, R., Yokoi, S., and Shimamoto, K. (2009). A gene network for long-day flowering activates RFT1 encoding a mobile flowering signal in rice. _Development_ 136, 3443–3450. doi: 10.1242/dev.040170

Lifschitz, E., Eviatar, T., Rozman, A., Shalit, A., Goldshmidt, A., Amsellem, Z., et al. (2006). The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. _Proc. Natl. Acad. Sci. U.S.A._ 103, 6398–6403. doi: 10.1073/pnas.0601620103

Lin, C. (2000). Photoreceptors and regulation of flowering time. _Plant Physiol._ 163, 39–50. doi: 10.1104/pp.123.1.39

Liu, L., Liu, C., Hou, X., Xi, W., Shen, L., Tao, Z., et al. (2012). FTIP1 is an essential regulator required for florigen transport. *PLoS Biol.* 10:e1001313. doi: 10.1371/journal.pbio.1001313

Liu, Y., Schiff, M., Marathe, R., and Dinesh-Kumar, S. (2002). Tobacco Rar1, EDS1 and NPR1/NIM1 like genes are required for N-mediated resistance to tobacco mosaic virus. *Plant J.* 30, 415–429. doi: 10.1046/j.1365-313X.2002.01297.x

McGarry, R. C., and Ayre, B. G. (2012). Manipulating plant architecture with members of the CETS gene family. *Plant Sci.* 188–189, 71–81. doi: 10.1016/j.plantsci.2012.03.002

Mockler, T. C., Guo, H., Yang, H., Duong, H., and Lin, C. (1999). Antagonistic actions of Arabidopsis cryptochromes and phytochrome B in the regulation of floral induction. *Development* 126, 2073–2082.

Molinero-Rosales, N., Latorre, A., Jamilena, M., and Lozano, R. (2004). SINGLE FLOWER TRUSS regulates the transition and maintenance of flowering in tomato. *Planta* 218, 427–434. doi: 10.1007/s00425-003-1109-1

Moon, J., Suh, S. S., Lee, H., Choi, K. R., Hong, C. B., Paek, N. C., et al. (2003). The SOC1 MADS−box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. *Plant J.* 35, 613–623. doi: 10.1046/j.1365-313X.2003.01833.x

Navarro, C., Abelenda, J. A., Cruz-Oró, E., Cuéllar, C. A., Tamaki, S., Silva, J., et al. (2011). Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. *Nature* 478, 119–122. doi: 10.1038/nature10431

Pin, P. A., Benlloch, R., Bonnet, D., Wremerth-Weich, E., Kraft, T., Gielen, J. J., et al. (2010). An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science* 330, 1397–1400. doi: 10.1126/science.1197004

Pnueli, L., Carmel-Goren, L., Hareven, D., Gutfinger, T., Alvarez, J., Ganal, M., et al. (1998). The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. *Development* 125, 1979–1989.

Pnueli, L., Gutfinger, T., Hareven, D., Ben-Naim, O., Ron, N., Adir, N., et al. (2001). Tomato SP-interacting proteins define a conserved signaling system that regulates shoot architecture and flowering. *Plant Cell* 13, 2687–2702. doi: 10.1105/tpc.13.12.2687

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80, 847–857. doi: 10.1016/0092-8674(95)90288-0

Quail, P. H., Boylan, M. T., Parks, B. M., Short, T. W., Xu, Y., and Wagner, D. (1995). Phytochromes: photosensory perception and signal transduction. *Science* 268, 675–680. doi: 10.1126/science.7732376

Samach, A., Onouchi, H., Gold, S. E., Ditta, G. S., Schwarz-Sommer, Z., Yanofsky, M. F., et al. (2000). Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. *Science* 288, 1613–1616. doi: 10.1126/science.288.5471.1613

Shalit, A., Rozman, A., Goldshmidt, A., Alvarez, J. P., Bowman, J. L., Eshed, Y., et al. (2009). The flowering hormone florigen functions as a general systemic regulator of growth and termination. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8392–8397. doi: 10.1073/pnas.0810810106

Song, Y. H., Smith, R. W., To, B. J., Millar, A. J., and Imaizumi, T. (2012). FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science* 336, 1045–1049. doi: 10.1126/science.1219644

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303, 1003–1006. doi: 10.1126/science.1091761

Velásquez, A. C., Chakravarthy, S., and Martin, G. B. (2009). Virus-induced gene silencing (VIGS) in *Nicotiana benthamiana* and tomato. *J. Vis. Exp.* 28:1292. doi: 10.3791/1292

Weller, J. L., and Reid, J. B. (1993). Photoperiodism and photocontrol of stem elongation in two photomorphogenic mutants of *Pisum sativum* L. *Planta* 189, 15–23. doi: 10.1007/BF00201338

Weller, J. L., Schreuder, M. E., Smith, H., Koornneef, M., and Kendrick, R. E. (2000). Physiological interactions of phytochromes A, B1 and B2 in the control of development in tomato. *Plant J.* 24, 345–356. doi: 10.1046/j.1365-313x.2000.00879.x

Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., et al. (2005). Integration of spatial and temporal information during floral induction in Arabidopsis. *Science* 309, 1056–1059. doi: 10.1126/science.1114358

Yamaguchi, A., Kobayashi, Y., Goto, K., Abe, M., and Araki, T. (2005). TWIN SISTER OF FT (TSF) acts as a floral pathway integrator redundantly with FT. *Plant Cell Physiol.* 46, 1175–1189. doi: 10.1093/pcp/pci151

Yanovsky, M. J., and Kay, S. A. (2003). Living by the calendar: how plants know when to flower. *Nat. Rev. Mol. Cell Biol.* 4, 265–276. doi: 10.1038/nrm1077

Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W., and Chua, N.-H. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* 1, 641–646. doi: 10.1038/nprot.2006.97

Physiological Traits Associated with Wheat Yield Potential and Performance under Water-Stress in a Mediterranean Environment

Alejandro del Pozo [1]*, Alejandra Yáñez [1,2], Iván A. Matus [3], Gerardo Tapia [3], Dalma Castillo [3], Laura Sanchez-Jardón [4] and José L. Araus [5]

[1] Programa de Investigación de Excelencia Interdisciplinaria, Adaptación de la Agricultura al Cambio Climático (A2C2), Facultad de Ciencias Agrarias, Centro de Mejoramiento Genético y Fenómica Vegetal, Universidad de Talca, Talca, Chile, [2] Departamento de Ciencias Agrarias, Facultad de Ciencias Agrarias y Forestales, Universidad Católica del Maule, Curicó, Chile, [3] Centro Regional Investigación Quilamapu, Instituto de Investigaciones Agropecuarias, Chillán, Chile, [4] Centro Universitario de la Patagonia, Universidad de Magallanes, Coyhaique, Chile, [5] Unitat de Fisiologia Vegetal, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

Different physiological traits have been proposed as key traits associated with yield potential as well as performance under water stress. The aim of this paper is to examine the genotypic variability of leaf chlorophyll, stem water-soluble carbohydrate content and carbon isotope discrimination (Δ^{13}C), and their relationship with grain yield (GY) and other agronomical traits, under contrasting water conditions in a Mediterranean environment. The study was performed on a large collection of 384 wheat genotypes grown under water stress (WS, rainfed), mild water stress (MWS, deficit irrigation), and full irrigation (FI). The average GY of two growing seasons was 2.4, 4.8, and 8.9 Mg ha^{-1} under WS, MWS, and FI, respectively. Chlorophyll content at anthesis was positively correlated with GY (except under FI in 2011) and the agronomical components kernels per spike (KS) and thousand kernel weight (TKW). The WSC content at anthesis (WSCCa) was negatively correlated with spikes per square meter (SM2), but positively correlated with KS and TKW under WS and FI conditions. As a consequence, the relationships between WSCCa with GY were low or not significant. Therefore, selecting for high stem WSC would not necessary lead to genotypes of GY potential. The relationship between Δ^{13}C and GY was positive under FI and MWS but negative under severe WS (in 2011), indicating higher water use under yield potential and MWS conditions.

Keywords: carbohydrate, carbon isotope discrimination, chlorophyll, drought, stem reserves

INTRODUCTION

Since the Green Revolution the yields of wheat and other cereals have increased considerably in many regions of the world, including Chile (Calderini and Slafer, 1998; Engler and del Pozo, 2013; del Pozo et al., 2014), as a result of genetic improvement and better agronomic practices. The yield potential, i.e., the yield achieved when the best available technology is used, has also increased almost linearly since the sixties, particularly in more favorable environments where soil water availability is not limited (Zhou et al., 2007; Fischer and Edmeades, 2010; Matus et al., 2012; del Pozo et al., 2014). Yield under water-limiting conditions, such those of the rainfed

Mediterranean environments, has also increased during the past decades (Sánchez-García et al., 2013). Notwithstanding the possible need for phenological adjustment (earliness) a higher yield potential may also translate into a higher performance under water stress (Nouri et al., 2011; Hawkesford et al., 2013). However, the potential yield and water-limited yield of wheat needs to continue increasing in order to cope with future demand for food, which is a consequence of the growing population and changes in social habits (Fischer, 2007; Hawkesford et al., 2013), and also to reduce the negative impacts on crop productivity of global climate change (Lobell et al., 2008; Lobell and Gourdji, 2012).

The increase, in the yield potential and stress adaptation of wheat has been attained mainly through empirical selection for grain yield (GY). However, there is evidence that phenotyping using physiological traits, as a complement to agronomic traits, may help in identifying selectable features that accelerate breeding for yield potential and performance under drought (Araus et al., 2002, 2008; Fischer, 2007; Foulkes et al., 2007; Cattivelli et al., 2008; Fleury et al., 2010). The increases in yield potential of wheat since the sixties have been both positively correlated with shoot dry matter and harvest index (HI); the latter also being positively associated with water-soluble carbohydrate (WSC) content of stems at anthesis (Foulkes et al., 2007). Under water limiting conditions, various physiological process and traits have been associated with GY (e.g., Araus et al., 2002, 2008; Condon et al., 2004; Reynolds et al., 2006; Tambussi et al., 2007). Among them are traits related to pre-anthesis accumulation of WSC in stems and its further use during grain filling (Ehdaie et al., 2006a,b; Reynolds et al., 2006), delays in senescence

during grain filling assessed via changes in leaf color (Lopes and Reynolds, 2012), and those related to water use efficiency, in particular carbon isotope discrimination ($\Delta^{13}C$) in kernels (Richards et al., 2002; Araus et al., 2003, 2008).

WSCs are accumulated in stems prior to anthesis and then are remobilized to the grain during the grain-filling period (Blum, 1998; Bingham et al., 2007). Under water limiting conditions, where canopy photosynthesis is inhibited, the contribution of stem carbohydrate to grain growth could be very significant (Ehdaie et al., 2006a,b; Reynolds et al., 2006). Both spring and winter wheat lines have been shown to vary significantly for WSC concentration and WSC content in stems around anthesis (Ruuska et al., 2006; Foulkes et al., 2007; Yang et al., 2007), whereas positive correlations have been observed between accumulated WSC at anthesis and GY in winter wheat genotypes (Foulkes et al., 2007), as well as with kernel weight in recombinant inbred lines (RILs) from the Seri/Babax population (Dreccer et al., 2009). However, stem WSC concentrations can be negatively correlated with stem number m^{-2} (Dreccer et al., 2013).

Drought increases senescence, by accelerating chlorophyll degradation, leading to a decrease in leaf area and canopy photosynthesis. There is evidence that stay-green phenotypes with delayed leaf senescence can improve their performance under drought conditions (Rivero et al., 2007; Lopes and Reynolds, 2012).

$\Delta^{13}C$ can be used as a selection criterion for high water use efficiency (Condon et al., 2004; Richards, 2006), but also can provide an indirect determination of the effective water used by the crop (Araus et al., 2002, 2008; Blum, 2009). In fact,
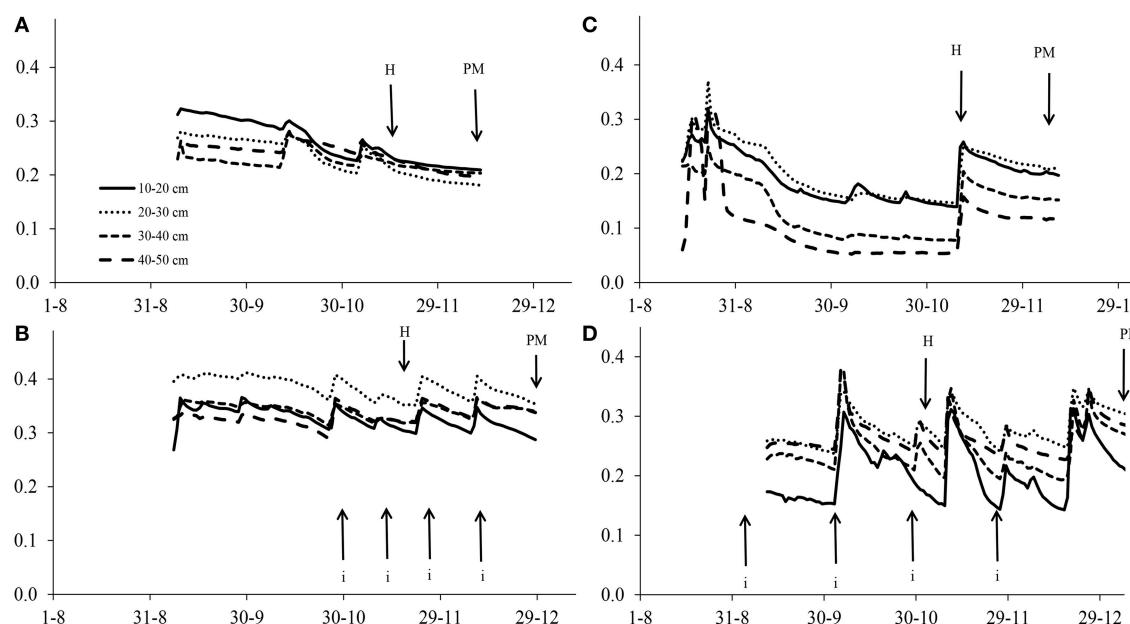
FIGURE 1 | Volumetric water content in the soil at severe water stress in Cauquenes (A,C) and at full irrigation in Santa Rosa (B,D) in 2011 (A,B) and 2012 (C,D). Each of the soil depth values are means from two sensors (replicates). Abbreviations H and PM refer to the dates of heading and physiological maturity, respectively. Dates of irrigation (i) at Santa Rosa are marked with arrows.

kernel Δ^{13}C can be positively or negatively correlated with GY depending on soil water availability. Indeed, under moderate stress to well-watered Mediterranean conditions Δ^{13}C has been reported to be positively correlated with GY in wheat (Araus et al., 2003, 2008 for wheat) and barley (Acevedo et al., 1997; Voltas et al., 1999; del Pozo et al., 2012), whereas the opposite trend has been reported under severe drought conditions (but see Araus et al., 1998).

In this study we investigated the genotypic variability of flag leaf chlorophyll content (measured with a portable leaf meter), stem WSC accumulation at anthesis and the Δ^{13}C of mature kernels, as well as the relationship of these traits with GY and its agronomical components, in spring bread wheat under contrasting water conditions in a Mediterranean environment. It is hypothesized that within a large set (384 genotypes) cultivars and advanced lines of spring bread wheat there is high genotypic variability for agronomic and physiological traits. In addition, the yield performance of genotypes under drought conditions is associated with stem WSC accumulation, delayed leaf senescence, and carbon discrimination in grains.

MATERIALS AND METHODS

Plant Material and Growing Conditions

A collection of 384 cultivars and advanced semidwarf lines of spring bread wheat (*Triticum aestivu*m L.), including 153 lines from the wheat breeding program of the Instituto de Investigaciones Agropecuarias (INIA) in Chile, 53 from the International Wheat and Maize Improvement Centre (CIMMYT) that were previously selected for adaptiveness to Chilean environments (these lines share common ancestors with the INIA-Chile breeding program), and 178 lines from INIA in Uruguay (Table S1). The objective with this set of lines was to create a germplasm base to breed for drier areas in Chile and subsequently other countries within the projects involved.

This large set of genotypes was evaluated in two Mediterranean sites of Chile: Cauquenes (35°58′ S, 72°17′ W; 177 m.a.s.l.) under the water stress (WS) typical of the rainfed at this site, and Santa Rosa (36°32′ S, 71°55′ W; 220 m.a.s.l.) under full irrigation (FI) and moderate water stress (MWS) conditions achieved through support irrigation. Trials were assayed during two consecutive (2011 and 2012) crop seasons, except for the MWS trial, which was only set up during 2011. Cauquenes corresponds to the Mediterranean drought-prone area of Chile; the average annual temperature is 14.7°C, the minimum average is 4.7°C (July) and the maximum is 27°C (January). The evapotranspiration is 1200 mm (del Pozo and del Canto, 1999) and the annual precipitation was 410 and 600 mm in 2011 and 2012, respectively. Santa Rosa corresponds to a high yielding area; the average annual temperature in this region is 13.0°C, the minimum average is 3.0°C (July) and the maximum is 28.6°C (January; del Pozo and del Canto, 1999). The annual precipitation was 736 and 806 mm, in 2011 and 2012, respectively.

The experimental design was an α-lattice with 20 incomplete blocks per replicate, each block containing 20 genotypes. In each replicate two cultivars (Don Alberto and Carpintero) were

included eight times. Two replicates per genotypes were used, except at Cauquenes and Santa Rosa SI in 2011 where a single replicate was established. Plots consisted of five rows of 2 m in length and 0.2 m distance between rows. The sowing rate was 20 g m^2 and sowing dates were: 07 September and 23 May, in 2011 and 2012, respectively at Cauquenes; 31 and 7 August, in 2011 and 2012, respectively at Santa Rosa. Because the sowing date in 2011 at Cauquenes was much later than in 2012, the water stress was more severe in the first year. Plots were fertilized with 260 kg ha^1 of ammonium phosphate (46% P$_2$O$_5$ and 18% N), 90 kg ha^{-1} of potassium chloride (60% K$_2$O), 200 kg ha^{-1} of sul-po-mag (22% K$_2$O, 18% MgO, and 22% S), 10 kg ha^{-1} of boronatrocalcite (11% B), and 3 kg ha^{-1} of zinc sulfate (35% Zn). Fertilizers were incorporated with a cultivator before sowing. During tillering an extra 153 kg ha^{-1} of N was applied. Weeds were controlled with the application of Flufenacet + Flurtamone + Diflufenican (96 g a.i.) as pre-emergence controls and a further application of MCPA (525 g a.i.) + Metsulfuron-metil (5 g a.i.) as post-emergents. Cultivars were disease resistance and no fungicide was used.

Furrow irrigation was used in Santa Rosa: one irrigation at the end of tillering (Zadocks Stage 21; Zadoks et al., 1974) in the
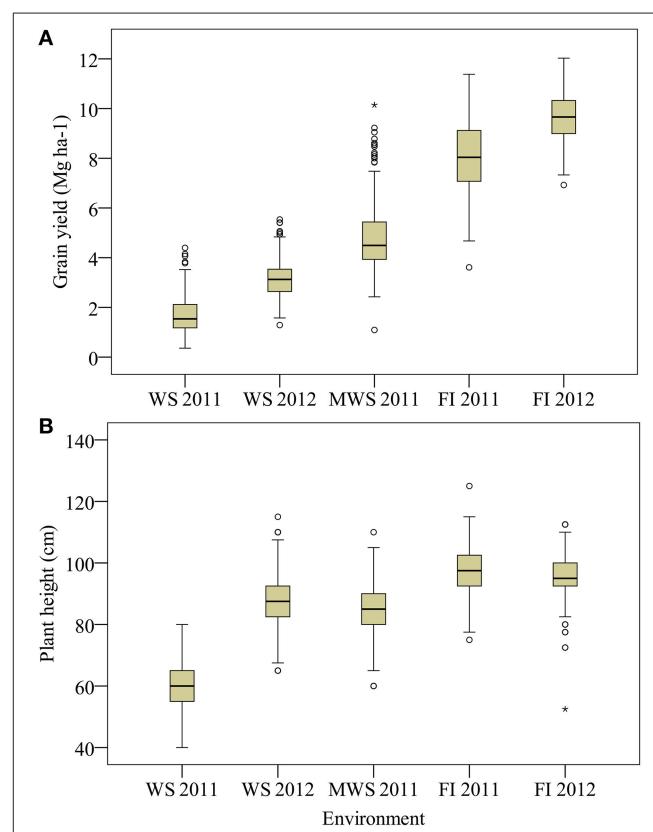
FIGURE 2 | Grain yield (A) and plant height (B) for 378 genotypes of wheat grown under water stress (Cauquenes WS), mild water stress (Santa Rosa MWS) and full irrigation (Santa Rosa FI) in two growing seasons (2011 and 2012), except at MWS. Box and whisker show population minimum, 25th percentile/median/75th percentile and maximum. The open symbols indicate outlier data.

MWS trial and four irrigations at the end of tillering, the flag leaf stage (Z37), heading (Z50), and middle grain filling (Z70) in the FI trial respectively. Soil moisture at 10–20, 20–30, 30–40, and 40–50 cm depth was determined by using 10HS sensors (Decagon Devices, USA) connected to an EM-50 data logger (Decagon Devices, USA). The 10HS sensor determines volumetric water content by measuring the dielectric constant of the soil using capacitance/frequency domain technology. Two sets of sensors were set up in each environment and mean values of two sensors per depth are presented in **Figure 1**.

Agronomical Traits

Days from emergence to heading (DH) were determined in Santa Rosa, through periodic (twice a week) observations, when approximately half of the spikes in the plot had already extruded. At maturity and for each plot the plant height (PH) of the different trials, up to the extreme of the spike (excluding awns), was measured, the number of spikes per m^2 (SM2) were determined for a 1 m length of an inside row, and the number of kernels per spike (KS) and 1000 kernel weight (TKW) were determined in 25 spikes taken at random. Grain yield was assessed by harvesting the whole plot.

Leaf Chlorophyll Content and Water-Soluble Carbohydrates

Chlorophyll content (SPAD index) was determined at anthesis and then during grain filling about 2 weeks after anthesis (both measured on given calendar dates) in five flag leaves

per plot using a SPAD 502 (Minolta Spectrum Technologies Inc., Plainfield, IL, USA) portable leaf chlorophyll meter. WSC concentration in stems (harvested at ground level and excluding leaf laminas and sheaths) was determined at anthesis and maturity, on five main stems per plot, using the anthrone reactive method (Yemm and Willis, 1954). The stem length was measured and then dried for 48 h at 60°C, weighed and ground. Next, a 100 mg subsample was used for WSC extraction, with 3 ml of extraction buffer containing 80% ethanol 10 mM Hepes-KOH (pH = 7.5), and incubated at 60°C overnight. Then, to separate the debris, the samples were centrifuged at 60 rpm for 30 min. The anthrone reagent was added to each supernatant and placed over a hotplate at 80°C for 20 min. Finally, the absorbance of the sample was measured at 620 nm in an EPOCH microplate UV-Vis Spectrophotometer (Biotek) using COSTAR 3636 96 well-plates (Corning) for the UV range. WSC content per whole stem and per unit land area were calculated as WSC concentration per unit stem weight (mg CHO g stem^{-1}), and WSC content per unit of stem (mg CHO stem^{-1}) and per unit grown area (g CHO m^{-2}), respectively. In addition, the apparent WSC remobilization was calculated as the differences from anthesis to maturity in WSC content on a stem and land area basis.

Stable Carbon Isotope Analysis

The stable carbon (^{13}C/^{12}C) isotope ratio was measured in mature kernels using an elemental analyser (ANCA-SL, PDZ Europa, UK) coupled with an isotope ratio mass spectrometer, at the Laboratory of Applied Physical Chemistry

TABLE 1 | *F*-values of ANOVA for agronomic and physiological traits, for 378 genotypes of wheat grown under severe water stress (Cauquenes WS) and full irrigation (Santa Rosa FI) in two growing seasons.

Agronomic traits	Year	Genotype (G)	Environment (E)	G × E	Physiological traits	Year	Genotype (G)	Environment (E)	G × E
GY	2011[a]	1.2*	5001.9***	0.6	SPADa	2011[a]	4.1***	211***	2.6***
	2012	2.5***	24217***	1.8***		2012	4.6***	4633***	2.0***
DH	2011	–	–	–	SPADgf	2011	–	–	–
	2012	–	–	–		2012	2.4***	1952***	1.8***
PH	2011	4.0***	12052***	1.4**	SWa	2011	0.7	76***	0.4
	2012	3.3***	543***	1.1		2012	5.3***	2398***	1.3**
SM2	2011	4.2***	1310***	1.4**	SWm	2011	0.5	128***	0.4
	2012	5.5***	7978***	2.1***		2012	6.2***	0.9	1.3**
KS	2011	6.6***	2383***	1.8***	WSCa	2011	1.3*	133***	1.0
	2012	3.8***	613***	1.6***		2012	1.2*	1249***	1.2*
TKW	2011	15.1***	6725***	1.4**	WSCm	2011	6.7***	2963***	6.6***
	2012	14.1***	906***	1.3**		2012	1.5***	13***	1.1
KM2	2011	2.2***	1860***	0.9	WSCCa	2011	0.7	3.8*	0.5
	2012	2.6***	6722***	1.7***		2012	2.4***	2305***	1.3**
					WSCCm	2011	1.3*	78***	1.1
						2012	1.8***	7.7**	1.2**
					Δ^{13}C	2011	1.1	7512***	0.7
						2012	4.3***	45500***	1.6***

GY, grain yield; DH, day to heading; PH, plant height; SM2, spikes m^{-2}; KS, kernel per spike; TKW, thousand kernel weight; KM2, kernels m^{-2}; SPADa, SPAD anthesis; SPADgf, SPAD grain filling; SWa, stem weight at anthesis; SWm, stem weight at maturity; WSCa, WSC at anthesis; WSCm, WSC at maturity; WSCCa, WSC content anthesis; WSCCm, WSC content maturity; Δ^{13}C, kernel Δ^{13}C.

[a]In 2011, 10 genotypes were discarded from the analysis due to low spike numbers.

*P < 0.05; **P < 0.001; ***P < 0.0001.
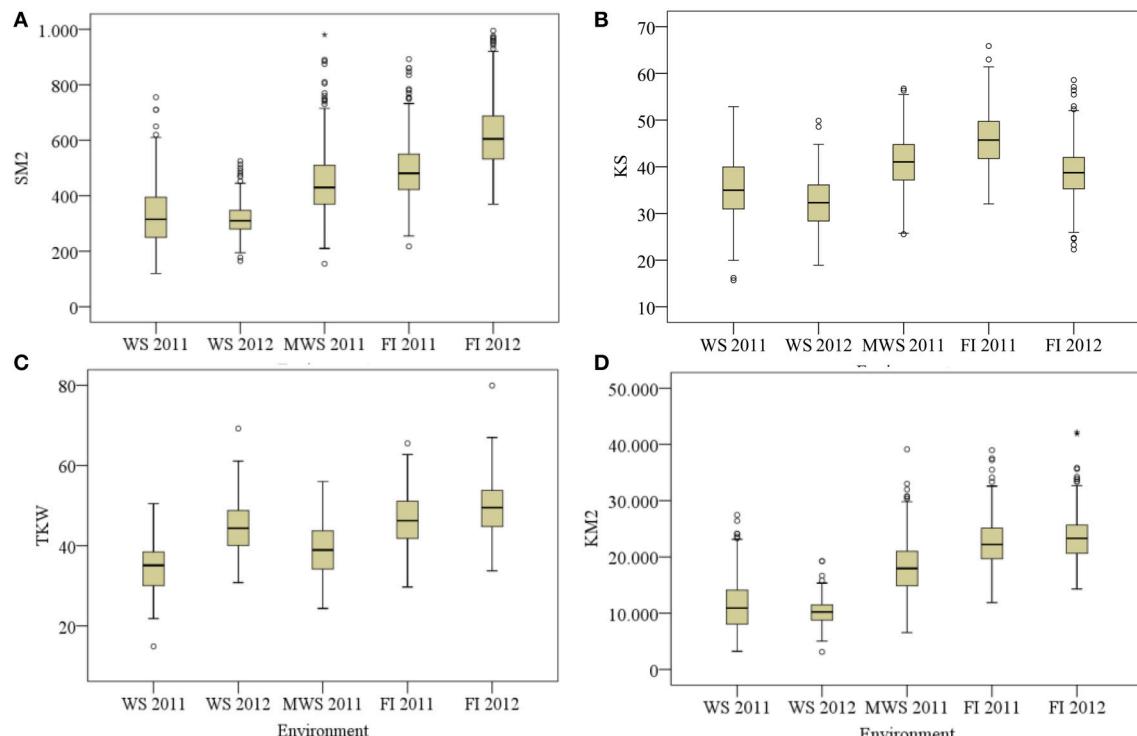
FIGURE 3 | Spikes per m² (SM2; A), kernel per spike (KS; B), thousand kernel weight [TKW (g); C] and kernels m⁻² (KM2; D) for 378 genotypes of wheat grown under water stress (Cauquenes WS), mild water stress (Santa Rosa MWS) and full irrigation (Santa Rosa FI) in two growing seasons (2011 and 2012), except at MWS. Box and whisker show population minimum, 25th percentile/median/75th percentile and maximum. The open symbols indicate outlier data.

at Ghent University (Belgium). The $^{13}C/^{12}C$ ratios were expressed in δ notation (Coplen, 2008) determined by: $\delta^{13}C = (^{13}C/^{12}C)_{sample}/(^{13}C/^{12}C)_{standard} - 1$ (Farquhar et al., 1989), where sample refers to plant material and standard to the laboratory standards that have been calibrated against international standards from Iso-Analytical (Crewe, Cheshire, UK). The precision of $\delta^{13}C$ analyses was 0.3‰ (SD, $n = 10$). Further, the carbon isotope discrimination ($\Delta^{13}C$) of kernels was calculated as: $\Delta^{13}C$ (‰) $= (\delta^{13}C_a - \delta^{13}C_p)/[1+ (\delta^{13}C_p)/1000]$, where a and p refer to air and the plant, respectively (Farquhar et al., 1989). $\delta^{13}C_a$ from the air was taken as $-8.0‰$.

Yield Tolerance Index

The yield tolerance index (YTI), which combines the relative performance of a genotype under drought with its potential yield under irrigated conditions (Ober et al., 2004), was calculated as:

$$YTI = \left(\frac{Y_D}{\overline{Y}_D}\right)\left(\frac{Y_I}{\overline{Y}_I}\right)\left(\frac{\overline{Y}_D}{\overline{Y}_I}\right) = \left(\frac{Y_D Y_I}{\overline{Y}_I^2}\right) \qquad (1)$$

where Y_D and Y_I are the genotype mean yield under drought (Cauquenes) and irrigation conditions (Santa Rosa, fully irrigation), respectively, and $\overline{Y}_D$ and $\overline{Y}_I$ are the mean yield of all genotypes growing under drought and irrigated conditions, respectively.

Statistical Analysis

In 2011, 10 genotypes were discarded from analysis due to low emergence. In addition six genotypes from Uruguay were discarded from the analysis for having late heading time (more than 100 days) an plant height >120 cm. ANOVAs for physiological and yield-related traits were performed for the whole set of genotypes using PROC MIXED of the SAS Institute Inc. Genotypes and environment (Cauquenes WS and Santa Rosa FI) were considered fixed effects, whereas blocks and incomplete blocks within each replication (in an α-lattice design) were considered random effects. Data from Santa Rosa MWS where not considered in the ANOVAs because there was no replication and only one year (2011) of observations. Correlation analysis was performed between agronomic and physiological traits, and also stepwise regressions between grain yield and related agronomical and physiological traits. Principal component analysis (PCA) was carried out for the 378 genotypes using the mean values for physiological and agronomical traits evaluated under severe water stress in Cauquenes and full irrigation in Santa Rosa, in two growing seasons, using IBM SPSS Statistics 19.

RESULTS

Agronomical and Physiological Traits

For SM2, KS and TKW the genotype x environment (GxE) interaction was highly significant ($P < 0.001$) in both growing

TABLE 2 | Means ± standard deviation and ranges (minimum–maximum) for chlorophyll content in SPAD units, stem weight, water-soluble carbohydrate (WSC) concentration and content per stem at anthesis and maturity, and carbon isotope discrimination ($\Delta^{13}C$) in kernels, for 378 genotypes of wheat grown under severe water stress (Cauquenes WS), mild water stress (Santa Rosa MWS) and full irrigation (Santa Rosa FI) in two growing seasons, except at MWS.

Trait	Year	Cauquenes (WS)		Santa Rosa (MWS)		Santa Rosa (FI)	
		Mean	Range	Mean	Range	Mean	Range
Days to heading	2011	n.e.	n.e.	78 ± 3	71–86	79 ± 3	73–89
	2012	n.e.	n.e.	n.e.	n.e.	85 ± 4	78–100
SPAD anthesis	2011	47 ± 5	32–60	47 ± 3	34–56	45 ± 3	36–52
	2012	42 ± 3	32–50	n.e.	n.e.	49 ± 3	41–58
SPAD grain filling	2011	n.e.	n.e.	31 ± 12	4–52	43 ± 3	30–51
	2012	34 ± 8	4–47	n.e.	n.e.	48 ± 3	35–55
Stem weight anthesis (g)	2011	1.24 ± 0.31	0.50–2.44	1.68 ± 0.40	0.86–2.88	1.67 ± 0.36	0.91–2.84
	2012	1.81 ± 0.36	0.91–3.04	n.e.	n.e.	1.19 ± 0.27	0.54–1.95
Stem weight maturity (g)	2011	0.70 ± 0.19	0.26–1.48	0.98 ± 0.26	0.41–3.03	1.14 ± 0.25	0.57–2.06
	2012	1.06 ± 0.24	0.59–1.79	n.e.	n.e.	1.06 ± 0.23	0.54–2.13
WSC concentration anthesis (mg g^{-1})	2011	178.4 ± 56.0	63.9–548.3	152.6 ± 40.7	74.0–431.1	139.9 ± 41.7	29.1–349.7
	2012	226.7 ± 36.2	110.5–444.3	n.e.	n.e.	142.3 ± 38.3	64.3–381.7
WSC concentration maturity (mg g^{-1})	2011	43.4 ± 23.2	15.8–171.3	36.4 ± 26.7	3.7–224.2	17.7 ± 5.4	7.0–32.7
	2012	48.5 ± 18.1	20.7–146.6	n.e.	n.e.	41.7 ± 17.8	14.7–153.5
Total WSC content anthesis (mg stem^{-1})	2011	221.5 ± 91.5	49.8–770.6	256.8 ± 92.3	80.8–758.7	234.9 ± 95.5	58.6–636.4
	2012	413.2 ± 110.5	158.0–988.9	n.e.	n.e.	174.2 ± 73.1	37.8–583.5
Total WSC content maturity (mg stem^{-1})	2011	30.6 ± 19.6	7.3–157.7	36.9 ± 33.0	3.3–291.5	20.2 ± 8.4	3.2–53.2
	2012	53.1 ± 27.7	17.1–169.9	n.e.	n.e.	43.9 ± 24.7	6.5–195.1
Kernel $\Delta^{13}C$ ($^0/_{00}$)	2011	14.2 ± 0.6	11.9–16.3	16.5 ± 0.9	14.6–19.6	18.1 ± 0.6	16.2–19.6
	2012	14.9 ± 0.4	13.4–16.1	n.e.	n.e.	18.7 ± 0.4	17.4–20.1

seasons, whereas for GY, PH, and KM2 was only in one growing season (**Table 1**). Among the physiological traits, the SPAD index exhibited a significant ($P < 0.001$) GxE interaction in both growing seasons, but stem weight and WSC concentration and content, and $\Delta^{13}C$ of kernels was only in 2012 (**Table 1**).

Under FI in Santa Rosa, the average GY of the three sets of wheat genotypes (378 in total) was 8–10 Mg ha^{-1} but some genotypes produced up to 12 Mg ha^{-1} (**Figure 2A**). Under MWS in Santa Rosa the average GY was 4.8 Mg ha^{-1}. Under WS GY was significantly ($P < 0.0001$) reduced in Cauquenes, by 79 and 68% in 2011 and 2012, respectively, compared to Santa Rosa under FI (**Figure 2A**). Also, plant height was reduced under WS by 40 and 9% in 2011 and 2012, respectively (**Figure 2B**).

The reduction in SM2, KS and TKW under WS compared with FI was in general more pronounced in the first growing season; on average (of the two growing seasons) these traits were reduced by 25, 41, 21, and 18%, respectively, whereas KM2 was reduced by 53% (**Figure 3**).

The relationships for GY under FI and WS showed no significant correlation in both years ($P > 0.05$). The yield tolerance index (YTI) of the 378 genotypes based on GY under WS and FI presented a wide range of values in both years, from 0.05 (very susceptible) to 0.65 (very tolerant genotypes). The frequency distribution of YTI had a left-skewed deviation in 2011 (mean YTI = 0.21) compared to 2012 (mean YTI = 0.32).

Days to heading, determined under FI, differed by about 20 days between the earliest and latest genotypes (**Table 2**). A wide range of SPAD index values among genotypes was observed in environments (WS, MWS, and FI) and growing seasons (**Table 2**). A significant reduction ($P < 0.001$) in the SPAD index at anthesis and during grain filling was observed under WS in 2012.

Stem weight and stem WSC concentration and content were much higher at anthesis compared to maturity. Their average reductions over two growing seasons were about 43, 77, and 87%, respectively, under WS at Cauquenes, and 23, 79, and 84%, respectively, in Santa Rosa under FI (**Table 2**). The apparent WSC remobilization was on average 279, 220, and 170 mg per stem under WS, MWS, and FI, respectively (data not shown).

The WSC concentration and content per stem at anthesis and maturity presented large genotypic variabilities in all the environments (**Table 2**). The stem WSC per unit area (g m^{-2}) at anthesis was highly correlated to the WSC concentration ($r = 0.66$ and 0.84, $P < 0.001$, for WS for FI, respectively, in 2012) and the stem biomass (g m^{-2}; $r = 0.81$ and 0.66, $P < 0.001$, for a WS for FI, respectively, in 2012).

The $\Delta^{13}C$ also exhibited genotypic variability under WS, MWS, and FI (**Table 2**), but lower values were found under WS compared to FI.

TABLE 3 | Correlations between physiological traits and grain yield and their agronomical components of 378 genotypes of wheat grown under water stress (WS) in Cauquenes and full irrigation (FI) in Santa Rosa, in 2011 and 2012.

	Grain yield				Spikes m⁻²				Kernels spike⁻¹				Thousand kernel weight			
	WS 2011	WS 2012	FI 2011	FI 2012	WS 2011	WS 2012	FI 2011	FI2012	WS 2011	WS 2012	FI 2011	FI2012	WS 2011	WS 2012	FI 2011	FI 2012
DH	−0.03	0.08	−0.02	**0.17**	**0.18**	**0.16**	**0.35**	**0.32**	**0.12**	0.02	0.02	0.08	**−0.54**	**−0.35**	**−0.43**	0.08
SPADa	**0.39**	**0.32**	−0.04	**0.22**	**0.15**	**−0.27**	**−0.43**	**−0.38**	**0.41**	**0.21**	**0.25**	**0.23**	0.10	**0.36**	**0.44**	**0.36**
SPADgf	n.e.	**0.19**	0.09	**0.24**	n.e.	**−0.21**	**−0.33**	**−0.30**	n.e.	0.05	**0.30**	**0.21**	n.e.	**0.33**	**0.41**	**0.29**
SWD	0.03	0.08	**−0.18**	0.00	**−0.26**	**−0.21**	**−0.39**	**−0.26**	**0.21**	**0.29**	**0.17**	**0.17**	**0.28**	**0.10**	**0.34**	**0.19**
WSCa	**−0.26**	−0.09	**−0.37**	**−0.14**	**−0.32**	0.05	**−0.10**	**−0.28**	−0.02	−0.02	−0.07	0.09	0.05	**−0.11**	−0.09	**0.25**
WSCm	**−0.28**	−0.02	**−0.12**	0.00	**−0.22**	**−0.12**	**−0.14**	0.06	**−0.33**	−0.09	−0.01	**−0.14**	**0.18**	**0.15**	**0.12**	0.02
WSCCa	−0.07	**0.15**	**−0.33**	0.05	**−0.41**	**−0.30**	**−0.45**	**−0.53**	**0.22**	**0.28**	**0.19**	**0.27**	**0.38**	**0.31**	**0.30**	**0.50**
WSCCm	**−0.13**	**0.14**	−0.06	0.11	**−0.24**	**−0.27**	**−0.43**	**−0.17**	**−0.11**	0.06	**0.26**	0.03	**0.38**	**0.39**	**0.43**	**0.22**
WSCCD	−0.03	**0.12**	**−0.33**	0.01	**−0.38**	**−0.25**	**−0.42**	**−0.47**	**0.27**	**0.28**	**0.16**	**0.26**	**0.30**	**0.22**	**0.28**	**0.43**
Δ¹³C	**−0.12**	**0.53**	**0.51**	**0.42**	−0.06	0.09	0.05	−0.07	−0.01	**0.26**	**0.23**	**0.21**	**0.16**	**0.11**	**0.21**	**0.13**

Coefficients of significant correlations (at P < 0.05) are in bold characters. HD, day to heading; SPADa, SPAD at anthesis; SPADgf, SPAD at grain filling; SWD, stem weight difference (SWa–SWm); WSCa, WSC at anthesis; WSCm, WSC at maturity; WSCCa, WSC content at anthesis; WSCCm, WSC content at maturity; Δ¹³C, kernel Δ¹³C. n.e.: not evaluated.

Relationships between Yield, Agronomical, and Physiological Traits

GY was positively correlated with SM2 and KM2, but negatively correlated with TKW, in both water regimes and growing seasons (**Figure 4**). GY was also positively correlated ($r = 0.3–0.52$, $P < 0.001$) with plant height in all the environments.

Days to heading (determined at FI) was not correlated with GY, but it was positively correlated with SM2 and negatively correlated with TKW, except under FI in 2012 (**Table 3**). The SPAD index was positive and significantly correlated with GY (except under FI in 2011) and the agronomical components KS and TKW (**Table 3**). The WSC content at anthesis (WSCCa) was negatively correlated with SM2, but positively correlated with KS and TKW under WS and FI conditions (**Figure 5**). As a consequence, GY exhibited a low positive correlation with WSCCa under WS in 2012, and non or negative correlation under FI (**Table 3**).

The relationship between $\Delta^{13}C$ and GY was slightly negative under WS in 2011, but positive and highly significant in 2012, and also positive under MWS and FI in 2011 and 2012 (**Table 3**; **Figure 6A**). Indeed, Pearson correlation values of the relationship between $\Delta^{13}C$ vs. GY depended on the environment, increasing from low to medium yields and further declining at higher GY (**Figure 6B**). The correlation between $\Delta^{13}C$ and STI under SWS was not significant in 2011 but was positive and significant in 2012 ($r = 0.51$; $P < 0.01$).

PCA analysis indicated that the two first principal components (PC) explained >50% of the observed variability, under WS and FI conditions (**Figure 7**). KS was the agronomical component more close related with GY under WS and FI (except in 2011). Among the physiological traits, $\Delta^{13}C$ presented the strongest association with GY, except under the severe WS in 2011 (**Figure 7**). The SPAD index at anthesis was close associated with GY under WS in 2011, but with TKW under WS in 2012 and FI. WSCCa was also close related to TKW in all the environments, and days to heading was associated SM2.

The stepwise regression analysis between GY and related agronomical (SM2, TKW, and KS) and physiological (SPADa, WSCCa, and $\Delta^{13}C$) traits indicated that under water stress conditions, the contribution of the agronomical trails was greater than the physiological ones, but under full irrigation conditions WSCCa and $\Delta^{13}C$ contributed similarly to the agronomical traits to GY (**Table 4**).

DISCUSSION

The set of 378 wheat genotypes tested in this work exhibited a high phenotypic variability for physiological and agronomic traits. The water stress in Cauquenes was very severe as reflected in the low average GY (1.7 Mg ha⁻¹ in 2011). However, some genotypes were able to produce more than 4 Mg ha⁻¹ under such WS conditions and showed high values of YTI (>0.50). Actually, YTI was highly correlated ($r > 0.92$; $P < 0.0001$ in both years) with GY under WS in Cauquenes. Under the full irrigation conditions of Santa Rosa some genotypes achieved extremely high yields (12 Mg ha⁻¹),

	Grain yield				
	WS 2011	WS 2012	MWS 2011	FI 2011	FI 2012
Spikes m^{-2}	**0.60**	0.10	**0.35**	**0.30**	−0.14
Kernels spike^{-1}	**0.47**	**0.51**	**0.27**	**0.19**	**0.35**
Thousand kernel weight	0.01	**0.32**	**0.29**	0.13	**0.28**
Kernels m^{-2} (x10^3)	**0.75**	**0.46**	**0.53**	**0.50**	0.16

Coefficients of significant correlations (at $P < 0.05$) are in bold characters.
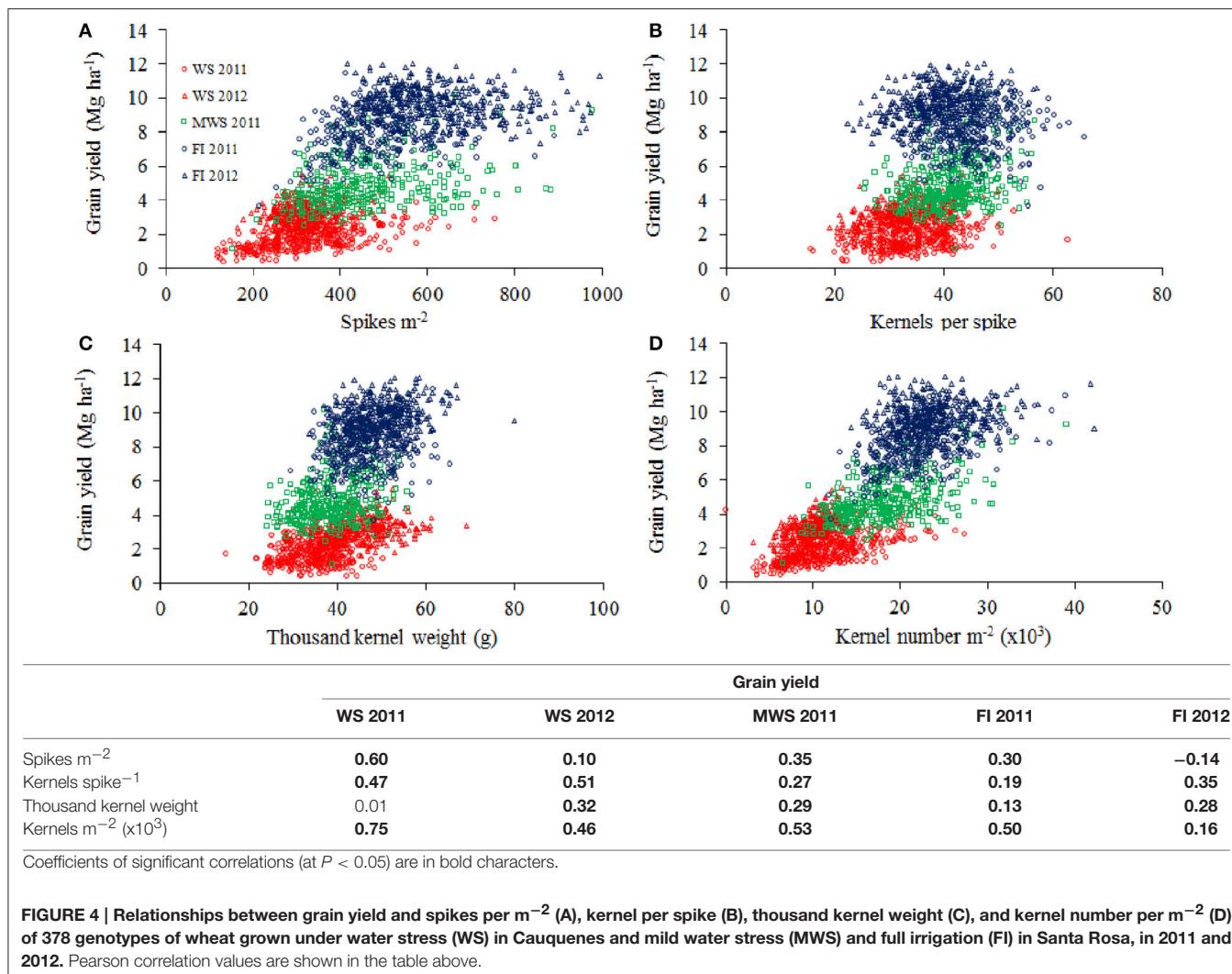
FIGURE 4 | Relationships between grain yield and spikes per m^{-2} (A), kernel per spike (B), thousand kernel weight (C), and kernel number per m^{-2} (D) of 378 genotypes of wheat grown under water stress (WS) in Cauquenes and mild water stress (MWS) and full irrigation (FI) in Santa Rosa, in 2011 and 2012. Pearson correlation values are shown in the table above.

for a Mediterranean environment. Large genotypic variability in GY and its agronomical components has also been found in 127 recombinant inbred lines (Dharwar Dry × Sitta) of wheat growing under severe water stress in Obregon, Mexico (Kirigwi et al., 2007), and in 105 lines of the double-haploid population (Weebil × Bacanora) in four contrasting high-yielding environments (García et al., 2013).

The strong reduction in GY under WS was mainly a consequence of the decline in SM2 (41%), followed by KS (21%), and as a consequence the number of kernels per m^2 was reduced (53%; **Table 2**). Thus, kernels per m^2 is the agronomical component most affected by drought, as previously reported by other authors (Estrada-Campuzano et al., 2012). In addition the TKW also decreased, but to a lesser extent (18%). As a consequence GY was positively correlated with the number of kernels m^{-2} (**Figure 4**; $r = 0.81$, $P < 0.0001$ for all the environments), but the correlation coefficients for each environment were not as high as has been reported by several authors (see Sinclair and Jamieson, 2006). In fact, a trade-off among the agronomical components was observed where SM2

was negatively correlated with KS under FI ($r = −0.50$ and $−0.58$ in 2011 and 2012, respectively) and TKW in WS ($r = −0.36$ and $−0.49$ in 2011 and 2012, respectively) and FI ($r = −0.60$ and $−0.58$ in 2011 and 2012, respectively) conditions. The PCA indicated that KS was better associated with GY in both WS and FI conditions (**Figure 7**). Other studies have also shown that KS but not TKW was associated with GY under water stress conditions (Denčić et al., 2000) and also a high-yielding environment (García et al., 2013).

Chlorophyll Content

Chlorophyll content at anthesis was positively correlated with GY and the agronomical components KS and TKW, particularly under WS (**Table 3**). Drought increases senescence by accelerating chlorophyll degradation leading to a decrease in leaf area and photosynthesis. There is evidence that stay-green phenotypes with delayed leaf senescence can improve their performance under drought conditions (Rivero et al., 2007; Lopes and Reynolds, 2012). In wheat and sorghum, genotypic variability has been detected in chlorophyll content as well as in the rate of
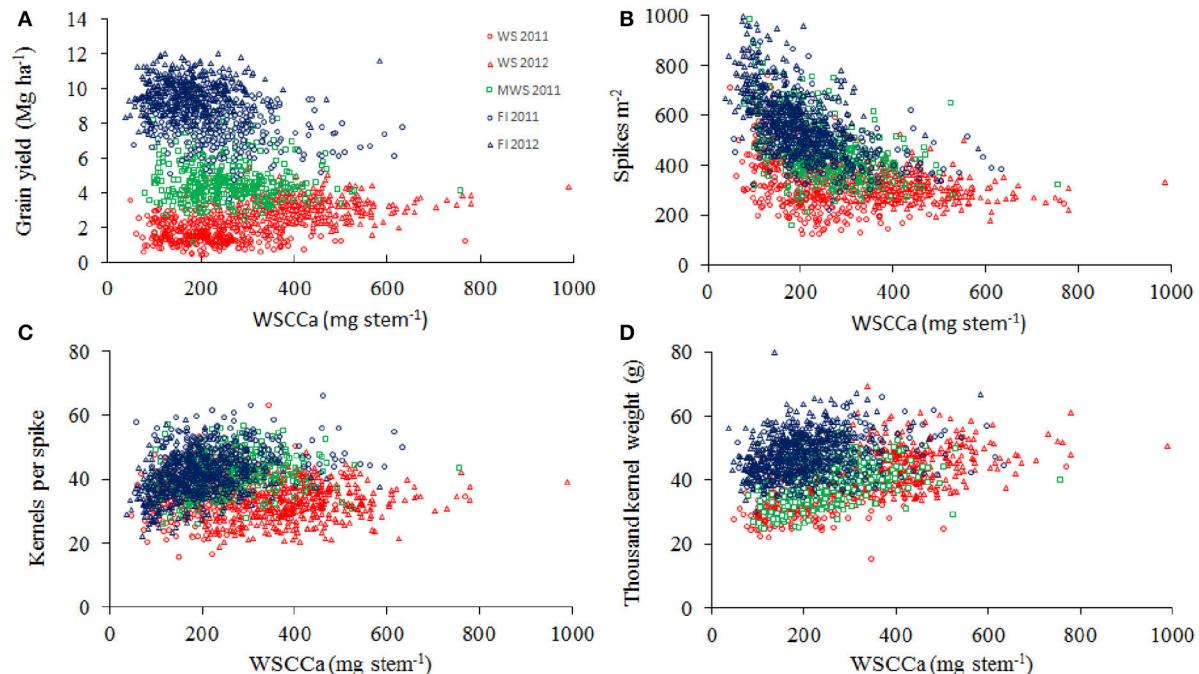
FIGURE 5 | Relationships between stem water-soluble carbohydrate content at anthesis (WSCCa) and grain yield (A), spikes per m^{-2} (B), kernel per spike (C), and thousand kernel weight (D) in 378 genotypes of wheat grown under water stress (WS) in Cauquenes and mild water stress (MWS) and full irrigation (FI) in Santa Rosa, in 2011 and 2012. Pearson correlation values are in **Table 3**.

TABLE 4 | Stepwise regression analysis between grain yield (GY) and related agronomical (TKW, SM2, and KS) and physiological (SPADa, WSCCa, and Δ^{13}C) traits of 378 genotypes of wheat grown under water stress (WS) in Cauquenes and full irrigation (FI) in Santa Rosa, in 2011 and 2012.

Trait	WS 2011–2012				FI 2011–2012			
	Coefficient B	SE	Contribution		Coefficient B	SE	Contribution	
			Share	Unique			Share	Unique
Constant	−7.218	0.618			−19.270	1.138		
TKW	0.068	0.004	30.9	13.9	0.089	0.007	16.8	8.1
SM2	0.006	0.000	39.2	19.9	0.005	0.000	16.4	7.9
KS	0.049	0.004	18.6	7.1	0.072	0.007	12.6	5.8
WSCCa	0.002	0.000	7.5	2.5	−0.004	0.000	16.8	8.1
Δ^{13}C	0.281	0.039	6.6	2.2	0.878	0.068	18.7	9.2
SPADa	−0.026	0.005	3.4	1.1	0.051	0.011	2.8	1.2
R^2	0.69				0.60			
N	722				734			
Mean square	84.21				183.14			
P	< 0.0001				P < 0.0001			

leaf senescence (measured with a portable leaf chlorophyll meter) during grain-filling (Harris et al., 2007; Lopes and Reynolds, 2012). In durum wheat (*Triticum turgidum* ssp. durum) stay-green mutants growing under glasshouse conditions remained green for longer and had higher rates of leaf photosynthesis and seed weight (Spano et al., 2003). These mutants with the stay-green characteristic also had higher levels of expression of the Rubisco small subunit of (RBCS) and chlorophyll a/b

binding protein (Rampino et al., 2006). Bread wheat genotypes with functional stay-green characteristics have also shown higher GY and total biomass in field conditions (Chen et al., 2010). Another study on Canadian spring wheat revealed that GY was positively correlated with green flag leaf duration and total flag leaf photosynthesis (Wang et al., 2008). Studies on spring wheat in the USA found a positive correlation between the stay-green trait and GY and grain weight in both water-limited and

FIGURE 6 | (A) Relationships between grain yield (GY) and carbon discrimination in grains (Δ^{13}C) for the set of 384 spring wheat genotypes grown under water stress (○, △), mild water stress (□) and full irrigation (●, ▲) conditions in 2011 and 2012. The negative regression is for water stress in 2011 and the positive one is for the rest of the environments. (B) The relationship between GY and correlation coefficients between GY and Δ^{13}C for each replicate (block) and environment (WS, MWS and FI in 2011, and WS and FI in 2012).

well-watered conditions (Blake et al., 2007). Therefore, a delay in leaf senescence would increase the amount of fixed carbon available for grain filling.

Stem Water-Soluble Carbohydrate

Large genotypic variability in stem WSC concentration and content was found at anthesis and maturity, in both environments (**Table 2**; **Figure 5**). Other studies conducted in spring and winter wheat lines have also found large variability in WSC concentration and WSC content on an area basis in stems around the time of anthesis (Ruuska et al., 2006; Foulkes et al., 2007; Yang et al., 2007). WSCs are accumulated in stems prior to anthesis and are then remobilized to the grain during the grain-filling period (Blum, 1998; Bingham et al., 2007). Indeed under water limiting conditions, where canopy photosynthesis is inhibited, the contribution of stem carbohydrate to grain growth could be very significant (Ehdaie et al., 2006a,b; Reynolds et al., 2006). In our study, more carbohydrate was accumulated at anthesis under WS than under FI, and the decline in stem WSC from anthesis to maturity was greater under WS, particularly in

2012 (360 vs. 130 mg per stem under WS and FI, respectively). This suggests that there was a larger remobilization of reserves during grain filling under WS. However, there were no clear relationships between the stem WSCCa, or the apparent WSC remobilization and GY, varying the correlation values from not significant to negative on the different environments (**Table 3**; **Figure 5**). Zhang et al. (2015) found also no significant correlation between stem WSC and GY in 20 genetically diverse double haploids derived from the cross of cvs. Westonia × Kauz, growing under drought, and irrigated conditions in Western Australia. These results differ from those found by Foulkes et al. (2007) in winter wheat under non water-stressed conditions in England.

It seems that there is a trade-off between the stem WSCCa and some of the agronomical yield components. In fact, negative correlations exist with SM2 in all the environments, but the correlations were positive with KS and TKW (**Table 3**; **Figure 5**). The PCA analyses also showed a high association between WSCCa and TKW (**Figure 7**). This negative relationship between WSC and either number of stems or number spikes per m^2 at maturity has also been reported for other wheat genotypes (Rebetzke et al., 2008a; Dreccer et al., 2009, 2013). Why genotypes with lower number of stems present higher stem WSC concentration and content? A possible explanation is that genotypes with lower number of stems per unit area have bigger stems; if fact, our results indicated a significant ($p < 0.001$) negative correlation ($r = -0.29$ and -0.36 under WS, and -0.58 and -0.56 under FI, in 2011 and 2012, respectively) between SM2 and stem weight at anthesis. Thus, genotypes with lower number of stems have probably more light transmission through the canopy and therefore higher rates of photosynthesis per stem, leading to higher stem weight and WSC content (more reserves), and greater numbers of grains per spike and kernel size. A significant and positive correlation between accumulated WSC at anthesis and kernel weight has been also observed in recombinant inbred lines (RILs) from the Seri/Babax population (Dreccer et al., 2009). Another hypothesis (complementary of the previous one) may be that those genotypes able to produce less tillers (because poorer adaptation to growing conditions—such as water stress-) are those which accumulate more carbohydrate since these photo-assimilates are not used for growth. Therefore, selecting for high stem WSC, either under near optimal agronomical conditions or under water stress, would probably lead to genotypes with lower tillering capacity and GY potential. The study conducted by Dreccer et al. (2013) in RILs of contrasting tillering and WSC concentration in the stem, and grown at different plant densities or on different sowing dates, indicates that genotypic rankings for stem WSC persisted when RILs were compared at similar stem density.

Carbon Isotope Discrimination

The genotypic differences in carbon isotope discrimination found among the 384 genotypes (**Table 3**) agree with other studies conducted in Mediterranean conditions. For example, higher Δ^{13}C (or lower carbon isotope composition, δ^{13}C) in modern cultivars compared with old varieties has been found
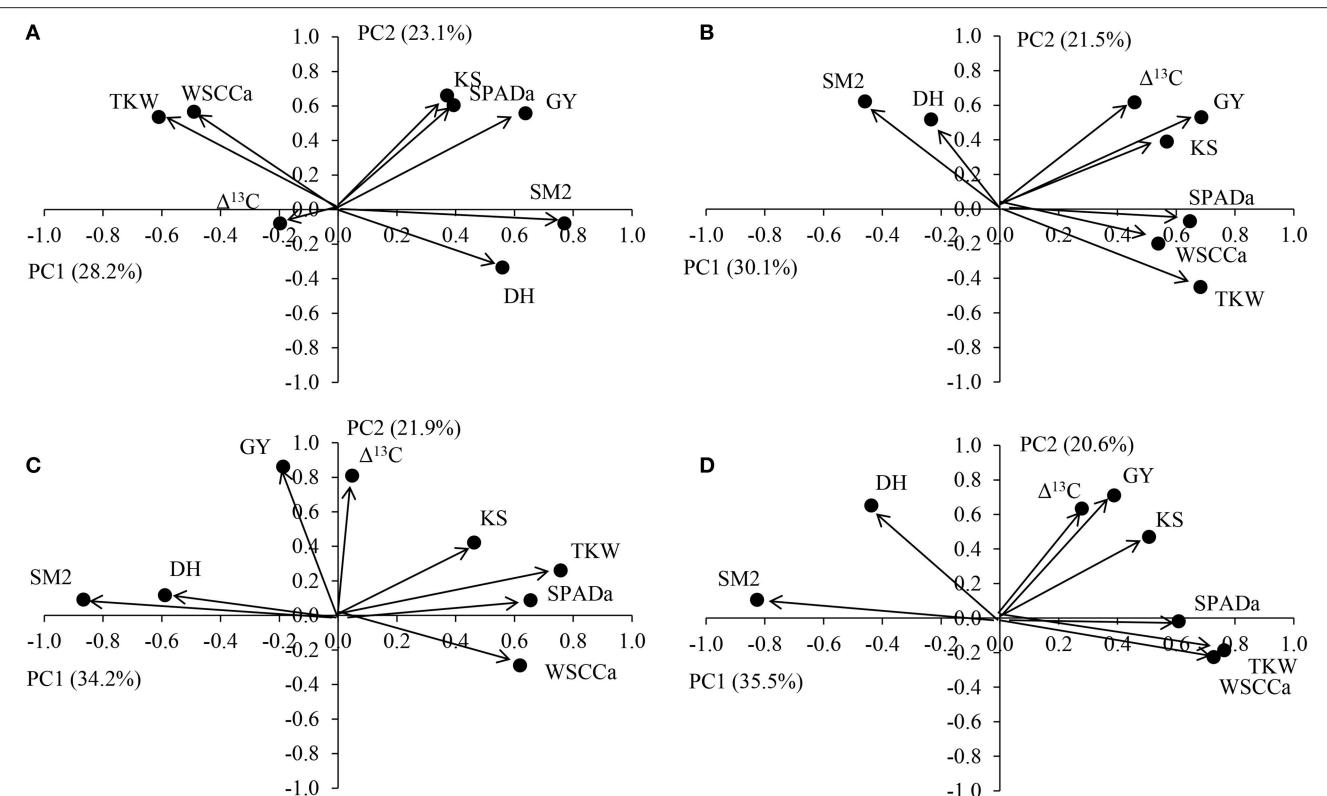
FIGURE 7 | PCA analysis of physiological and agronomical traits measured in 378 genotypes of wheat grown under severe water stress in Cauquenes (A,B) and full irrigation in Santa Rosa (C,D), in 2011 (A,C) and 2012 (B,D). Traits included in the PCA are day to heading (DH), grain yield (GY), kernel per spike (KS), thousand kernel weight (TKW), spikes per kernels m^{-2} (KM2), chlorophyll content at anthesis (SPAD), water soluble carbohydrates content at anthesis (WSCCa), and carbon isotope discrimination (Δ^{13}C) in kernels.

in bread (del Pozo et al., 2014) and durum wheats (Araus et al., 2013).

The relationship between Δ13C and GY was positive under MWS or FI but was negative under WS (**Figure 5**). Other studies in wheat (Araus et al., 2003, 2008) and barley (del Pozo et al., 2012) have also shown that Δ^{13}C in kernels can be positively or negatively correlated with GY depending on soil water availability. Positive relationships between Δ^{13}C (or negative with δ^{13}C) and GY have been frequently reported for cereals under Mediterranean conditions (see Rebetzke et al., 2008b for bread wheat and Araus et al., 2003, 2013 for durum wheat), and this can be explained by the fact that genotypes maintaining a larger transpiration and thus water use during the crop cycle will be the most productive (Araus et al., 2003, 2008, 2013; Blum, 2005, 2009). In fact, negative relationships between kernel oxygen isotope composition (δ^{18}O) or enrichment (Δ^{18}O) and grain yield have been reported in bread wheat under fully irrigated conditions (Cabrera-Bosquet et al., 2011; del Pozo et al., 2014) as well as for durum wheat under Mediterranean conditions (Araus et al., 2013) and subtropical maize under well irrigated and moderate stress (Cabrera-Bosquet et al., 2009). Indeed, carbon isotope composition can be used as a selection criterion for high water use efficiency (Condon et al., 2004; Richards, 2006), but also can provide an indirect determination of the effective water used by the crop (Araus et al., 2002, 2008; Blum, 2009). The effect of

phenology on Δ^{13}C (earlier genotypes exhibiting higher Δ^{13}C) may be discarded, since heading date was not correlated with Δ^{13}C ($P > 0.05$) in none of the environments. Actually, the positive correlations between Δ^{13}C and GY was also found when the relationship were studied within subset of 212 genotypes with similar heading duration (80–85 days); $r = 0.50$ for WS and 0.42 for WI in 2012.

CONCLUSIONS

The identification of genotypic variability for agronomical and physiological traits under water stress conditions and full irrigation is of great interest for breeders because selected genotypes with favorable traits can be used as parents in future crosses. Among these, genotypes with higher numbers of fertile tillers would lead to higher numbers of kernels per m^2 and GY under terminal water stress and non-stress conditions. Additionally, genotypes with delay in leaf senescence (a higher SPAD index) would lead to higher KS and TKW, particularly under water stress, and to a lesser extent at full irrigation. In the case of yield potential conditions, this is probably the consequence of greater amounts of fixed carbon available for grain filling, whereas under water stress stay-green it is an indicator of resilience to stress conditions. In addition, genotypes

with higher carbon discrimination values are associated with higher GY under MWS and full irrigation, indicating that more water is used by the crop. In addition, selection for a higher WSC at anthesis may bring negative consequences in terms of yield potential and adaptation to MWS conditions. This study clearly illustrates the importance of defining the target environment for wheat breeding before determining the set of phenotyping traits for selection.

AUTHOR CONTRIBUTIONS

AD and IM designed the experiments, selected the germplasm and participated on field evaluations. AY and GT were in charge of carbohydrate determinations. DC was in charge of the management of the experiments and evaluation of agronomic traits. LS and JA contributed to analysis of the data. AD was in charge of the writing up but all the authors contributed to the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00987

REFERENCES

Acevedo, E., Baginsky, C., Solar, B., and Ceccarelli, S. (1997). Discriminación isotópica de ^{13}C y su relación con el rendimiento y la eficiencia de transpiración de genotipos locales y mejorados de cebada bajo diferentes condiciones hídricas. *Inv. Agríc. (Chile)* 17, 41–54.

Araus, J. L., Amaro, T., Casadesus, J., Asbati, A., and Nachit, M. M. (1998). Relationships between ash content, carbon isotope discrimination and yield in durum wheat. *Aust. J. Plant Phys.* 25, 835–842.

Araus, J. L., Cabrera-Bosquet, L., Serret, M. D., Bort, J., and Nieto-Taladriz, M. T. (2013). Comparative performance of δ^{13}C, δ^{18}O and δ^{15}N for phenotyping durum wheat adaptation to a dryland environment. *Funct. Plant Biol.* 40, 595–608. doi: 10.1071/FP12254

Araus, J. L., Slafer, G. A., Reynolds, M. P., and Royo, C. (2002). Plant breeding in C3 cereals: what should we look for? *Ann. Bot.* 89, 925–940. doi: 10.1093/aob/mcf049

Araus, J. L., Slafer, G. A., Royo, C., and Serret, M. D. (2008). Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant Sci.* 27, 377–412. doi: 10.1080/07352680802467736

Araus, J. L., Villegas, D., Aparicio, L. F., García del Moral, S., El Hani, S., Rharrabti, Y., et al. (2003). Environmental factors determining carbon isotope discrimination and yield in durum wheat under Mediterranean conditions. *Crop Sci.* 43, 170–180. doi: 10.2135/cropsci2003.1700

Bingham, I. J., Blake, J., Foulkes, M. J., and Spink, J. (2007). Is barley yield in the UK sink limited? Post-anthesis radiation interception, radiation-use efficiency and source–sink balance. *Field Crop. Res.* 101, 198–211. doi: 10.1016/j.fcr.2006.11.005

Blake, N. K., Lanning, S. P., Martin, J. M., Sherman, J. D., and Talbert, L. E. (2007). Relationship of flag leaf characteristics to economically important traits in two spring wheat crosses. *Crop Sci.* 47, 491–496. doi: 10.2135/cropsci2006.05.0286

Blum, A. (1998). Improving wheat grain filling under stress by stem reserve mobilisation. *Euphytica* 100, 77–83. doi: 10.1023/A:1018303922482

Blum, A. (2005). Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive? *Aust. J. Agric. Res.* 56, 1159–1168. doi: 10.1071/AR05069

Blum, A. (2009). Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. *Field Crop. Res.* 112, 119–123. doi: 10.1016/j.fcr.2009.03.009

Cabrera-Bosquet, L., Albrizio, R., Nogués, S., and Araus, J. L. (2011). Dual Δ^{13}C/δ^{18}O response to water and nitrogen availability and its relationship with yield in field-grown durum wheat. *Plant Cell Environ.* 34, 418–433. doi: 10.1111/j.1365-3040.2010.02252.x

Cabrera-Bosquet, L., Sánchez, C., and Araus, J. L. (2009). How yield relates to ash content, Δ^{13}C and Δ^{18}O in maize grown under different water regimes. *Ann. Bot.* 104, 1207–1216. doi: 10.1093/aob/mcp229

Calderini, D. F., and Slafer, G. A. (1998). Changes in yield and yield stability in wheat during the 20th Century. *Field Crop. Res.* 57, 335–347.

Cattivelli, L., Rizza, F., Badeck, F. W., Mazzucotelli, E., Mastrangelo, A. M., Francia, E., et al. (2008). Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. *Field Crop. Res.* 105, 1–2. doi: 10.1016/j.fcr.2007.07.004

Chen, J. B., Liang, Y., Hu, X. Y., Wang, X. X., Tan, F. Q., Zhang, H. Q., et al. (2010). Physiological characterization of 'stay green' wheat cultivars during the grain filling stage under field growing conditions. *Acta Physiol. Plant.* 32, 875–882. doi: 10.1007/s11738-010-0475-0

Condon, A. G., Richards, R. A., Rebetzke, G. J., and Farquhar, G. D. (2004). Breeding for high water-use efficiency. *J. Exp. Bot.* 55, 2447–2460. doi: 10.1093/jxb/erh277

Coplen, T. B. (2008). "Explanatory glossary of terms used in Expression of relative isotope ratios and gas Ratios," in *IUPAC Recommendations 2008*. International Union of Pure and Applied Chemistry Inorganic Chemistry Division. Commission on Isotopic Abundances and Atomic Weights. 27. Available online at: http://media.iupac.org/reports/provisional/abstract08/coplen_prs.pdf

del Pozo, A., Castillo, D., Inostroza, L., Matus, I., Méndez, A. M., and Morcuende, R. (2012). Physiological and yield responses of recombinant chromosome substitution lines of barley to terminal drought in a Mediterranean-type environment. *Ann. Appl. Biol.* 160, 157–167. doi: 10.1111/j.1744-7348.2011.00528.x

del Pozo, A., and del Canto, P. (1999). *Áreas Agroclimáticas y Sistemas Productivos de la VII y VII región.* Instituto de Investigaciones Agropecuarias, Centro de Investigaciones Quilamapu. Serie Quilamapu N° 113.

del Pozo, A., Matus, I., Serret, M. D., and Araus, J. L. (2014). Agronomic and physiological traits associated with breeding advances of wheat under high-productive Mediterranean conditions. The case of Chile. *Env. Exp. Bot.* 103, 180–189. doi: 10.1016/j.envexpbot.2013.09.016

Denčić, S., Kastori, R., Kobiljski, B., and Duggan, B. (2000). Evaluation of grain yield and its components in wheat cultivars and landraces under near optimal and drought conditions. *Euphytica* 113, 43–52. doi: 10.1023/A:1003997700865

Dreccer, M. F., Chapman, S. C., Rattey, A. R., Neal, J., Song, Y., Christopher, J. T., et al. (2013). Developmental and growth controls of tillering and water-soluble car-bohydrate accumulation in contrasting wheat (*Triticum aestivum* L.) genotypes:can we dissect them? *J. Exp. Bot.* 64, 143–160. doi: 10.1093/jxb/ers317

Dreccer, M., Van Herwaarden, A., and Chapman, S. (2009). Grain number and grain weight in wheat lines contrasting for stem water soluble carbohydrate concentration. *Field Crop. Res.* 112, 43–54. doi: 10.1016/j.fcr.2009.02.006

Ehdaie, B., Alloush, G. A., Madore, M. A., and Waines, J. G. (2006a). Genotypic variation for stem reserves and mobilization in wheat: I. postanthesis changes in internode dry matter. *Crop Sci.* 46, 735–746. doi: 10.2135/cropsci2005.04-0033

Ehdaie, B., Alloush, G. A., Madore, M. A., and Waines, J. G. (2006b). Genotypic variation for stem reserves and mobilization in wheat: II. postanthesis changes in internode water-soluble carbohydrates. *Crop Sci.* 46, 2093–2103. doi: 10.2135/cropsci2006.01.0013

Engler, A., and del Pozo, A. (2013). Assessing long and short run trends in cereal yields: the case of Chile between 1929-2009. *Cien. Inv. Agr.* 40, 397–410. doi: 10.4067/S0718-16202013000100005

Estrada-Campuzano, G., Slafer, G. A., and Miralles, D. (2012). Differences in yield, biomass and their components between triticale and wheat grown under contrasting water and nitrogen environments. *Field Crop. Res.* 128, 167–179. doi: 10.1016/j.fcr.2012.01.003

Farquhar, G. D., Ehleringer, J. R., and Hubick, K. T. (1989). Carbon isotope discrimination and photosynthesis. *Annu. Rev. Plant Phys.* 40, 503–537.

Fischer, R. A. (2007). Understanding the physiological basis of yield potential in wheat. *J. Agric. Sci.* 145, 99–113. doi: 10.1017/S0021859607006843

Fischer, R. A., and Edmeades, G. O. (2010). Breeding and cereal yield progress. *Crop Sci.* 50, S85–S98. doi: 10.2135/cropsci2009.10.0564

Fleury, D., Jefferies, S., Kuchel, H., and Langridge, P. (2010). Genetic and genomic tools to improve drought tolerance in wheat. *J. Exp. Bot.* 61, 3211–3222. doi: 10.1093/jxb/erq152

Foulkes, M. J., Snape, J. W., Shearman, V. J., Reynolds, M. P., Gaju, O., and Sylverstar-Bradley, R. (2007). Genetic progress in yield potential in wheat: recent advances and future prospects. *J. Agric. Sci.* 145, 17–29. doi: 10.1017/S0021859607006740

García, G. A., Hasan, A. K., Puhl, L. E., Reynolds, M. P., Calderini, D. F., and Miralles, D. J. (2013). Grain yield potential strategies in an elite wheat double-haploid population grown in contrasting environments. *Crop Sci.* 53, 2577–2587. doi: 10.2135/cropsci2012.11.0669

Harris, K., Subudhi, P. K., Borrel, A., Jordan, D., Rosenow, D., Nguyen, H., et al. (2007). Sorghum stay-green QTL individually reduce post-flowering drought-induced leaf senescence. *J. Exp. Bot.* 58, 327–338. doi: 10.1093/jxb/erl225

Hawkesford, M., Araus, J., Park, R., Calderini, D., Miralles, D., Shen, T., et al. (2013). Prospects of doubling global wheat yields. *Food Energy Secur.* 2, 34–48. doi: 10.1002/fes3.15

Kirigwi, F. M., Van Ginkel, M., Brown-Guedira, G., Gill, B. S., Paulsen, G. M., and Fritz, A. K. (2007). Markers associated with a QTL for grain yield in wheat under drought. *Mol. Breeding* 20, 401–413. doi: 10.1007/s11032-007-9100-3

Lobell, D. B., Burke, M. B., Tebaldi, C., Mastrandrea, M. D., Falcon, W. P., and Naylor, R. L. (2008). Priorizing climate change adaptation needs for food security in 2030. *Science* 319, 607–610. doi: 10.1126/science.1152339

Lobell, D. B., and Gourdji, S. M. (2012). The influence of climate change on global crop productivity. *Plant Physiol.* 160, 1686–1697. doi: 10.1104/pp.112.208298

Lopes, M. S., and Reynolds, M. P. (2012). Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. *J. Exp. Bot.* 63, 3789–3798. doi: 10.1093/jxb/ers071

Matus, I., Mellado, M., Pinares, M., Madariaga, R., and del Pozo, A. (2012). Genetic progress in winter wheat cultivars released in Chile from 1920 and 2000. *Chil. J. Agr. Res.* 72, 303–308. doi: 10.4067/S0718-58392012000300001

Nouri, A., Etminan, A., Teixeira da Silva, J. A., and Mohammadi, R. (2011). Assessment of yield, yield-related traits and drought tolerance of durum wheat genotypes (*Triticum turjidum* var. *durum Desf.*). *Aust. J. Crop Sci.* 5, 8–16.

Ober, E. S., Clark, C. J. A., Le Bloa, M., Royal, A., Jaggard, K. W., and Pidgeon, J. D. (2004). Assessing the genetic resources to improve drought tolerance in sugar beet: agronomic traits of diverse genotypes under droughted and irrigated condition. *Field Crop. Res.* 90, 213–234. doi: 10.1016/j.fcr.2004.03.004

Rampino, P., Spano, G., Pataleo, S., Mita, G., Napier, J. A., Di Fonzo, N., et al. (2006). Molecular analysis of a durum wheat 'stay green' mutant: Expression pattern of photosynthesis-related genes. *J. Cereal Sci.* 43, 160–168. doi: 10.1016/j.jcs.2005.07.004

Rebetzke, G. J., Condon, A. G., Farquhar, G. D., Appels, R., and Richards, R. A. (2008b). Quantitative trait loci for carbon isotope discrimination are repeatable across environments and wheat mapping populations. *Theor. Appl. Genet.* 118, 123–137. doi: 10.1007/s00122-008-0882-4

Rebetzke, G. J., van Herwaarden, A. F., Jenkeins, C., Weiss, M., Lewis, D., Ruuska, S., et al. (2008a). Quantitative trait loci for water-soluble carbohydrates and associations with agronomic traits in wheat. *Aust. J. Exp. Agric.* 59, 891–905. doi: 10.1071/AR08067

Reynolds, M., Dreccer, F., and Trethowan, R. (2006). Drought-adaptive traits derived from wheat wild relatives and landraces. *J. Exp. Bot.* 58, 177–186. doi: 10.1093/jxb/erl250

Richards, R. A. (2006). Physiological traits used in the breeding of new cultivars for water-scarce environments. *Agric. Water Manag.* 80, 197–211. doi: 10.1016/j.agwat.2005.07.013

Richards, R. A., Rebetzke, G. J., Condon, A. G., and Herwaarden, A. F. V. (2002). Breeding opportunities for increasing the efficiency of water use and crop yield in temperate cereals. *Crop Sci.* 42, 111–121. doi: 10.2135/cropsci2002.0111

Rivero, R. M., Kojima, M., Gepstein, A., Sakakibara, H., Mittler, R., Gepstein, S., et al. (2007). Delayed leaf senescence induces extreme drought tolerance in a flowering plant. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19631–19636. doi: 10.1073/pnas.0709453104

Ruuska, S., Rebetzke, G., Van Harwaarden, A., Richards, R., Fettel, N., Tabe, L., et al. (2006). Genotypic variation in wheat. *Funct. Plant Biol.* 33, 799–809. doi: 10.1071/FP06062

Sánchez-García, M., Royo, C., Aparicio, N., Martín-Sánchez, A., and Álvaro, F. (2013). Genetic improvement of bread wheat yield and associated traits in Spain during the 20th century. *J. Agric. Sci.* 151, 105–118. doi: 10.1017/S0021859612000330

Sinclair, T. R., and Jamieson, P. D. (2006). Grain number, wheat yield, and bottling beer: an analysis. *Field Crop. Res.* 98, 60–67. doi: 10.1016/j.fcr.2005.12.006

Spano, G., Di Fonzo, N., Perrotta, C., Platani, C., Ronga, G., Lawlor, D. W., et al. (2003). Physiological characterization of 'stay green' mutants in durum wheat. *J. Exp. Bot.* 54, 1415–1420. doi: 10.1093/jxb/erg150

Tambussi, E. A., Bort, J., and Araus, J. L. (2007). Water use efficiency in C3 cereals under Mediterranean conditions: a review of physiological aspects. *Ann. Appl. Biol.* 150, 307–321. doi: 10.1111/j.1744-7348.2007.00143.x

Voltas, J., Romagosa, I., Lafarga, A., Armesto, A. P., Sombrero, A., and Araus, J. L. (1999). Genotype by environment interaction for grain yield and carbon isotope discrimination of barley in Mediterranean Spain. *Crop Pasture Sci.* 50, 1263–1271.

Wang, H., McCaig, T. N., Depauw, R. M., and Clarke, J. M. (2008). Flag Leaf physiological traits in two high-yielding Canada Western Red Spring wheat cultivars. *Can. J. Plant Sci.* 88, 35–42. doi: 10.4141/CJPS07055

Yang, D., Jing, R., Chang, X., and Li, W. (2007). Identification of quantitative trait loci and environmental interactions for accumulation and remobilization of water-soluble carbohydrates in wheat (*Triticum aestivum* L.) stems. *Genetics* 176, 571–584. doi: 10.1534/genetics.106.068361

Yemm, E. W., and Willis, A. J. (1954). The estimation of carbohydrates in plant extracts by anthrone. *Biochem. J.* 57, 508–514.

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421.

Zhang, J., Xu, Y., Chen, W., Dell, B., Vergauwen, R., Biddulph, B., et al. (2015). A wheat 1-FEH w3 variant underlies enzyme activity for stem WSC remobilization to grain under drought. *New Phytol.* 205, 293–305. doi: 10.1111/nph.13030

Zhou, Y., He, Z. H., Chen, X. M., Wang, D. S., Yan, J., Xia, X. C., et al. (2007). "Genetic improvement of wheat yield potential in North China," in *Wheat Production in Stressed Environments,* eds H. T. Buck, J. E. Nisi, and N. Salomón (New York, NY: Springer-Verlag), 583–589.

Linking Dynamic Phenotyping with Metabolite Analysis to Study Natural Variation in Drought Responses of *Brachypodium distachyon*

Lorraine H. C. Fisher[1][†], Jiwan Han[2], Fiona M. K. Corke[2], Aderemi Akinyemi[1], Thomas Didion[3], Klaus K. Nielsen[3], John H. Doonan[2], Luis A. J. Mur[1]* and Maurice Bosch[1]*

[1] Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, UK, [2] The National Plant Phenomics Centre, Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, UK, [3] DLF Seeds A/S, Store Heddinge, Denmark

Drought is an important environmental stress limiting the productivity of major crops worldwide. Understanding drought tolerance and possible mechanisms for improving drought resistance is therefore a prerequisite to develop drought-tolerant crops that produce significant yields with reduced amounts of water. *Brachypodium distachyon* (Brachypodium) is a key model species for cereals, forage grasses, and energy grasses. In this study, initial screening of a Brachypodium germplasm collection consisting of 138 different ecotypes exposed to progressive drought, highlighted the natural variation in morphology, biomass accumulation, and responses to drought stress. A core set of ten ecotypes, classified as being either tolerant, susceptible or intermediate, in response to drought stress, were exposed to mild or severe (respectively, 15 and 0% soil water content) drought stress and phenomic parameters linked to growth and color changes were assessed. When exposed to severe drought stress, phenotypic data and metabolite profiling combined with multivariate analysis revealed a remarkable consistency in separating the selected ecotypes into their different pre-defined drought tolerance groups. Increases in several metabolites, including for the phytohormones jasmonic acid and salicylic acid, and TCA-cycle intermediates, were positively correlated with biomass yield and with reduced yellow pixel counts; suggestive of delayed senescence, both key target traits for crop improvement to drought stress. While metabolite analysis also separated ecotypes into the distinct tolerance groupings after exposure to mild drought stress, similar analysis of the phenotypic data failed to do so, confirming the value of metabolomics to investigate early responses to drought stress. The results highlight the potential of combining the analyses of phenotypic and metabolic responses to identify key mechanisms and markers associated with drought tolerance in both the Brachypodium model plant as well as agronomically important crops.

Keywords: *Brachypodium distachyon*, drought, grasses, hormones, metabolite profiling, natural variation, phenotyping, stress

INTRODUCTION

Drought, the sub-optimal supply of water, is an important environmental stress that limits the productivity of major crops worldwide. Climate change models predict greater variability in rainfall patterns and increased periods of summer drought, posing serious risks for food production. The cereals, wheat, rice, and maize, provide fifty percent of human dietary energy supply and climate change is projected to negatively impact on their production (IPCC, 2014). Likewise, a decline in the productivity of cereal crops and forages will impact on animal feed supplies and therefore affect the amount and price of milk and meat available for human consumption (Wheeler and Reynolds, 2013). As the global population increases, and hence the demand for staple cereal crops, drought will have widespread implications both on the environment and related socio-economic factors, agronomy, employment, migration, food security, health, and mortality (Stanke et al., 2013). Understanding drought tolerance and possible mechanisms for improving drought resistance in agronomically important crops is therefore a prerequisite to develop drought-tolerant crops that produce significant yields with reduced amounts of water.

Drought affects morphological, physiological, biochemical, and molecular processes in plants, resulting in growth inhibition. The extent of these changes is dependent on time, stage, and intensity of the drought stress (Chaves et al., 2009). Drought tolerance is therefore considered a complex trait under polygenic control and involving complex morpho-physiological mechanisms. Progress in improving drought tolerance in cereal crops and forages has been hampered by the lack of genetic and genomic tools, their large, and often polyploid, genomes and, critically, their large physical size and relatively long life cycles. *Brachypodium distachyon* (Brachypodium) contains many of the desirable properties required for a model system for the monocotyledon grasses (Opanowicz et al., 2008) and a number of genetic and genomic tools have been developed (Mur et al., 2011), including the genome sequence of accession Bd21 (The International Brachypodium Initiative, 2010). This has positioned Brachypodium as a powerful model species to accelerate trait improvement in cereals, forages and biomass crops (Rancour et al., 2012; Slavov et al., 2013; Girin et al., 2014). The Bd21 accession has been the main focus for experiments focussing on molecular analysis of drought stress (Bertolini et al., 2013; Verelst et al., 2013). The exploitation of genetic diversity in wild relatives and diversity germplasm of crops species holds the key to improving important agronomic traits. However, the domestication of cereal and forage crops has resulted in a decrease in genetic diversity compared to their wild ancestors (Buckler et al., 2001), decreasing our capacity for trait improvement in response to environmental stress including drought. Brachypodium grows in a wide variety of habitats and was never domesticated. This provides an excellent opportunity to identify drought associated gene-trait associations and transfer this knowledge to improve drought associated agronomic traits in important food and forage crops. The natural variation in aboveground and belowground physiology upon drought induced stress has been described for Brachypodium germplasm

collections containing mainly material originating from Turkey (Luo et al., 2011; Chochois et al., 2015).

Conventional methods to phenotype drought related traits are time-consuming, labor intense, low throughput and often involve destructive harvest of plants making repeated measurements on the same plant impossible. Dynamic plant phenotyping methods enable controlled irrigation combined with automated imaging. These are particularly suited for studies on drought tolerance as these require accurate watering regimes and monitoring of responses over time to dissect the dynamic nature of drought development and the resulting stress response (Berger et al., 2010; Chen et al., 2014; Honsdorf et al., 2014).

While such phenotyping facilities have gained much interest in the context of increasing our understanding of the genetics of drought tolerance and for making gene-trait associations, the advantages it offers for metabolite–phenotype associations have yet to be explored. This is the case even though metabolite levels can be more closely linked to the macroscopic phenotype than genes as they reflect the integration of gene expression, enzyme activity, and other processes (Arbona et al., 2013) and, as such, are often used as predictive biomarkers. While metabolite profiling, as a diagnostic and predictive tool to identify novel markers for diseases, is widely used in the medical field, the application of biomarkers is not a common strategy to assist crop improvement (Steinfath et al., 2010). Drought stress leads to the accumulation of metabolites that function as osmolytes, antioxidants, or scavengers that help plants to avoid or tolerate stresses (Seki et al., 2007). Given the importance of drought induced biochemical changes, metabolic analyses provide a powerful approach to elucidate tolerance mechanisms and identify metabolic markers that may assist in developing drought-tolerant crops. For instance, a metabolomics study in oats (*Avena sativa*) has defined key processes involved in drought tolerance (Sanchez-Martin et al., 2015), while the combination of metabolite and phenotypic analysis identified correlations between certain metabolites and important fruit quality and yield traits in tomato (Schauer et al., 2006), highlighting the potential of metabolic markers for crop selection, evaluation and improvement.

In this study, we combine metabolic profiling with dynamic phenotyping of a diversity panel of Brachypodium accessions. We utilized an expanded germplasm collection of 138 Brachypodium diploid inbred lines (**Supplementary Table S1**) including ecotypes collected from a range of geographies and ecological niches from Northern Spain (Mur et al., 2011) to assess the natural variation in drought responses. Following initial drought screens, we used the facilities at the UK National Plant Phenomics Centre (NPPC-Aberystwyth) for high-resolution temporal imaging of 10 ecotypes with contrasting responses to drought stress to determine phenotypic trait associations to drought tolerance. Metabolite profiling of these samples not only informed on metabolic pathways involved in conferring drought tolerance in these selected Brachypodium ecotypes, but also enabled identification of key metabolite–phenotype associations. Our results identify several novel drought induced correlations between hormone pathways and phenotypic traits and provide a platform for improving our understanding of the

genetics underlying drought associated metabolite–phenotype correlations essential for improving drought tolerance in cereal crops and forage grasses.

MATERIALS AND METHODS

Plant Material

One hundred and thirty-eight Brachypodium diploid inbred lines were selected for the initial large drought screen, including ecotypes collected from a range of geographies and ecological niches (**Supplementary Table S1**), with 101 lines collected in Spain, 31 in Turkey, 3 in Iraq and one each from France, Italy, and Croatia.

Initial Drought Screens

For the large drought screen, six replicates of 138 Brachypodium ecotypes were sown in small pots (7.5 cm diameter, 9 cm height) in 4:1 John Innes No 1 potting compost: grit mix. The pots were placed in a glasshouse at 21–22°C with 16 h of light (natural light supplemented with artificial light from 400-W sodium lamps). Two weeks after sowing plants were fully randomized. As it was our aim to induce drought stress in actively growing plants, well before they normally would start flowering, plants were not vernalized (this also applies to the second drought screen and the detailed plant phenotyping experiment). One ecotype, Arc23, was excluded as only three out of the six replicates had germinated. For the remaining 137 ecotypes, at least four replicate plants were available (114 with six replicates, 17 with five replicates, and 6 with four replicates). All seedlings were watered equal amounts prior treatment. Those replicates (2–3 plants per ecotype) to be exposed to water deficit were raised on circular blocks, approximately 1.5 cm in height, to prevent contact with any water draining through the soil of well-watered control plants. Total water withdrawal began 28 days after sowing and lasted 6 days. Leaf wilting was scored by two individuals according to a 1–6 scale for this trait (1 = no effect, 6 = severe effect) based on the visual assessment categories described by Engelbrecht et al. (2007). Plants were weighed immediately after harvest of the above ground biomass.

A second, refined drought screen was performed on 48 selected Brachypodium ecotypes (see **Supplementary Table S1**). Seedlings were germinated in square pots (9 cm × 9 cm × 7.5 cm), containing the same soil mix as before, in replicates of 12 (six controls and six to be stressed). Plants were grown under controlled environment conditions in a growth chamber at 21°C and under a 16 h photoperiod with 176 µmol m^{-2}s^{-1} photon flux density supplied by white fluorescent tubes (OSRAM, Garching, Germany). All plants germinating within 1–2 days of each other (with the exceptions of ecotypes Kah6 and Uni14, which developed approximately 5 days later). Plants were randomized in a block design with buffer plants placed on outskirts to reduce changes in stress resulting from interactions with microclimate variations. Plants were all given equal volumes of water and water levels were monitored throughout, averaging at 0.25 m^3 m^{-3} prior stress treatment and 0.29 m^3 m^{-3} in soils of control plants during the

stress treatment period. Total water withdrawal of six replicates per ecotype started 33 days after sowing and lasted for 12 days. The average soil moisture content upon drought treatment dropped to 0.024 m^3 m^{-3} on day 42 while at the end of the experiment on day 45 no reading could be obtained.

Leaf wilting observations were taken after 12 days of drought treatment, using the same criteria as used in the large screen. On day 45, plants were harvested and fresh weight and dry weight biomass recorded; the latter after drying at 70°C for 3 days. Fresh and dry weight (DW) figures were then used to calculate above ground plant water content (PWC) on a fresh weight (FW) basis using the following equation: PWC (%) = [(FW$_{(g)}$ − DW$_{(g)}$)/FW$_{(g)}$] × 100%.

Phenotyping of Drought Induced Physiological Changes in Selected *Brachypodium* Ecotypes

Ten Brachypodium ecotypes were selected for detailed dynamic plant phenotyping using the facilities of the National Plant Phenomics Centre (NPPC) at Aberystwyth University. Seedlings were grown in 7.5 cm square pots, each containing 225 ± 0.5 g 4:1 Levington F2: grit sand in a regular glasshouse (for conditions see large drought screen). Twenty-eight days after sowing, the plants were transferred to the NPPC Smarthouse. Four plants of each ecotype were placed in a single tray lined with blue germination paper to ensure even distribution of water between the pots. Dividers were placed between the plants to enable imaging of individual plants, each of the trays was placed on a cart of the conveyer belt system (**Figure 2A**). All plants were watered at calculated 75% field capacity for an additional 2 days with the start of the treatments 33 days after sowing (das). Treatments were comprised of two levels of drought stress: 24 plants for each ecotype (six trays/carts) exposed to 15% soil water content (SWC), another set of 24 plants were not watered at all (0% SWC). Twelve control plants per ecotype (three trays/carts) continued to be watered to 75% SWC. Weighing and watering of the plants was fully automated. Samples of the compost were soaked to determine field capacity or dried to determine dry matter content. The water content was the difference between these two values. Target watering weight was calculated based upon a percentage of the water content. Plants from a parallel experiment were weighed at maximum size and this showed that plant fresh weight accounted for around 2% of the daily water usage, so no adjustment for plant biomass was deemed necessary. The different treatments lasted for 12 days with the aboveground biomass of four plants (one tray) per ecotype being harvested after 12 days for each of the treatments. Growth conditions in the Smarthouse were 22°C 16 h day/20°C night with supplementary top-up lighting. Watering was calculated based on the field capacity and dry matter content of the compost. Automated, target weight based watering was provided on a daily basis. Imaging was also performed daily starting from imposition of drought treatment (34 das) of all plants until the last day (45 das) of the treatments.

The timing for the initiation of the treatments for the two drought screens and for the phenotyping experiment (28, 33,

and 33 das, respectively), as well as for the end-point of the treatments (34, 45, and 45 das, respectively), was based on our observations that Bd21, which does not require vernalization, started flowering ∼45 das. The chosen time-points therefore represented a reasonable equivalent for all ecotypes to be at a pre-flowering stage had they been vernalized.

Feature Extraction

Images of each tray were taken from four side view angles with an interval of 90° and one top view once a day. The camera used to image the plants exports 24 bit RGB color images, i.e., each channel has 256 class color levels. **Figure 2B** shows the position of four plants in each tray and the images acquired from side view and top view cameras. Images were processed to segment the plant from the background and to extract plant height, top view and side view projection area and color information. In order to acquire relatively uniform processing results for further analysis, images were filtered by a retina filter (Benoit et al., 2010), which enhances the contrast between plant and background, improves color consistency and provides spatial noise removal and luminance correction. Each side view image contains two individual plants, which are separated by a panel. Therefore after this pre-processing step, regions of interests (ROI) containing only one single plant were acquired with the bottom position of each ROI fixed just beyond the top of the pot from where the plant height and side view area are measured. **Figure 2C** illustrates the image processing on a side view image. In this image, the ROIs are fixed and height and total pixels (plant area) are acquired for each individual.

Within each ROI, plants were segmented according to the pixel colors in RGB color space. Although the pre-processing improved image quality and removed some noise, some small background pixel patches (clustering around 10–20 pixels) with the same colors of a plant, were removed by morphological operations.

Four features were extracted directly from the side view images of segmented plants: height, total pixels, yellow pixels, and grey pixels. The yellow color is defined as a color range where the red channel value is 10 levels higher than that of the green channel. The grey color range is defined by the following formula in which V denotes the value:

$$\frac{V_{Green}}{V_{Blue}} < 2 \text{ and } \frac{V_{Green}}{V_{Blue}} > 1.15 \text{ and } \frac{V_{Green}}{V_{Red}} < 1.4 \text{ and } V_{Green} < 150$$

The top view image processing extracted total pixel numbers for each of the four plants in a tray separated by two panels. For top view image processing, the color description of the plants needed to be more complex to extract as much as possible of the plant pixels while minimizing noise. **Figure 2D** shows the top view segmentation and pixel count. The colors of the plants were defined as:

$$V_{Green} < 200 \text{ and } V_{Green} > 20 \text{ and } V_{Red} > 20 \text{ and } V_{Red} < 200$$

$$\text{and } (V_{Green} - V_{Blue}) > 12 \text{ and } \frac{V_{Red}}{V_{Green}} < 1.5$$

$$\text{and } \frac{V_{Green}}{V_{Blue}} > 1.5$$

All extracted values for height, estimated plant area and color information were written into a csv file for further analysis.

Image processing and feature extraction was achieved using C++ and OpenCV, an open source computer vision library[1].

Stomatal Conductance

Stomatal conductance was measured for 9 of the selected Brachypodium ecotypes (Bd21 was excluded as it started to flower during the course of the drought experiment) over a period of 7 days following drought treatment as well as for well-watered control plants at 5 weeks after sowing. Plants were grown under a 12 h light period (8 am – 4 pm) at 250 μmol m^{-2} s^{-1} photon flux density supplied by white fluorescent tubes (OSRAM) in a growth chamber at 20°C. Measurements of stomatal conductance were made on the second fully emerged true leaf between 12 noon and 1 pm using an AP4 porometer (Delta-T devices Ltd, Cambridge, UK).

Metabolite Profiling

For each treatment (0%, 15%, and 75% SWC), leaf samples harvested 12 days after initiation of drought treatments, were pooled for each of the ecotypes and metabolites extracted following the procedure described by Allwood et al. (2006). Glass vials were capped and analyzed in random order on a LTQ linear ion trap (Thermo Electron Corporation). Data were acquired in alternating positive and negative ionization modes over four scan ranges (15–110, 100–220, 210–510, and 500–1200 m/z), with an acquisition time of 5 min. Discriminatory metabolites were selected and tentatively identified by following statistical analyses and interrogation of KEGG: Kyoto Encyclopedia of Genes and Genomes[2] and MZedDB[3]. To substantiate these identifications, nominal mass signals were investigated further by targeted nano-flow Fourier Transform-Ion Cyclotron Resonance Ultra-Mass-Spectrometry (FT-ICR-MS) using TriVersa NanoMate (Advion BioSciences Ltd) on a LTQ-FT-ULTRA (Thermo Scientific) to obtain ultra-high accurate mass information and MSn ion-trees. Based on an accuracy of 1 ppm for the FT-ICR-MS, the top ranking metabolite with this range was indicated as the identification for each discriminatory negative ionization mode flow injection electrospray (FIE)-MS metabolite.

Statistical Analysis

FIE-MS data was normalized with the total ion count for each sample used to transform the intensity value for each metabolite in to a percentage of the total ion count, after the removal of metabolites below 50 m/z. ANOVA, Pearsons correlation analyses, Principal Component Analyses (PCA), and Hierarchical Cluster Analyses (HCA) were completed using the R-based MetaboAnalyst 2.0 interface (Xia et al., 2015).

---

[1]http://www.opencv.org/

[2]http://www.genome.jp/kegg/

[3]http://maltese.dbs.aber.ac.uk:8888/hrmet/index.html

RESULTS

Brachypodium Drought Screens

Two progressive drought screens were conducted with the aim to identify a core set of Brachypodium ecotypes for further detailed analysis. The initial large drought screen of 138 different Brachypodium ecotypes highlighted the considerable variation in the size and stature between the ecotypes, both in well-watered and drought stressed conditions. Wilting scores are the most widely used indicator for plant drought stress and have been shown to allow for robust ranking of survival when exposed to drought (Engelbrecht et al., 2007). Hence, visual assessment of plant wilting (**Supplementary Table S2**) formed the basis for the selection of those ecotypes to be included for a second, more refined drought screen. Other considerations taken into account for further selection were based on inclusion of ecotypes for resequencing[4], available seed stock and developmental aspects.

A total of 48 different ecotypes were selected for the second drought screen, featuring bigger pots and controlled environment conditions as well as an increased number of replicates compared to the first screen. **Figure 1A** shows the distribution of the above ground plant water content (PWC) figures for the 48 ecotypes after withholding water for 12 days. During this period, the soil moisture content dropped from an average of ~ 0.25 m^3 m^{-3} before the treatment to ~ 0.024 m^3 m^{-3} 9 days after withholding water, while 12 days after treatment no readings could be obtained anymore. PWC ranged from just below 30% for Koz1 to almost 80% for ABR8. The ranking of the PWC was used as the main criterion for the selection of 10 ecotypes for further detailed analysis. Ecotypes ABR8, Pal6, and Mur3 were selected as drought resistant ecotypes as they showed the highest PWC upon drought treatment (**Figure 1A**). These ecotypes also showed a low wilting score in the second screen (**Figure 1B**). Koz1, Gal10, Per3 and ABR3 were selected as drought-susceptible ecotypes, since they featured amongst the lowest PWC scores and ranked highly in the wilting scores (**Figures 1A,B**). ABR4, Bd21, and Luc21 were included as having an intermediate (INT) response. As expected, comparison between the PWC data (**Figure 1A**) with the wilting scores (**Figure 1B**) shows that there is a strong negative correlation between these two measures ($r = -0.869$).

Phenotypic Analyses of 10 Selected Brachypodium Lines

Having selected 10 Brachypodium lines that were either classified as tolerant (TOL), INT, or susceptible (SUS) to the drought conditions imposed during the preceding drought screens (**Figure 1**); we next performed an integrated phenotype and metabolite analysis. Plants were exposed to two levels of drought targeting 15 and 0% SWC over a period of 12 days. RGB images were acquired daily from side views at different angles and top view and phenotypic features (height, area, and color) were extracted from the plant images (**Figure 2**). Plant area, as estimated from side view images, provides a proxy for growth

---

[4]http://jgi.doe.gov/our-science/science-programs/plant-genomics/brachypodium/

(Neilson et al., 2015) and response to drought. An example of side-area over time for the 10 ecotypes is shown in **Figure 3**. Tolerant genotypes (blue lines) show higher accumulation of side area under water stress when compared to SUS (green) and INT (red) genotypes. Under well-watered control conditions there is no obvious distinction in side area accumulation according to the pre-classified tolerance groups (**Figure 3**). It should be noted that Bd21, included as an INT ecotype, started to flower during the course of the drought experiment and, therefore, was excluded from further phenotypic and metabolomic data analyses. To allow balanced multivariate analyses of our data, ABR3 was re-classified (as a line with borderline phenotype) and moved from the SUS to the INT drought tolerance group. The targeted genotypes were also assessed for stomatal performance under conditions tending toward 0% SWC (**Supplementary Figure S1**). Examination of stomatal performance under drought with each genotype and phenotypic class suggested that none exhibited significantly ($P = 0.982$) different responses. Thus, differential drought tolerance across the Brachypodium genotypes did not arise from stomatal effects.

Unsupervised PCA was applied to the phenotypic data obtained at day 12 for the plants exposed to 15 and 0% SWC and compared to the fully watered 75% SWC controls. Data obtained from different genotypes were classified as TOL, INT, or SUS phenotypes based on our previous screens (**Figure 1**), and re-classification of ABR3 as mentioned before. PCA of phenotypic data obtained for 75% SWC controls suggest that most genotypes were morphologically similar with the possible exception of two TOL lines; Pal6 and ABR8 which lay outside the 95% confidence interval circles (**Figure 4A**). However, further analyses of these genotypes using ANOVA suggested that their most prominent feature ("area-side") was not significantly different ($P = 0.41$) from the other genotypes. With imposition of drought to 15% SWC, phenotypic features were again not discriminating between the SUS and INT groups but were discriminating with two TOL genotypes Pal6 and ABR8 across PC2 (**Figure 4B**). ANOVA suggested that these two and indeed Mur3 were significantly taller ($P = 0.04$) as also indicated from box and whisker plots for plant height (**Figure 4B** inset). Considering phenotypic data for plants exposed to 0% SWC, clear differentiation was observed between TOL and SUS genotypes (**Figure 4C**). This aspect was also observed using HCA where each phenotypic group were clustered (**Figure 4D**). HCA and ANOVA (**Figure 4E**) demonstrated that drought tolerance was associated with significantly increased height and 'area side' but with reduced grey and yellow pixels.

Metabolite Analyses Indicate Differential Responses to Drought within Each Phenotypic Group

Metabolite analyses based on FIE-MS were performed on leaf samples collected 12 days after initiation of the two drought treatments and compared to well-watered controls. Images of the 9 genotypes recorded immediately before harvest for metabolite analysis are shown in **Supplementary Figure S2**. Non-supervised PCA of the well-watered control samples revealed no distinct

FIGURE 1 | Assessment of drought stress for selected Brachypodium ecotypes exposed to water deficit. Plant water content (PWC) measures are ranked from high to low **(A)** and wilting scores from low (no sign of wilting) to high **(B)** for the 48 ecotypes included in the screen. Arrows indicate the ecotypes selected for further detailed analysis; dashed lines are included to compare the relative rankings of the selected ecotypes between **(A)** and **(B)**. Both PWC and wilting score data are based on six biological replicates, except for Foz8, Sar24, and Tek3 (five replicates) and Kah6 (four replicates). Origin: dark grey = Spain, light grey = Turkey, black bars = Other. Horizontal bars: PWC > 70% (Blue) = ecotypes classified as tolerant (TOL); PWC 55–70% (red) = ecotypes classified as intermediate (INT); PWC < 55% (green) = ecotypes classified as susceptible (SUS). Error bars indicate standard deviation.

FIGURE 2 | Plant Phenotyping and image processing. Setup of high-throughput phenotyping system. A total of 600 Brachypodium plants (10 genotypes) where phenotyped over 12 days **(A)**. Each tray holds four replicate individuals. The tray will be rotated four times with an angle of 90°. Therefore, each individual plant can have two side views from two orthogonal angles. The top view image is taken when the tray is at 0° **(B)**. The image processing is to segment the plant from background, such as the panels and wall. The program automatically finds the central separating panel to setup two regions of interests (ROI). The measurement happens in each ROI. The program copes with various lighting conditions and image quality **(C)**. The top view image processing finds the panels and uses them to separate those four plants in the tray. Then the total pixels of each plant are acquired. These plants are denoted by different colors **(D)**.

FIGURE 3 | Normalized shoot area curves. Shoot areas for each of the 10 ecotypes for each treatment (0, 15, and 75% SWC) as extracted from RGB side images of the plants from day 34 to day 45 after sowing. Data was normalized against the values at 34 das (days after sowing). Error bars are not included to enable visual identification of the individual ecotypes (a list of normalized data including standard deviations are included in **Supplementary Table S3**. Blue colored lines = ecotypes classified as TOL; Red colored lines = ecotypes classified as INT; Green colored lines = ecotypes classified as susceptible (SUS).

clusters in the metabolite profiles of the three pre-defined phenotypic groups (**Figure 5A**). However, distinct clusters linked to relative drought tolerance emerge for the metabolites extracted from plants exposed to 15% SWC (**Figure 5B**), in particular for the TOL ecotypes. More severe drought to 0% SWC resulted in three clearly distinct clusters corresponding to TOL, SUS, and INT phenotypes (**Figure 5C**). The major source of variation were extracted for treatments to 15 and 0% SWC based on PCA loading vectors and significant differences as identified using ANOVA. These were tentatively identified by database interrogating based on high-resolution MS of the targeted *m/z*. For the 15% SWC samples the major sources of variation were tentatively associated with tricarboxylic acid (TCA) cycle intermediates malate and citrate, the phospholipid-derived hormone jasmonate, and unidentified metabolites with *m/z* of 236.01 and 386.9. HCA suggested that these metabolites alone could discriminate between the TOL and other groups at 15% SWC (**Figure 5D**). All with the exception of 236.01 *m/z* were relatively increased in the TOL grouping. Upon more severe drought (0% SWC) more metabolites were targeted which again could discriminate the TOL genotypes from the others (**Figure 5E**). These metabolites included jasmonate as well as other plant hormones gibberellin (GA17) and salicylate, metabolites associated with amino acid metabolism (aspartate, glutamate), chorismate, and lipid metabolism (caprylate). In each case, accumulation of the metabolites was relatively increased in the TOL lines.

To further highlight the possible importance of the pathways, *m/z* corresponding to metabolites within the pathways were extracted (**Figure 6**). In the case of the TCA cycle (**Figure 6A**), the putative TCA metabolites allowed the separate clustering of the TOL genotypes from other phenotypic classes. These data, as well as box and whisker plots, suggested significant increases in the TCA cycle in TOL genotypes. The TCA intermediates 2-oxoglutarate and oxaloacetate also contribute carbon skeletons for amino acid biosynthesis and examination of *m/z* tentatively linked to amino acid, suggested that lower levels of amino acid accumulation might be linked to a SUS phenotype (**Figure 6B**).

Indeed, glutamate, glutamine, and the drought responsive amino acid proline were significantly elevated in both TOL and INT phenotypes. With alanine, significant increases were only associated with the TOL phenotypes.

Considering *m/z* tentatively linked with the biosynthesis of jasmonate or salicylate, HCA separated clustered TOL genotypes (**Figures 6C,D**). The TOL phenotypes exhibited significant increases in key oxylipins, linolenic acid, and the active jasmonate hormones, jasmonic acid (JA) and jasmonate-isoleucine (JA-Ile) (**Figure 6C**). With the salicylate pathway only salicylic acid (SA) itself (**Figure 6D**) was significantly increased in TOL phenotypes.

Correlation between Metabolite and Phenotypic Traits

Since both the phenotypic data and the metabolite analyses reveal distinct clusters that match the designation of the different drought tolerance groups established during the initial screening, we wished to determine potential correlation between the metabolite pathway data and phenotypic variables. To assess this, a new matrix was derived which incorporated both metabolite and phenotypic data. Given the very different types of data, data was mean-centered and divided by the standard deviation of each variable. The data from the resulting matrix were established to be normally distributed.

The outcome of multivariate correlation analyses (based on Pearson's *r* as a distance measure) is provided in **Supplementary Figure S3**. These analyses suggested that the putative jasmonate pathway correlated only with itself and no phenotypic measure. However, a key cluster (arrowed in **Supplementary Figure S3**) suggested positive correlations with phenotypic and metabolite indicators and a negative correlation with "Grey"/"Yellow" pixel counts.

To facilitate visualization of these trends, key variables in this cluster were extracted and separately analyzed (**Figure 7**). This highlighted that plant area ("top" and "side") and height positively correlated with the relative concentrations

FIGURE 4 | Phenotypic descriptions of drought responses by Brachypodium accessions. Principal Component Analyses of phenotypic data derived for conditions of **(A)** 75%, **(B)** 15%, and **(C)** 0% SWC. Data points are annotated according to genotypes and drought tolerance classes TOL (blue) = tolerant (Pal = Pal6; A8 = ABR8; Mur = Mur3); INT (red) = intermediate (Luc = Luc21; A3 = ABR3; A4 = ABR4) and SUS (green) = susceptible (Koz = Koz1; Gal = Gal10; Per = Per3). Larger circles indicate 95% confidence intervals for each phenotypic group that are appropriately color-coded. Inset within B is a box and whisker projection of the only statistically significant variable for 15% SWC conditions; "area side." **(D)** Hierarchical cluster analyses (HCA) and **(E)** box and whisker projections of phenotypic data derived from 0% SWC condition. Levels of significance are indicated for each phenotypic character.

of TCA intermediates, alanine and the hormone salicylate. This would suggest that tolerance in these Brachypodium genotypes was related to the accumulation of these metabolites. Linked to this, negative correlations were observed with grey and yellow pixel contents- features linked to susceptibility to drought.

DISCUSSION

The Poaceae (or grasses), including cereals and grasses of grasslands and pastures, constitute the major source of dietary calories for human and livestock, are increasingly important as lignocellulosic biomass for biofuels, and define many natural ecosystems and agricultural landscapes. Drought is an important environmental factor limiting the productivity of crops worldwide. The greater the diversity within a species, the

better the potential for the species to overcome and adapt to changes in the environment, such as drought. Brachypodium displays considerable phenotypic variation across its geographic range, and is a key model species for cereals, forage grasses and energy grasses (Brkljacic et al., 2011). Its extensive natural variation encompasses many traits of agronomic or adaptive significance and, together with genetic and genomic tools available, is therefore well positioned for improving agronomic traits.

Screening and Ecotype Selection

To assess the natural variation to drought stress and select ecotypes for further detailed studies, we performed two successive screens in which, respectively, 138 and 48 diploid Brachypodium ecotypes, sourced from different geographical locations, were exposed to progressive drought stress. These screens

FIGURE 5 | Metabolomic responses to drought by Brachypodium accessions. Principal Component Analyses of metabolomic data derived for conditions of **(A)** 75%, **(B)** 15%, and **(C)** 0% SWC. Data points are annotated according to genotypes and drought tolerance classes TOL (blue) = tolerant (Pal = Pal6; A8 = ABR8; Mur = Mur3); INT (red) = intermediate (Luc = Luc21; A3 = ABR3; A4 = ABR4) and SUS (green) = susceptible (Koz = Koz1; Gal = Gal10; Per = Per3). Larger circles indicate 95 % confidence intervals for each phenotypic group that are appropriately color-coded. Major source of variation in the datasets as defined by ANOVA for **(D)** 15% and **(E)** 0% SWC.

demonstrated the natural variation among Brachypodium ecotypes, in terms of morphology, biomass accumulation and response to drought stress. For instance, final dry weight biomass of the watered control replicates ranged from 0.032 to 0.317 g in the first screen and from 0.183 to 0.318 g in the second screen (data not shown). This was in line with the large phenotypic variation in response to drought stress observed amongst 57 Brachypodium ecotypes that mostly originated from Turkey (Luo et al., 2011).

When assessed for wilting, leaf water content (LWC) and chlorophyll fluorescence (F_v/F_m), these Brachypodium ecotypes were classified into four different groups ranging from TOL to most-SUS to drought stress (Luo et al., 2011). Assessment of genotypic variation has suggested greater genetic diversity in wild Brachypodium individuals from the Western Mediterranean region compared to those from the Eastern (Mur et al., 2011). Here we screened geographically more diverse ecotypes (101 lines collected in Spain, 31 in Turkey, 3 in Iraq, and 1 each from France, Italy, and Croatia), with our second screen only having six ecotypes (Adi10, Bd21, Koz1, Gaz8, Kah6, and Bd2-3) in common with the study by Luo et al. (2011).

Ecotypes were tentatively classified as TOL when exhibiting a PWC higher than 70% and SUS with a PWC lower than 55% with the ecotypes in-between classified as INT (**Figure 1A**). Brachypodium is found in many climate zones throughout temperate regions, spanning habitats near sea level to over 1800 m altitude (Des Marais and Juenger, 2016), suggesting that populations are locally adapted. It was therefore anticipated that similarly classified ecotypes might share similarities with regards to ecological niche. However, based on highland/lowland descriptions or altitude, such shared responses were not apparent in our screens. For example, two out of the three selected drought resistant ecotypes (Mur3 and Pal6) were from the highlands in Spain (ABR8 was from a hillside origin close to Siena in Italy), as were three of the four selected drought-susceptible ecotypes (ABR3, Per3 and Gal10; Koz1 is from the highlands in Turkey).
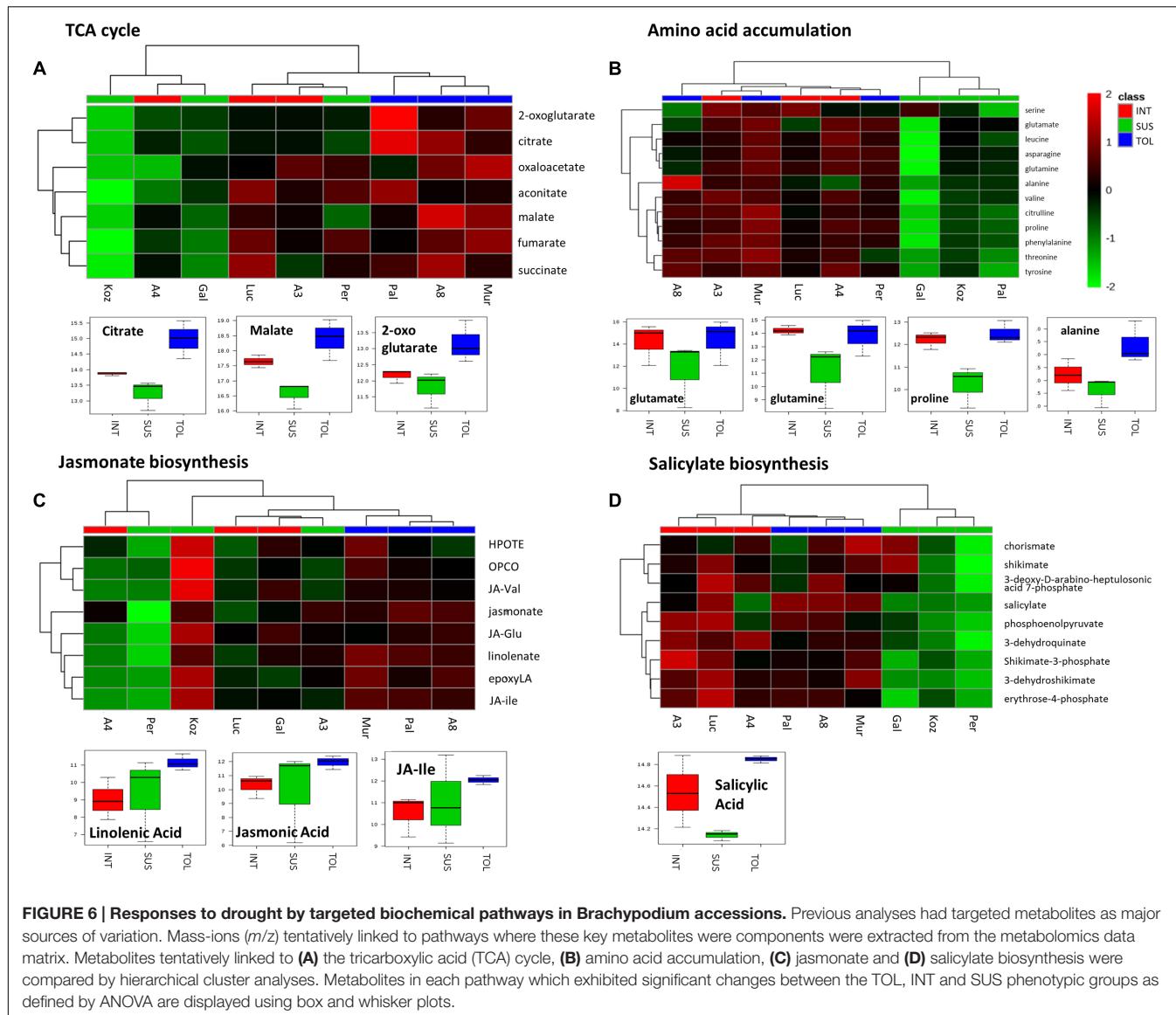
FIGURE 6 | Responses to drought by targeted biochemical pathways in Brachypodium accessions. Previous analyses had targeted metabolites as major sources of variation. Mass-ions (*m/z*) tentatively linked to pathways where these key metabolites were components were extracted from the metabolomics data matrix. Metabolites tentatively linked to **(A)** the tricarboxylic acid (TCA) cycle, **(B)** amino acid accumulation, **(C)** jasmonate and **(D)** salicylate biosynthesis were compared by hierarchical cluster analyses. Metabolites in each pathway which exhibited significant changes between the TOL, INT and SUS phenotypic groups as defined by ANOVA are displayed using box and whisker plots.

A study looking at root phenotypes amongst 81 Brachypodium accessions also found no correlation between the phenotypic diversity and geographical origins (Chochois et al., 2015). This lack of correlation may suggest that the phenotypic plasticity of Brachypodium to changing environmental conditions (i.e., acclimation responses, Des Marais and Juenger, 2016) prevails over ecotypic differentiation (local adaptation, or within-species niche divergence leading to ecotypic differentiation, Liancourt and Tielbörger, 2009). It should be noted, however, that other features associated with ecological niches, such as soil type, acidity, and climate may still attribute to a shared drought response, but these parameters were not identified in this study.

Phenotypical Characteristics of the Different Tolerance Groups

Maintaining plant growth and yield under drought represents a major objective for plant breeding. Growth, caused by cell division and cell elongation, is reduced under drought owing to impaired enzyme activities, loss of turgor, and decreased energy supply (Farooq et al., 2012). Most studies on drought are performed by withholding water, leading to progressive drought stress. However, under natural conditions plants are often exposed to moderate drought stress, in particular in temperate climates. As such, enhanced survival of *Arabidopsis thaliana* (Arabidopsis) under severe drought was shown not to be a good indicator for improved growth performance under mild drought conditions (Skirycz et al., 2011). In addition to progressive drought (0% SWC), we therefore exposed Brachypodium ecotypes to mild drought stress (15% SWC), utilizing the gravimetric watering control of the NPPC. Image based projected shoot area allows estimation of above ground biomass and, when assessed over time, can serve as a useful proxy for overall plant growth (Honsdorf et al., 2014; Neilson et al., 2015). On average, based on projected shoot area, exposure to moderate drought

FIGURE 7 | Correlation analyses of phenotypic and metabolomic data.
Metabolites tentatively linked to four key pathways as targeted by unbiased
chemometric analyses were correlated with phenotypic datasets
(**Supplementary Figure S3**), based on which key relationships were
identified. These form the basis of this Pearson's correlation analyses the
strength of which are indicated based on the intensity of red (positive) or green
(negative) color. r^2 values are provided in **Supplementary Table S4.**

stress resulted in 85% more biomass accumulation during the
treatment compared to exposure to severe stress (ranging from
43% for Pal6 to 211% for Gal10). Similarly, moderate drought
imposed an average yield penalty of 38% (ranging from 28% for
Mur3 to 46% for Gal10) when compared to well-watered controls.
Results suggest that under mild drought stress, assessment of
plant height, another parameter used to estimate plant biomass in
several crop species (Tilly et al., 2015), could be a good predictor
for ecotypes that perform well under more severe drought stress.
However, the phenotypic features measured after exposure to
moderate drought were unable to separate the Brachypodium
ecotypes in the TOL, INT, and SUS groups.

In contrast to mild drought, PCA on the phenotypic data after
progressive drought showed a clear differentiation between TOL
and SUS genotypes. The TOL group was not only characterized
by having significantly increased measures for height and side
area, but also showed distinct color-related properties with
significantly reduced proportions of yellow and grey pixels.
While leaf 'greenness' has been shown to correlate with foliar
nitrogen and chlorophyll, the relative proportion of yellow pixels
is indicative of leaf senescence characterized by yellowing or
chlorosis (Rajendran et al., 2009; Li et al., 2014; Neilson et al.,
2015). Interestingly, we also identified a "grey" pixel color
range that reflected the relationship between greenness and red
and blue color pixels. The physical basis of this trait was not
determined but could be due to altered pigmentation or water
content. Thus, our results suggest reduced or delayed drought

induced senescence in the TOL lines. Overall, our phenotypic
data analyses upon exposure to progressive drought confirm the
classification of the selected ecotypes into TOL, SUS, and INT,
and therefore the robustness of the preceding drought screens.

Metabolites Associated with Drought Stress

The interaction between genotype and environment is complex.
The ability for metabolites to integrate these two components
reflects an increasing tendency to use metabolites as selection
markers in crop breeding programs to accelerate the development
of improved cultivars tolerant to drought (Sanchez-Martin
et al., 2015). Metabolites with a higher or indeed lower relative
accumulation in the TOL lines when exposed to drought
not only provide insights into the regulation of metabolic
networks under drought stress, but could also lead to metabolite
marker development. In this study, we used our well-established
metabolomic analyses pipeline (e.g., Lloyd et al., 2011) to mine
high-resolution FIE-MS datasets using multivariate approaches
to identify major sources of variation in the experimental
parameters. Through interrogation of the loading vectors linked
to PCA, coupled with ANOVA of the datasets, we identified
metabolites which exhibited differential accumulation in the TOL
class in response to 0% SWC. Database interrogation with the
highly resolved (to 1ppm) m/z suggested that these metabolites
included three phytohormones, metabolites associated with
amino acid metabolism and putative TCA cycle metabolites.

Phytohormones play critical roles in regulating plant
responses to stress. Expression profiling of markers for
defense-related phytohormones showed that Brachypodium has
phytohormone responses more similar to those of rice than of
Arabidopsis, suggesting that monocots share a common defense
system that is different from that of dicots (Kouzai et al., 2016).
Our analyses identified relative higher levels of GA, SA, and JA
in the TOL ecotypes when exposed to severe drought with the
latter also higher in TOL lines under mild drought. A central
role for the GA class of growth hormones in the response to
abiotic stress is becoming increasingly evident (Colebrook et al.,
2014). For instance exogenous application of GA can alleviate
drought-imposed adverse effects in maize (Akter et al., 2014),
although application of GA has also been shown to increase
shoot height of dwarf Barley lines, negating the increased stress
tolerance exhibited by the dwarf plants (Vettakkorumakankav
et al., 1999). SA accumulation has been reported to improve
drought tolerance in Arabidopsis by inducing stomatal closure
(Khokon et al., 2011) and inhibiting stomatal opening (Okuma
et al., 2014). Similarly, in oats, tolerance to drought has been
at least partially associated with the accumulation of SA, again
by influencing stomatal opening (Sanchez-Martin et al., 2015).
However, like for GA, the effect of SA on drought tolerance
is complex and others have reported a reduction of drought
tolerance by SA application (Miura and Tada, 2014). Since
no differential effect on stomatal closure was observed upon
progressive drought between the Brachypodium ecotypes, SA
may be involved in alternative drought-induced regulatory
mechanisms.

An interesting question is how far SA could be influencing the two primary metabolism pathways targeted in our study; the TCA cycle and amino acid metabolism. The TCA cycle is a crucial component of respiratory metabolism and is often altered in plants experiencing stress. Drought induced accumulation of TCA cycle metabolites (Urano et al., 2009) and the upregulation of TCA cycle-related genes in Arabidopsis shoots (Cavalcanti et al., 2014), have been reported. Metabolic differences in the stress tolerance of four different lentil genotypes were related to a reduction in the levels of TCA cycle intermediates suggesting an impaired energy metabolism with consequences on the ability of seedlings to acquire water and to support transport processes (Muscolo et al., 2015). SA has a well-characterized role in maintaining mitochondrial electron flow under stress conditions by inducing the expression of alternative oxidase (AOX) (Feng et al., 2009). This will influence the NADH oxidation by mitochondrial complex I which is coupled to the TCA cycle (Vanlerberghe, 2013). Thus, TOL Brachypodium genotypes could be exhibiting a SA-AOX mechanism of maintaining bioenergetic metabolism during drought.

Amino acid accumulation, in particular proline, is considered a protective mechanism in many water-stressed plants (Rai, 2002). Proline has been shown to increase in several different plant species under drought stress, including maize, wheat, and Miscanthus (Rampino et al., 2006; Witt et al., 2012; Ings et al., 2013), and is thought to function primarily as an osmoprotectant, thereby protecting cells from damage caused by stress (Delauney and Verma, 1993). Indeed, overproduction of proline has been shown to result in increased tolerance to osmotic stress in transgenic plants (Kishor et al., 1995; Zhu et al., 1998; Yamada et al., 2005). The lower levels of both proline and glutamate, which can act as the precursor for proline, in the SUS class, suggests increased sensitivity to drought induced osmotic damage to the ecotypes within this class. In this context, there is evidence suggesting SA can influence the accumulation of some amino acids upon drought stress. A recent study focusing on drought in Creeping Bentgrass (*Agrostis stolonifera*) showed that SA conferred drought tolerance and also the accumulation of proline, serine, threonine and alanine (Li et al., 2016). However, an SA-influenced increased accumulation of carbohydrates, as noted in the Li et al. (2016) study was not prominent in our results.

Despite increasing evidence for the involvement of JA in drought stress, there is little knowledge about its actual role in drought stress signaling, particularly when compared to its involvement in the response to biotic stresses (Du et al., 2013). In some studies, JA seems to improve drought tolerance while in others it has been reported to cause a reduction in growth and yield (Riemann et al., 2015). Interestingly, JA was one of the five metabolites contributing to the discrimination between TOL and SUS/INT under mild stress. A JA-synthesizing lipoxygenase was among the most interacting genes in a regulatory interaction network analyses of Arabidopsis exposed to mild drought stress (Clauw et al., 2015), in agreement with a role of JA in response to mild drought. Higher levels of GA, SA and JA in the TOL Brachypodium ecotypes when exposed to severe drought might suggest a protective

role for these hormones. Clearly, cross-talk between these and other hormones (including ABA, ethylene, and auxins) and their associated signal transduction elements may play important roles in the response to drought stress. Further studies are necessary to confirm that increases in GA, SA, and JA are indeed involved in establishing the drought tolerance phenotype exhibited by the TOL Brachypodium ecotypes. In particular, the possible role of JA in influencing primary metabolism to confer drought tolerance in grasses needs to be assessed.

Since analysis of both phenotypic and metabolite data, obtained from the same plant material upon progressive drought, revealed the TOL-SUS-INT clusters, there was an opportunity to assess potential correlations between the two data-sets to suggest metabolic events that could be contributing to the phenotypic changes. Biomass yield and delayed leaf senescence (stay-green) rank among the most important traits for improvement of crop plants under drought stress (Rivero et al., 2007; Salekdeh et al., 2009). Importantly, our analyses identified a number of metabolites (TCA-cycle intermediates, alanine, and SA; as well as JA; see **Supplementary Figure S3**) that correlate positively with phenotypic measures for biomass yield (area and height) and negatively with measures for stress (yellow and grey pixels). These observations would strongly suggest that SA and JA stress hormone signaling is important for maintaining some plant growth and reducing stress phenotypes. Although there is limited biological replication of the metabolite data, such conclusions highlight the value of integrative 'omic' approaches in providing novel insights into plant phenomena.

CONCLUSION

Drought screens of a diverse Brachypodium ecotype collection, integrated with phenotypic and metabolite profiling of selected Brachypodium ecotypes, highlight the variation in the response of Brachypodium ecotypes to water stress. Combined with its genotypic diversity (Gordon et al., 2014), this confirms the value of Brachypodium as a powerful model for the improvement of cereal, bioenergy, forage, and turf grasses to changing environmental conditions, including drought stress. The combination of phenotypic analysis and metabolite profiling revealed a remarkable consistency in separating the selected ecotypes into their pre-defined drought tolerance groups, highlighting the value of multivariate analysis as a robust approach to analyze complex phenotypic and metabolite data sets. The relative abundance of several metabolites, including for the phytohormones SA and JA, appear to correlate with biomass yield and reduced stress features. Although further studies are necessary to validate these findings, the results highlight the potential advantage of combining the analyses of phenotypic and metabolic responses to identify key mechanisms and markers associated with drought tolerance in crops. Overall, this work shows that different phenotyping assays could reliably identify drought-tolerant and sensitive Brachypodium lines, while

metabolite analysis may be predictive. Whilst the current study only provides correlations, future work will be aimed at gaining a better understanding of the genetic basis of the observed differences in the responses to drought stress between selected Brachypodium ecotypes and to investigate causation using recombinant inbred populations (Bettgenhaeuser et al., 2016) and reverse genetics approaches.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01751/full#supplementary-material

FIGURE S1 | Stomatal performance in selected Brachypodium ecotypes. At 5 weeks old, the nine Brachypodium accessions were either watered on a daily basis or droughted through non-watering. Stomatal conductance was determined at mid-point in the light period for each plant using a porometer every other day. Results are presented as mean conductance ($n = 5$ plants $\pm$ SE). Results are presented in three registers corresponding to susceptible (SUS), intermediate (INT) and tolerant (TOL) genotypes, respectively.

FIGURE S2 | Brachypodium ecotypes 12 days after initiation of treatments (45 das). Side image photographs show two out of the four plants harvested for metabolite profiling for each of the ecotypes. TOL, tolerant; INT, intermediate; SUS, susceptible.

FIGURE S3 | Correlation analyses of phenotypic and metabolomic data. Metabolites tentatively linked to four key pathways as targeted by unbiased chemometric analyses were correlated with phenotypic datasets. The degree of Pearson's correlation is indicated as the intensity of red (positive) or green (negative) color. r^2 values are provided in **Supplementary Table S4**.

TABLE S1 | Overview of the Brachypodium distachyon germplasm collection used for the two drought screens. Origin and further geographical information available is listed for each of the 138 ecotypes. All accession are from IBERS collection, except * Benavente Collection; ** ABR collection; ˆ Garvin collection; ~ Vogel collection. Ecotypes included in second drought screen are highlighted in grey.

TABLE S2 | Mean wilting scores obtained for each ecotype included in first drought screen. Scores are averages of three individual plants except *, based on two plants. Ecotypes included in second drought screen are highlighted in grey. SD, standard deviation.

TABLE S3 | Normalized shoot area data. Normalized shoot area data, derived from side images, used to prepare graphs shown in **Figure 3**. Standard deviation of the measurements is included (SD). Abbreviations: das, days after sowing; S-area, side area; SWC, soil water content.

TABLE S4 | Multivariate regression analyses (R^2) of phenomic and metabolomics data.

REFERENCES

Akter, N., Rafiqul Islam, M., Abdul Karim, M., and Hossain, T. (2014). Alleviation of drought stress in maize by exogenous application of gibberellic acid and cytokinin. *J. Crop Sci. Biotechnol.* 17, 41–48. doi: 10.1007/s12892-013-0117-3

Allwood, J. W., Ellis, D. I., Heald, J. K., Goodacre, R., and Mur, L. A. J. (2006). Metabolomic approaches reveal that phosphatidic and phosphatidyl glycerol phospholipids are major discriminatory non-polar metabolites in responses by *Brachypodium distachyon* to challenge by *Magnaporthe grisea*. *Plant J.* 46, 351–368. doi: 10.1111/j.1365-313X.2006.02692.x

Arbona, V., Manzi, M., Ollas, C., and Gomez-Cadenas, A. (2013). Metabolomics as a tool to investigate abiotic stress tolerance in plants. *Int. J. Mol. Sci.* 14, 4885–4911. doi: 10.3390/ijms14034885

Benoit, A., Caplier, A., Durette, B., and Herault, J. (2010). Using human visual system modeling for bio-inspired low level image processing. *Comp. Vis. Image Unders.* 114, 758–773. doi: 10.1016/j.cviu.2010.01.011

Berger, B., Parent, B., and Tester, M. (2010). High-throughput shoot imaging to study drought responses. *J. Exp. Bot.* 61, 3519–3528. doi: 10.1093/jxb/erq201

Bertolini, E., Verelst, W., Horner, D. S., Gianfranceschi, L., Piccolo, V., Inzé, D., et al. (2013). Addressing the role of microRNAs in reprogramming leaf growth during drought stress in *Brachypodium distachyon*. *Mol. Plant* 6, 423–443. doi: 10.1093/mp/sss160

Bettgenhaeuser, J., Corke, F. M. K., Opanowicz, M., Green, P., Hernández-Pinzón, I., Doonan, J. H., et al. (2016). Natural variation in *Brachypodium* links vernalization and flowering time loci as major flowering determinants. *Plant Physiol.* doi: 10.1104/pp.16.00813 [Epub ahead of print].

Brkljacic, J., Grotewold, E., Scholl, R., Mockler, T., Garvin, D. F., Vain, P., et al. (2011). *Brachypodium* as a model for the grasses: today and the future. *Plant Physiol.* 157, 3–13. doi: 10.1104/pp.111.179531

Buckler, E. S. I. V., Thornsberry, J. M., and Kresovich, S. (2001). Molecular diversity, structure and domestication of grasses. *Genet. Res.* 77, 213–218. doi: 10.1017/S0016672301005158

Cavalcanti, J. H. F., Esteves-Ferreira, A. A., Quinhones, C. G. S., Pereira-Lima, I. A., Nunes-Nesi, A., Fernie, A. R., et al. (2014). Evolution and functional implications of the tricarboxylic acid cycle as revealed by phylogenetic analysis. *Genome Biol. Evol.* 6, 2830–2848. doi: 10.1093/gbe/evu221

Chaves, M. M., Flexas, J., and Pinheiro, C. (2009). Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Ann. Bot.* 103, 551–560. doi: 10.1093/aob/mcn125

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., et al. (2014). Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *Plant Cell* 26, 4636–4655. doi: 10.1105/tpc.114.129601

Chochois, V., Vogel, J. P., Rebetzke, G. J., and Watt, M. (2015). Variation in adult plant phenotypes and partitioning among seed and stem-borne roots across *Brachypodium distachyon* accessions to exploit in breeding cereals for well-watered and drought environments. *Plant Physiol.* 168, 953–967. doi: 10.1104/pp.15.00095

Clauw, P., Coppens, F., De Beuf, K., Dhondt, S., Van Daele, T., Maleux, K., et al. (2015). Leaf responses to mild drought stress in natural variants of *Arabidopsis*. *Plant Physiol.* 167, 800–816. doi: 10.1104/pp.114.254284

Colebrook, E. H., Thomas, S. G., Phillips, A. L., and Hedden, P. (2014). The role of gibberellin signalling in plant responses to abiotic stress. *J. Exp. Biol.* 217, 67–75. doi: 10.1242/jeb.089938

Delauney, A. J., and Verma, D. P. S. (1993). Proline biosynthesis and osmoregulation in plants. *Plant J.* 4, 215–223. doi: 10.1046/j.1365-313X.1993.04020215.x

Des Marais, D. L., and Juenger, T. E. (2016). "*Brachypodium* and the abiotic environment," in *Genetics and Genomics of Brachypodium*, ed. P. J. Vogel (Cham: Springer International Publishing).

Du, H., Liu, H., and Xiong, L. (2013). Endogenous auxin and jasmonic acid levels are differentially modulated by abiotic stresses in rice. *Front. Plant Sci.* 4:397. doi: 10.3389/fpls.2013.00397

Engelbrecht, B. M. J., Tyree, M. T., and Kursar, T. A. (2007). Visual assessment of wilting as a measure of leaf water potential and seedling drought survival. *J. Trop. Ecol.* 23, 497–500. doi: 10.1017/S026646740700421X

Farooq, M., Hussain, M., Wahid, A., and Siddique, K. H. M. (2012). "Drought stress in plants: an overview," in *Plant Responses to Drought Stress: From Morphological to Molecular Features*, ed. R. Aroca (Berlin: Springer).

Feng, H. Q., Li, H. Y., and Sun, K. (2009). Enhanced expression of alternative oxidase genes is involved in the tolerance of rice (*Oryza sativa* L.) seedlings to drought stress. *Z. Naturforsch. C* 64, 704–710. doi: 10.1515/znc-2009-9-1016

Girin, T., David, L. C., Chardin, C., Sibout, R., Krapp, A., Ferrario-Mery, S., et al. (2014). *Brachypodium*: a promising hub between model species and cereals. *J. Exp. Bot.* 65, 5683–5696. doi: 10.1093/jxb/eru376

Gordon, S. P., Priest, H., Des Marais, D. L., Schackwitz, W., Figueroa, M., Martin, J., et al. (2014). Genome diversity in *Brachypodium distachyon*: deep sequencing of highly diverse inbred lines. *Plant J.* 79, 361–374. doi: 10.1111/tpj.12569

Honsdorf, N., March, T. J., Berger, B., Tester, M., and Pillen, K. (2014). High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. *PLoS ONE* 9:e97047. doi: 10.1371/journal.pone.0097047

Ings, J., Mur, L. A., Robson, P. R., and Bosch, M. (2013). Physiological and growth responses to water deficit in the bioenergy crop *Miscanthus* x giganteus. *Front. Plant Sci.* 4:468. doi: 10.3389/fpls.2013.00468

IPCC (2014). "Summary for Policymakers," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, et al. (Cambridge: Cambridge University Press).

Khokon, A. R., Okuma, E., Hossain, M. A., Munemasa, S., Uraji, M., Nakamura, Y., et al. (2011). Involvement of extracellular oxidative burst in salicylic acid-induced stomatal closure in *Arabidopsis*. *Plant Cell Environ.* 34, 434–443. doi: 10.1111/j.1365-3040.2010.02253.x

Kishor, P., Hong, Z., Miao, G. H., Hu, C., and Verma, D. (1995). Overexpression of [delta]-Pyrroline-5-Carboxylate synthetase increases proline production and confers osmotolerance in transgenic plants. *Plant Physiol.* 108, 1387–1394. doi: 10.1104/pp.108.4.1387

Kouzai, Y., Kimura, M., Yamanaka, Y., Watanabe, M., Matsui, H., Yamamoto, M., et al. (2016). Expression profiling of marker genes responsive to the defence-associated phytohormones salicylic acid, jasmonic acid and ethylene in

Brachypodium distachyon. *BMC Plant Biol.* 16:59. doi: 10.1186/s12870-016-0749-9

Li, L., Zhang, Q., and Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors* 14, 20078–20111. doi: 10.3390/s141120078

Li, Z., Yu, J., Peng, Y., and Huang, B. (2016). Metabolic pathways regulated by abscisic acid, salicylic acid and γ-aminobutyric acid in association with improved drought tolerance in creeping bentgrass (*Agrostis stolonifera*). *Physiol. Plant.* doi: 10.1111/ppl.12483 [Epub ahead of print].

Liancourt, P., and Tielbörger, K. (2009). Competition and a short growing season lead to ecotypic differentiation at the two extremes of the ecological range. *Funct. Ecol.* 23, 397–404. doi: 10.1111/j.1365-2435.2008.01497.x

Lloyd, A. J., William Allwood, J., Winder, C. L., Dunn, W. B., Heald, J. K., Cristescu, S. M., et al. (2011). Metabolomic approaches reveal that cell wall modifications play a major role in ethylene-mediated resistance against *Botrytis cinerea*. *Plant J.* 67, 852–868. doi: 10.1111/j.1365-313X.2011.04639.x

Luo, N., Liu, J., Yu, X., and Jiang, Y. (2011). Natural variation of drought response in *Brachypodium distachyon*. *Physiol. Plant.* 141, 19–29. doi: 10.1111/j.1399-3054.2010.01413.x

Miura, K., and Tada, Y. (2014). Regulation of water, salinity, and cold stress responses by salicylic acid. *Front. Plant Sci.* 5:4. doi: 10.3389/fpls.2014.00004

Mur, L. A., Allainguillaume, J., Catalan, P., Hasterok, R., Jenkins, G., Lesniewska, K., et al. (2011). Exploiting the *Brachypodium* tool box in cereal and grass research. *New Phytol.* 191, 334–347. doi: 10.1111/j.1469-8137.2011.03748.x

Muscolo, A., Junker, A., Klukas, C., Weigelt-Fischer, K., Riewe, D., and Altmann, T. (2015). Phenotypic and metabolic responses to drought and salinity of four contrasting lentil accessions. *J. Exp. Bot.* 66, 5467–5480. doi: 10.1093/jxb/erv208

Neilson, E. H., Edwards, A. M., Blomstedt, C. K., Berger, B., Møller, B. L., and Gleadow, R. M. (2015). Utilization of a high-throughput shoot imaging system to examine the dynamic phenotypic responses of a C4 cereal crop plant to nitrogen and water deficiency over time. *J. Exp. Bot.* 7, 1–16. doi: 10.1093/jxb/eru526

Okuma, E., Nozawa, R., Murata, Y., and Miura, K. (2014). Accumulation of endogenous salicylic acid confers drought tolerance to *Arabidopsis*. *Plant Signal. Behav.* 9:e28085. doi: 10.4161/psb.28085

Opanowicz, M., Vain, P., Draper, J., Parker, D., and Doonan, J. H. (2008). *Brachypodium distachyon*: making hay with a wild grass. *Trends Plant Sci.* 13, 172–177. doi: 10.1016/j.tplants.2008.01.007

Rai, V. K. (2002). Role of amino acids in plant responses to stresses. *Biol. Plant.* 45, 481–487. doi: 10.1023/A:1022308229759

Rajendran, K., Tester, M., and Roy, S. J. (2009). Quantifying the three main components of salinity tolerance in cereals. *Plant Cell Environ.* 32, 237–249. doi: 10.1111/j.1365-3040.2008.01916.x

Rampino, P., Pataleo, S., Gerardi, C., Mita, G., and Perrotta, C. (2006). Drought stress response in wheat: physiological and molecular analysis of resistant and sensitive genotypes. *Plant Cell Environ.* 29, 2143–2152. doi: 10.1111/j.1365-3040.2006.01588.x

Rancour, D. M., Marita, J. M., and Hatfield, R. D. (2012). Cell wall composition throughout development for the model grass *Brachypodium distachyon*. *Front. Plant Sci.* 3:266. doi: 10.3389/fpls.2012.00266

Riemann, M., Dhakarey, R., Hazman, M., Miro, B., Kohli, A., and Nick, P. (2015). Exploring jasmonates in the hormonal network of drought and salinity responses. *Front. Plant Sci.* 6:1077. doi: 10.3389/fpls.2015.01077

Rivero, R. M., Kojima, M., Gepstein, A., Sakakibara, H., Mittler, R., Gepstein, S., et al. (2007). Delayed leaf senescence induces extreme drought tolerance in a flowering plant. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19631–19636. doi: 10.1073/pnas.0709453104

Salekdeh, G. H., Reynolds, M., Bennett, J., and Boyer, J. (2009). Conceptual framework for drought phenotyping during molecular breeding. *Trends Plant Sci.* 14, 488–496. doi: 10.1016/j.tplants.2009.07.007

Sanchez-Martin, J., Heald, J., Kingston-Smith, A., Winters, A., Rubiales, D., Sanz, M., et al. (2015). A metabolomic study in oats (*Avena sativa*) highlights a drought tolerance mechanism based upon salicylate signalling pathways and the modulation of carbon, antioxidant and photo-oxidative metabolism. *Plant Cell Environ.* 38, 1434–1452. doi: 10.1111/pce.12501

Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., et al. (2006). Comprehensive metabolic profiling and phenotyping of interspecific

introgression lines for tomato improvement. *Nat. Biotechnol.* 24, 447–454. doi: 10.1038/nbt1192

Seki, M., Umezawa, T., Urano, K., and Shinozaki, K. (2007). Regulatory metabolic networks in drought stress responses. *Curr. Opin. Plant Biol.* 10, 296–302. doi: 10.1016/j.pbi.2007.04.014

Skirycz, A., Vandenbroucke, K., Clauw, P., Maleux, K., De Meyer, B., Dhondt, S., et al. (2011). Survival and growth of *Arabidopsis* plants given limited water are not equal. *Nat. Biotechnol.* 29, 212–214. doi: 10.1038/nbt.1800

Slavov, G., Allison, G., and Bosch, M. (2013). Advances in the genetic dissection of plant cell walls: tools and resources available in *Miscanthus*. *Front. Plant Sci.* 4:217. doi: 10.3389/fpls.2013.00217

Stanke, C., Kerac, M., Prudhomme, C., Medlock, J., and Murray, V. (2013). Health effects of drought: a systematic review of the evidence. *PLoS Curr.* 5:ecurrents.dis.7a2cee9e980f91ad7697b570bcc4b004. doi: 10.1371/currents.dis. 7a2cee9e980f91ad7697b570bcc4b004

Steinfath, M., Strehmel, N., Peters, R., Schauer, N., Groth, D., Hummel, J., et al. (2010). Discovering plant metabolic biomarkers for phenotype prediction using an untargeted approach. *Plant Biotechnol. J.* 8, 900–911. doi: 10.1111/j.1467-7652.2010.00516.x

The International Brachypodium Initiative (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463, 763–768. doi: 10.1038/nature08747

Tilly, N., Hoffmeister, D., Cao, Q., Lenz-Wiedemann, V., Miao, Y., and Bareth, G. (2015). Transferability of models for estimating paddy rice biomass from spatial plant height data. *Agriculture* 5:538. doi: 10.3390/agriculture5030538

Urano, K., Maruyama, K., Ogata, Y., Morishita, Y., Takeda, M., Sakurai, N., et al. (2009). Characterization of the ABA-regulated global responses to dehydration in Arabidopsis by metabolomics. *Plant J.* 57, 1065–1078. doi: 10.1111/j.1365-313X.2008.03748.x

Vanlerberghe, G. C. (2013). Alternative oxidase: a mitochondrial respiratory pathway to maintain metabolic and signaling homeostasis during abiotic and biotic stress in plants. *Int. J. Mol. Sci.* 14, 6805–6847. doi: 10.3390/ijms14046805

Verelst, W., Bertolini, E., De Bodt, S., Vandepoele, K., Demeulenaere, M., Enrico Pè, M., et al. (2013). Molecular and physiological analysis of growth-limiting drought stress in *Brachypodium distachyon* leaves. *Mol. Plant* 6, 311–322. doi: 10.1093/mp/sss098

Vettakkorumakankav, N. N., Falk, D., Saxena, P., and Fletcher, R. A. (1999). A crucial role for gibberellins in stress protection of plants. *Plant Cell Physiol.* 40, 542–548. doi: 10.1093/oxfordjournals.pcp.a029575

Wheeler, T., and Reynolds, C. (2013). Predicting the risks from climate change to forage and crop production for animal feed. *Anim. Front.* 3, 36–41. doi: 10.2527/af.2013-0006

Witt, S., Galicia, L., Lisec, J., Cairns, J., Tiessen, A., Araus, J. L., et al. (2012). Metabolic and phenotypic responses of greenhouse-grown maize hybrids to experimentally controlled drought stress. *Mol. Plant* 5, 401–417. doi: 10.1093/mp/ssr102

Xia, J., Sinelnikov, I. V., Han, B., and Wishart, D. S. (2015). MetaboAnalyst 3.0– making metabolomics more meaningful. *Nucleic Acids Res.* 43, W251–W257. doi: 10.1093/nar/gkv380

Yamada, M., Morishita, H., Urano, K., Shiozaki, N., Yamaguchi-Shinozaki, K., Shinozaki, K., et al. (2005). Effects of free proline accumulation in petunias under drought stress. *J. Exp. Bot.* 56, 1975–1981. doi: 10.1093/jxb/eri195

Zhu, B., Su, J., Chang, M., Verma, D. P. S., Fan, Y.-L., and Wu, R. (1998). Overexpression of a Δ1-pyrroline-5-carboxylate synthetase gene and analysis of tolerance to water- and salt-stress in transgenic rice. *Plant Sci.* 139, 41–48. doi: 10.1016/S0168-9452(98)00175-7

Public Availability of a Genotyped Segregating Population May Foster Marker Assisted Breeding (MAB) and Quantitative Trait Loci (QTL) Discovery: An Example Using Strawberry

James F. Hancock[1]*, Suneth S. Sooriyapathirana[2], Nahla V. Bassil[3], Travis Stegmeir[4], Lichun Cai[1], Chad E. Finn[5], Eric Van de Weg[6] and Cholani K. Weebadde[7]

[1] Department of Horticulture, Michigan State University, East Lansing, MI, USA, [2] Department of Molecular Biology and Biotechnology, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka, [3] United States Department of Agriculture – Agricultural Research Service, National Clonal Germplasm Repository, Corvallis, OR, USA, [4] Lassen Canyon Nursery Breeding Program, Redding, CA, USA, [5] United States Department of Agriculture – Agricultural Research Service, Horticultural Crops Research Unit, Corvallis, OR, USA, [6] Wageningen UR Plant Breeding, Wageningen, Netherlands, [7] Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI, USA

Much of the cost associated with marker discovery for marker assisted breeding (MAB) can be eliminated if a diverse, segregating population is generated, genotyped, and made available to the global breeding community. Herein, we present an example of a hybrid, wild-derived family of the octoploid strawberry that can be used by other breeding programs to economically find and tag useful genes for MAB. A pseudo test cross population between two wild species of *Fragaria virginiana* and *F. chiloensis* (FVC 11) was generated and evaluated for a set of phenotypic traits. A total of 106 individuals in the FVC 11 were genotyped for 29,251 single nucleotide polymorphisms (SNPs) utilizing a commercially available, genome-wide scanning platform (Affymetrix Axiom IStraw90[TW]). The marker trait associations were deduced using TASSEL software. The FVC 11 population segregating for daughters per mother, inflorescence number, inflorescence height, crown production, flower number, fruit size, yield, internal color, soluble solids, fruit firmness, and plant vigor. Coefficients of variations ranged from 10% for fruit firmness to 68% for daughters per mother, indicating an underlying quantitative inheritance for each trait. A total of 2,474 SNPs were found to be polymorphic in FVC 11 and strong marker trait associations were observed for vigor, daughters per mother, yield and fruit weight. These data indicate that FVC 11 can be used as a reference population for quantitative trait loci detection and subsequent MAB across different breeding programs and geographical locations.

Keywords: *Fragaria* x *ananassa*, *F. chiloensis*, *F. virginiana*, association mapping, quantitative trait analysis (QTL)

INTRODUCTION

The expense of developing a dense linkage map of single sequence repeats (SSRs) (Sargent et al., 2012) and/or single nucleotide polymorphism (SNPs) (Bassil et al., 2015) to conduct marker assisted breeding (MAB) (Collard and Mackill, 2008) is prohibitive for many breeding programs, particularly in the developing countries. Much of this cost can be circumvented; however, if a diverse, segregating population is made available to the global breeding community that has already been genotyped using markers from a published linkage map. That segregating population could be evaluated by the recipient breeding program *in situ* for the traits of interest and then the published linkage map could be used to search for markers associated with traits of interest. The only costs incurred by the recipient breeding program would be those associated with the phenotyping and computer analysis and use of only the relevant markers. The high costs associated with the generation of a dense linkage map of a segregating population could be avoided, saving the recipient breeding program hundreds of thousands of dollars in development costs. Costs could be further reduced by utilizing publicly available association mapping software such as 'Tassel' (Bradbury et al., 2007; Khan, 2011).

In this paper, we describe a hybrid population of the octoploid strawberry that should be a rich source of novel genes for the global breeding community. We have genotyped this population for 29,251 SNPs that were previously mapped in another segregating population (Bassil et al., 2015). The objective of this paper is to show how this population can be used by other breeding programs to economically find and tag useful genes for MAB. A similar approach could be used in all crops to make MAB much more available to breeding programs with limited resources.

MATERIALS AND METHODS

Segregating Population

We have generated a hybrid population of the octoploid strawberry consisting of four subspecies. The primary cultivated strawberry, *Fragaria × ananassa* Duchesne ex Rozier, initially arose from a hybridization between *F. chiloensis* (L.) Miller subsp. *chiloensis* forma *chiloensis* and *F. virginiana* Miller subsp. *virginiana* in Europe 250 years ago (Hancock, 1999). Wild-collected clones of both species have been evaluated in multiple locations to identify the possible beneficial traits that could be incorporated into the cultivated strawberry and thereby select elite germplasm (Hancock et al., 2001a,b). Elite selections of *F. virginiana* and *F. chiloensis* were intercrossed in 23 combinations and evaluated in the field in Michigan and Oregon (Hancock et al., 2010). The most impressive family was FVC-11 [(Frederick 9 × LH 50–4) × (Scotts Creek × 2 MAR 1A)], which had the best combination of fruit size, color, and yield and was composed of four different subspecies: *F. virginiana* ssp. *virginiana* from Ontario (Frederick 9, PI 612493), *F. virginiana* ssp. *glauca* from Montana (LH 50-4, PI 612495), *F. chiloensis* ssp. *chiloensis* forma *patagonica* from Chile (2 MAR 1A, PI 602567),

and *F. chiloensis* ssp. *pacifica* from California (Scotts Creek, PI 612490). This population likely contains as much diversity as is possible in a single octoploid strawberry family as it is composed of four different subspecies from four distant ecological regions spanning two continents.

Marker Development

We genotyped 106 individuals in the FVC 11 family utilizing a commercially available, genome-wide scanning platform (Affymetrix Axiom IStraw90[TW]), according to manufactures instructions (Affymetrix, Inc., Santa Clara, CA, USA). This platform was developed as part of the international RosBREED project, focused on enabling marker-assisted breeding through identification and validation of QTLs for traits of importance to breeders (Iezzoni et al., 2010; Bassil et al., 2015). Short-read sequences from one diploid and 19 octoploid accessions were aligned to the diploid *F. vesca* 'Hawaii 4' reference genome to identify SNPs and indels for incorporation into the array. A total of 95,062 marker loci (SNPs, indels, and haploSNPs) were included on the array. After removing monomorphic, segregation wise distorted and ambiguous markers by applying necessary data filters in excel, we identified polymorphisms in the FVC 11 family in a subset of 6,594 SNPs that had already been placed in a dense genetic linkage map of the full sib-family 'Holiday' × 'Korona' (van Dijk et al., 2014; Bassil et al., 2015). This gave us the physical positions of our segregating markers to do a subsequent association analysis.

Phenotypic Data

Seventy-eight genotypes of the FVC 11 population were evaluated at Benton Harbor, MI in 2007 and 2008 for their daughter plant production, inflorescence number, inflorescence height, crown production, flower number, fruit size, yield, internal color, soluble solids, and fruit firmness. In June 2007, two to three replicates (runner plants) of each genotype were planted in the field in Benton Harbor, MI, in a randomized complete block design. Plants were set in rows at 1.2 m × 1.2 m spacing and all runners were trained by cross-cultivation into a 0.5 m wide square.

A total of three random inflorescences were selected per mother and daughter plant and their heights were measured from crown to tip and their flower numbers were counted. The number of crowns was also counted on each mother plant and the three daughter plants as well as the total number of daughters produced by each mother plant (original plants set in field). Overall plant vigor was estimated on a 1–7 (least to most vigorous) scale based on plot fill and individual plant vigor. The first five ripe fruit were harvested to determine an average fruit weight, and after another five fruits had ripened all ripe and unripe fruits were picked to determine yield. Fruit firmness (g mm^{-2}) was measured on five ripe fruit per plot (when available) using the compression test of BioWorks' FirmTech 2 (Wamego, KS, USA). Two ripe fruits from each replication were cut in half and percent internal color was estimated visually based on how deep the color penetrated the flesh. Soluble solids were taken by squeezing one drop of juice onto the handheld refractometer from the two fruits for two separate readings. These data were previously reported by Stegmeir et al. (2010).

Association Analysis

Marker-trait associations were determined by using TASSEL 5 software (Bradbury et al., 2007). Diallelic SNP markers were called for genotypes in which homozygous classes were designated as 0 or 1 and heterozygous class was designated as 0.5. The monomorphic markers, the markers with ≥15% missing data and markers whose genotypes were genetically ambiguous were removed prior to the analysis. The numerical marker and trait data were uploaded to Tassel and kinship was calculated as in a classical association analysis. Then MLM mapping was used to find significant marker trait associations. The probability (p) estimate of [−log10(p)] was used to represent the strength of the associations and the threshold was set to 3.00 for reporting most significant marker-trait associations.

RESULTS AND DISCUSSION

Phenotypic Variability

As was previously reported by Stegmeir et al. (2010) significant variation was observed among the progeny for their inflorescence number, inflorescence height, crown production, flower number,

fruit size, yield, internal color, soluble solids, fruit firmness, and plant vigor (**Table 1**). Coefficients of Variations (CVs) ranged from 10% for fruit firmness to 68% for yield. When progeny means were compared with those of the parental means, many traits exhibited transgressive segregation, most notably yield, and fruit weight. Of the two parents, the *F. virginiana* had the highest value for vigor, crown production, inflorescence number, inflorescence height, flower numbers, yield, and depth of fruit color, while the *F. chiloensis* had the highest values for fruit weight and firmness.

Genotypic Variability and Association Analysis

The FVC 11 segregating population proved to be highly polymorphic. Out of the 6,594 SNPs that had been placed on the genetic linkage map of 'Holiday' × 'Korona' by Bassil et al. (2015), we found 37.42% to be polymorphic in our FVC-11 population. In the subsequent QTL analysis, we found a number of SNP markers that were closely linked to important horticultural traits, with highly significant *p* values and −log10 (*p*) values of ≥3.0. We discovered a SNP on LG 6 (AX-89796183) that was associated with both plant vigor and the production of daughter plants.

TABLE 1 | The mean, range, and coefficient of variation (CV) for 15 traits evaluated in the strawberry FVC 11 family at Benton Harbor, MI in 2008 (Stegmeir et al., 2010).

Trait	Means of parents		FVC -11 family		
	Fragaria *chiloensis*	*F. virginiana*	Mean	Range	CV (%)
Plant vigor	3.0	5.7	3.9	1.5–6.0	27
Daughters per mother	10.0	10.7	8.3	0.0–20.0	58
Crowns per mother	1.7	3.3	4.6	1.0–11.3	48
Crowns per daughter	1.0	2.0	1.4	1.0–3.0	21
Inflorescences per mother	1.7	5.7	6.8	1.0–16.3	54
Inflorescences. per daughter	1.2	2.6	1.7	0.0–4.1	29
Inflorescences height (cm)	7.0	14.8	9.7	4.1–17.7	27
Flowers per infl.	3.7	6.6	5.5	3.5–8.8	20
Fruit wt. (g)	3.1	1.9	6.1	1.0–12.6	41
Yield per plant (g)	1.0	2.0	9.0	0.6–33.6	68
Total yield (g)	33.3	68.6	213.7	22.9–651.0	68
Soluble solids	11.1	10.9	8.4	5.1–11.0	14
Fruit firmness	157.0	141.1	153.2	120.9–192.2	10
Internal color (%)	17.5	62.9	54.2	32.5–80.0	19

TABLE 2 | Significant marker-trait associations found in the FVC 11 population.

Trait	Marker	Chromosome	Position	*p*	[−log10(p)]
Vigor	AX-89796183	6	1865527	0.000496	3.30
Daughters per mother	AX-89796183	6	1865527	0.000827	3.08
Yield per plant (g)	AX-89849271	5	26321147	0.000441	3.36
Total yield (g)	AX-89876035	1	7946370	0.000745	3.13
Total yield (g)	AX-89898803	6	35968025	0.000611	3.21
Fruit wt. (g)	AX-89904664	2	9843555	0.001058	2.98
Fruit wt. (g)	AX-89880679	2	9923170	0.001058	2.98
Fruit wt. (g)	AX-89823518	2	9927708	0.000231	3.64

SNPs in two regions of LG 2 were associated with fruit weight (AX-89904664; AX-89880679, and AX-89823518). A SNP on LG 5 (AX-89849271) was associated with yield per plant, while two other SNPs on LG 1 (AX-898760359) and LG 6 (AX-89898803) were significantly associated with yield per plot (**Table 2**).

CONCLUSION

Based on the results described above, it is clear that the FVC-11 population is highly polymorphic for numerous horticulturally important traits and could provide useful genetic variability to other strawberry breeding programs. It is likely that the population also segregates for a number of other important horticultural traits that we did not evaluate. The original *F. virginiana* parent Frederick 9 is resistant to powdery mildew and leaf scorch, and very highly resistant to the northern root knot nematode. The other original *F. virginiana* parent LH-50-4 is unusually cold-hardy and is also very highly resistant to the northern root knot nematode. The *F. chiloensis* parent Scotts Creek has high salt and drought tolerance, and low nutrient needs. The other *F. chiloensis* parent MAR 1A also has high salt and drought tolerance, and low nutrient requirements. The environmental adaptations of the parents ranged from the Mediterranean climates where Scotts Creek and MAR 1A grew, to the high elevation Rocky Mountain habitat of LH 50-4 and the mid-south Canadian, continental climate of Frederick 9.

We have demonstrated how an existing, soon to be published linkage map can be used to find QTL cheaply, using publically available software (e.g.,: 'TASSEL') (Bradbury et al., 2007; Khan, 2011) and a commercially available SNP array. We will maintain the FVC 11 family at Michigan State University and will make plants available to other interested strawberry breeders until at least December 31, 2017, so that other global programs can search for breeder-friendly markers and use these informative markers themselves. We ask only that the recipients pay for any required phytosanitary analysis and shipping costs.

AUTHOR CONTRIBUTIONS

JH – Designed study and was primary source of funding, wrote manuscript. SS – Collected leaf samples for SNP analysis, conducted QTL analysis, generated tables for manuscript NB – Supervised DNA extraction and development of SNPs. TS – Collected phenotypic data. LC – Organized SNP data for analysis. CF – Generated FVC family that was analyzed. EW – Provided unpublished linkage map for QTL analysis. CW – Was secondary source of funding, played integral role in planning, and interpretation of data.

FUNDING

REFERENCES

Bassil, N. V., Davis, T. M., Zhang, H., Ficklin, S., Mittmann, M., Webster, T., et al. (2015). Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry *Fragaria* × *ananassa*. *BMC Genomics* 16:155. doi: 10.1186/s12864-015-1310-1

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Collard, B. C., and Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 557–572. doi: 10.1098/rstb.2007.2170

Hancock, J. F. (1999). *Strawberries*. Wallingford: CABI Publishing.

Hancock, J. F., Callow, P. W., Dale, A., Luby, J. J., Finn, C. E., Hokanson, S. C., et al. (2001a). From the Andes to the Rockies: native strawberry collection and utilization. *HortScience* 36, 221–225.

Hancock, J. F., Finn, C. E., Hokanson, S. C., Luby, J. J., Goulart, B. L., Demchak, K., et al. (2001b). A multistate comparison of native octoploid strawberries from North and South America. *J. Am. Soc. Hortic. Sci.* 126, 579–586.

Hancock, J. F., Finn, C. E., Luby, J. J., Dale, A., Callow, P. W., and Serçe, S. (2010). Reconstruction of the strawberry, *Fragaria* × *ananassa*, using genotypes of *F. virginiana* and *F. chiloensis*. *HortScience* 45, 1006–1013.

Iezzoni, A., Weebadde, C., Luby, J., Chengyan, Y., Van De Weg, E., Fazio, G., et al. (2010). RosBREED: enabling marker assisted breeding in Rosaceae. *Acta Hortic.* 859, 389–394. doi: 10.17660/ActaHortic.2010.859.47

Khan, M. A. (2011). *A Tutorial to Perform Association Mapping Analysis using TASSEL V 3.0 Software*. Available at: http://pbgworks.org/sites/pbgworks.org/files/associationmappingtutorialusingtassel-111119213639-phpapp01.pdf

Sargent, D. J., Passey, T., Surbanovski, N., Lopez Girona, E., Kuchta, P., Davik, J., et al. (2012). A microsatellite linkage map for the cultivated strawberry (*Fragaria x ananassa*) suggests extensive regions of homozygosity in the genome that may have resulted from breeding and selection. *Theor. Appl. Genet.* 124, 1229–1240. doi: 10.1007/s00122-011-1782-6

Stegmeir, T. L., Finn, C. E., Warner, R. M., and Hancock, J. F. (2010). Performance of an elite strawberry population derived from wild germplasm of *Fragaria chiloensis* and *F. virginiana*. *HortScience* 45, 1140–1145.

van Dijk, T., Pagliarani, G., Pikunova, A., Noordijk, Y., Yilmaz-Temel, H., and Meulenbroek, B. (2014). Genomic rearrangements and signatures of breeding in the allo-octoploid strawberry as revealed through an allele dose based SSR linkage map. *BMC Plant Biol.* 14:55. doi: 10.1186/1471-2229-14-55

Utilization of Molecular, Phenotypic, and Geographical Diversity to Develop Compact Composite Core Collection in the Oilseed Crop, Safflower (*Carthamus tinctorius* L.) through Maximization Strategy

Shivendra Kumar[1‡], **Heena Ambreen**[1‡], **Murali T. Variath**[1†], **Atmakuri R. Rao**[2], **Manu Agarwal**[1], **Amar Kumar**[1], **Shailendra Goel**[1*] and **Arun Jagannath**[1*]

[1] Department of Botany, University of Delhi, New Delhi, India, [2] Centre for Agricultural Bioinformatics, Indian Council of Agricultural Research-Indian Agricultural Statistics Research Institute, New Delhi, India

Safflower (*Carthamus tinctorius* L.) is a dryland oilseed crop yielding high quality edible oil. Previous studies have described significant phenotypic variability in the crop and used geographical distribution and phenotypic trait values to develop core collections. However, the molecular diversity component was lacking in the earlier collections thereby limiting their utility in breeding programs. The present study evaluated the phenotypic variability for 12 agronomically important traits during two growing seasons (2011–12 and 2012–13) in a global reference collection of 531 safflower accessions, assessed earlier by our group for genetic diversity and population structure using AFLP markers. Significant phenotypic variation was observed for all the agronomic traits in the representative collection. Cluster analysis of phenotypic data grouped the accessions into five major clusters. Accessions from the Indian Subcontinent and America harbored maximal phenotypic variability with unique characters for a few traits. MANOVA analysis indicated significant interaction between genotypes and environment for both the seasons. Initially, six independent core collections (CC1–CC6) were developed using molecular marker and phenotypic data for two seasons through POWERCORE and MSTRAT. These collections captured the entire range of trait variability but failed to include complete genetic diversity represented in 19 clusters reported earlier through Bayesian analysis of population structure (BAPS). Therefore, we merged the three POWERCORE core collections (CC1–CC3) to generate a composite core collection, CartC1 and three MSTRAT core collections (CC4–CC6) to generate another composite core collection, CartC2. The mean difference percentage, variance difference percentage, variable rate of coefficient of variance percentage, coincidence rate of range percentage, Shannon's diversity index, and Nei's gene diversity for CartC1 were 11.2, 43.7, 132.4, 93.4, 0.47, and 0.306, respectively while the corresponding values for CartC2 were 9.3, 58.8, 124.6, 95.8, 0.46, and 0.301. Each composite core collection represented the complete range of phenotypic and genetic variability of the crop including 19 BAPS clusters. This is the

first report describing development of core collections in safflower using molecular marker data with phenotypic values and geographical distribution. These core collections will facilitate identification of genetic determinants of trait variability and effective utilization of the prevalent diversity in crop improvement programs.

INTRODUCTION

Safflower (*Carthamus tinctorius* L.) is a dryland oilseed crop widely adapted to grow over a broad range of geographical locations extending from Far East to American region (Dajue and Mündel, 1996). It was initially cultivated for extraction of dyes and subsequently gained importance as a source of edible oil due to its nutritionally desirable composition of plant-based unsaturated fatty acids namely, oleic, and linoleic acid (Ashri et al., 1977; Dajue and Mündel, 1996; Khan et al., 2009). In addition, the medicinal properties of safflower and its use as a system for production of pharmaceutical products are well documented (Weiss, 1983; McPherson et al., 2009; Carlsson et al., 2014). Safflower is severely affected by several biotic and abiotic stresses and is characterized by low yield and spiny nature which have discouraged farmers from adopting its cultivation in several countries including India (Nimbkar, 2008). Moreover, the breeding lines and cultivars of safflower harbor low genetic diversity (Kumar et al., 2015), which restricts their utility in breeding programs. Therefore, an extensive characterization of the prevalent genetic and phenotypic diversity among the global germplasm of the crop is required to facilitate development of effective crop improvement strategies.

Germplasm resources act as a reservoir for trait variability and are of prime importance for crop improvement. However, their large size and heterogeneous structure restricts their accessibility and application (Brown, 1989a,b; Noirot et al., 1996; van Hintum, 2000). For effective management and utilization of these resources, Frankel (1984) introduced the concept of "core collection." A core collection is a representative subset of minimum number of non-redundant individuals capturing maximum variability prevalent in the entire germplasm collection. Characterization and evaluation of core collection is an easier task compared to the entire germplasm collection. Initially, core collections were developed using morphological parameters and/or geographical distribution (Huaman et al., 1999; Tai and Miller, 2001; Upadhyaya and Ortiz, 2001; Upadhyaya et al., 2003, 2009; Li et al., 2005; Bhattacharjee et al., 2007; Mahalakshmi et al., 2007). Subsequently, availability of molecular markers and their greater efficacy in elucidating genetic diversity have facilitated the development of more robust core collections using molecular markers either alone (Zhang et al., 2009) or in conjunction with phenotypic data in various crop species (Wang et al., 2006; Ebana et al., 2008; Shehzad et al., 2009; Belaj et al., 2012; Díez et al., 2012; Liu et al., 2015).

Until now, efforts to consolidate safflower genetic resources into core collections were based on assessment of morphological traits and geographical distribution. Johnson et al. (1993)

developed the first core collection in safflower consisting of 210 accessions by evaluating a germplasm collection of 2042 accessions from ~50 countries. Dwivedi et al. (2005) developed another core collection comprising 570 accessions from a total collection of 5522 safflower accessions from 38 countries. However, since most agronomically important traits are quantitative in nature, they are significantly influenced by genotype × environment (GE) interactions. Therefore, the data types (morphological and geographical information) used for development of the initial core collections in safflower would have under-represented the genetic diversity present in the crop due to lack of allelic information. Efforts are required to include genetic diversity based on molecular markers for development of a more effective and robust core collection in safflower.

The present study describes the phenotypic evaluation of a global representative collection of 531 safflower accessions and development of a robust core collection in safflower using maximization strategy. To the best of our knowledge, this is the first report of a composite core collection in safflower utilizing molecular variability along with geographical distribution and phenotypic data. This collection will be useful in designing crop improvement programs in a more effective manner and in dissecting the molecular determinants of trait variability.

MATERIALS AND METHODS

Germplasm Resources

The safflower germplasm used in the present study comprised of 531 accessions. The details of the accessions including their PI numbers, country of origin and regional pool along with the strategy used for their selection has been described by Kumar et al. (2015).

Measurement of Phenotypic Data

The accessions were grown and characterized in two consecutive seasons (2011–12 and 2012–13) at Agricultural Research Station, University of Delhi, Bawana Road, New Delhi, India (Latitude: 28° 38′ N, longitude: 77° 12′ E and altitude: 252 m). Ten seeds of each accession were sown in a single row of 2 m with an average distance of 0.2 m between plants and a gap of 0.6 m between each row. Locally adopted agronomic practices were followed for raising a healthy crop.

Phenotypic characterization was done following the guidelines of International Plant Genetic Resources Institute (IPGRI) for safflower. Each accession was characterized for 12 traits which included 8 pre-harvest and 4 post-harvest traits. The pre-harvest traits were growth habit (GH), plant height (PH), spininess (SP), number of primary branches (PB), branch location (BL), number

of heads per plant (HD), flower color (FC), and days to 50% flowering (DTF). The post-harvest traits were 100-seed weight (SW), seed oil content (OC), oleic acid content (OA), and linoleic acid content (LA). The data was recorded for three healthy plants of each accession.

Growth habit of the plant was recorded as "erect" or "sprawling" on ground. For plant height, main shoot length was measured from soil surface to the highest inflorescence of the plant. Spininess of the accessions were recorded at the onset of flowering and reported as "present" or "absent." Number of branches originating from the main axis was counted as number of primary branches. Distribution of primary branches on the main shoot determined branch location in safflower and was categorized as basal, upper one-third, upper two third, and from base to apex of the plant. The total number of inflorescences (primary, secondary, and tertiary) per plant was recorded as number of heads per plant. Flower color was documented as yellow, orange, red and off-white at full bloom stage. For each accession, the number of days from planting to onset of flowering in 50% plants was considered as days to 50% flowering. Seed weight of 100 achenes from each plant was measured in grams and recorded as 100-seed weight. Oil content was measured by Near-Infrared Reflectance Spectroscopy (NIRS) (Foss, Germany). Oil content in seed samples of 300 safflower accessions was estimated by Soxhlet method and used for the development and calibration of NIRS equations for oil content measurement in safflower (manuscript under preparation). Fatty acid composition (oleic and linoleic acid content), was determined by methyl esterification followed by gas chromatography using *Clarus* 580 (Perkin Elmer, USA) as per manufacturer's instructions.

Statistical Analysis of Phenotypic Data

Phenotypic correlations between different quantitative traits (computed as Pearson correlation coefficient, r), cluster analysis based on Euclidean distance and two-dimensional Principal coordinate analysis (PCoA) were performed using PAST *version* 3.10 (Hammer et al., 2001). Frequency distribution of accessions for different classes of traits was calculated. Evaluation of seasonal variation for the traits under consideration was conducted through Multivariate Analysis of Variance (MANOVA) using SPSS version 18 (Statistical Package for the Social Sciences; SPSS Inc. Released, 2009. PASW Statistics for Windows, Version 18.0. Chicago: SPSS Inc.).

Development of Core Collections

MSTRAT (Gouesnard et al., 2001) and POWERCORE (Kim et al., 2007) were used for development of independent core collections using phenotypic data of seasons 2011–12, 2012–13 and genotypic data reported by Kumar et al. (2015). In MSTRAT, 20 replicates and 100 iterations were tested at a fixed sample size of 10%. The core collection with highest Shannon's diversity index was selected. POWERCORE was used as described in the user's manual (Kim et al., 2007).

Evaluation of Core Collections

Core collections were evaluated by estimating Shannon's diversity index (*I*) and Nei's gene diversity (*H*) using POPGENE version

1.32 (Yeh et al., 1999). Additionally, *mean difference percentage* (MD%), *variance difference percentage* (VD%), *variable rate of coefficient of variance* (VR%), and *coincidence rate of range* (CR%) were calculated to assess the level of diversity captured in core collection with respect to the entire collection (Hu et al., 2000). *T*-test and *F*-test were performed to study difference in mean and variance of traits between the entire collection and composite core collections. The "*coverage*" criterion described by Kim et al. (2007) was used to evaluate the percentage diversity captured for each variable in the composite core collections.

RESULTS

Analysis of Pre-harvest Traits

Analysis of pre-harvest traits revealed significant phenotypic variability among the safflower accessions used in the current study. Erect growth was observed in 529 accessions while two accessions (PI-305204 and PI-306912) showed sprawling growth in both the seasons (2011–12 and 2012–13). Plant height of the studied accessions ranged from 94 to 226 cm in 2011–12 and from 73 to 211 cm in 2012–13 growing seasons (Supplementary Figures 1A,B). Although these values suggest a minor shift in the overall range between the two seasons, plant height of individual accessions did not show a markable difference. In our study, around 21% of the accessions (111) were non-spiny while 79% of accessions (420) were spiny in nature. The number of primary branches in the studied accessions ranged from 4 to 34 in 2011–12 season and from 5 to 33 for 2012–13 season. The position of branch emergence is associated with the bushy nature in safflower. A large number of accessions (38%) had branches located in the upper one third portion of the plant followed by 31% of accessions with branches in the upper two third portion. The remaining 31% of accessions had branches originating from the base till the apex giving it a more bushy appearance.

The number of heads per plant varied from 11 to 203 and from 9 to 189 for 2011–12 and 2012–13 growing seasons, respectively. Safflower shows different shades for its corolla color varying from yellow, orange, red to off-white. In our study, yellow was the most common color (76% of accessions) followed by orange (11% of accessions). Days to 50% flowering was recorded for each accession as described above. The trait distribution was observed to be asymptotically normal in both the seasons (Supplementary Figures 1C,D). Based on these observations, accessions were categorized as early flowering (tail of the distribution curve; 119–128 days and 137–146 days for 2011–12 and 2012–13 seasons, respectively), mid flowering (129–151 days for 2011–12 and 147–174 days for 2012–13, respectively) and late flowering (tail of the distribution curve; 152–160 days and 175–182 days for 2011–12 and 2012–13 season, respectively; Supplementary Figures 1C,D). Although days to 50% flowering shifted between the two seasons, no change was observed in the associated categories of accessions between the seasons. Based on the above analysis, we identified 14 early-flowering, 490 mid-flowering, and 27 late-flowering accessions.

Analysis of Post-harvest Traits

The hundred seed weight value ranged from 1 to 8 g for 2011–12 season and from 2 to 8 g in 2012–13 season. No significant

difference was observed in the phenotypic range between the two seasons. Estimation of oil content was performed using NIRS. The oil content among the analyzed accessions ranged from 16 to 50% in 2011–12 while for the 2012–13 season it ranged from 15 to 47% (Supplementary Figures 1E,F). Accessions with oil content <22% were considered as "low oil content" while those with >40% oil content were categorized as "high oil content" (Supplementary Figures 1E,F). Accessions with low and high oil content remained consistent in both the seasons. The oleic acid content ranged from 9 to 82% with most accessions (93%) falling in the lower range of oleic acid content (below 25%) and a few (7%) having medium and high oleic acid content (>75%). Linoleic acid content varied from 13 to 87% with most accessions (90%) showing high linoleic acid content (65–80%) and few accessions having very high (3%), medium or low linoleic acid content (6.6%). **Table 1** includes list of accessions with high oil content (>40%), high oleic acid (>75%), and very high linoleic acid (≥80%) observed in the current study.

Correlation Analysis between Traits

Correlation analysis indicated a significant negative correlation ($r = -0.99$) between oleic and linoleic acid content of safflower. The correlation values for all other traits were below the significance level of 0.50. The highest positive correlation value was observed between number of heads per plant and number of primary branches (0.45). The correlation coefficient values for the analyzed traits are listed in **Table 2**.

Distribution of Traits within and between Regional Gene Pools

The 531 accessions used in this study represented all the 10 regional gene pools defined by Ashri (1975) based on morphological parameters. The distribution of different

phenotypic classes among the safflower regional gene pools is given in Supplementary File 1. Although morphological delineation was not prominently observed between different regional gene pools for most traits, a few character states were more pronounced in some gene pools. Accessions with increased plant height (>155 cm) were limited to Iran-Afghanistan, Turkey, Far East, and Europe. The majority of accessions with low head count per plant were from the Far East. A higher number of primary branches (25–33) was found only among accessions from the Indian subcontinent, Far East, America, and Iran-Afghanistan. Early flowering accessions were found only among genotypes from Far East, Indian subcontinent, Egypt, and America. On the other hand, all other pre-harvest traits namely growth habit, spines, location of branches on the main axis of plant and flower color did not show any preferential distribution to any regional gene pool.

Among post-harvest traits, high oil content was observed only in accessions from the American region while some accessions from the Indian subcontinent had up to 40% of oil content (Supplementary File 1). High oleic acid content (>75%) was found only in accessions from America and Indian subcontinent. All accessions from Near East, Turkey, Egypt, Sudan, Europe, and Iran-Afghanistan had low oleic acid content. Higher ranges of 100 seed weight (6–8 gm) were found predominantly among Indian accessions and to a limited extent from American region.

Cluster Analysis and Principal Coordinate Analysis (PCoA)

The inter-relationships and genetic distance between safflower accessions based on phenotypic data was assessed through unweighted pair group method with arithmetic mean (UPGMA) clustering using Euclidean distance matrix (**Figure 1**). Safflower accessions were grouped in five major clusters designated as CL

TABLE 1 | List of safflower accessions with high oil content (>40%), linoleic acid content (≥80%), and oleic acid content (>75%).

Oil content (>40%)			Linoleic acid (≥80%)			High Oleic acid (>75%)		
PI number	Country of origin	Oil content (%)	PI number	Country of origin	Linoleic acid content (%)	PI number	Country of origin	Oleic acid content (%)
537635	USA	50	250081	Egypt	87	613394	USA	82
537701	USA	48	544025	China	82	560177	USA	81
560169	USA	47	560188	USA	81	560165	USA	81
537662	USA	46	305198	India	81	560173	USA	81
560175	USA	45	543992	China	81	560166	USA	80
560172	USA	45	514624	China	80	560169	USA	79
560168	USA	43	613459	Portugal	80	401474	Bangladesh	78
537693	USA	43	537654	USA	80	560172	USA	77
537110	USA	43	560185	USA	80	560168	USA	77
560177	USA	42	560176	USA	80	401589	India	77
537677	USA	42	537645	USA	80	537712	USA	77
560171	USA	42	537653	USA	80	470942	Bangladesh	77
537696	USA	41	251987	Turkey	80	401470	Bangladesh	76
537656	USA	41	426186	Afghanistan	80	401477	Bangladesh	76
537645	USA	41	306685	Israel	80	401476	Bangladesh	76
–	–	–	–	–	–	401479	Bangladesh	76

TABLE 2 | Correlation coefficient between eight quantitative traits studied in the entire safflower collection.

	OC	OA	LA	SW	PH	NH	NB	DTF
OC	*							
OA	0.133	*						
LA	−0.142	−0.996	*					
SW	0.044	−0.088	0.094	*				
PH	−0.130	−0.122	0.120	−0.185	*			
NH	0.065	−0.049	0.039	0.077	0.005	*		
NB	0.109	0.024	−0.028	−0.039	0.134	0.450	*	
DTF	0.164	0.027	−0.033	−0.053	0.098	0.033	0.090	*

OC, oil content; OA, oleic acid content; LA, linoleic acid content; SW, 100-seed weight; PH, plant height; NH, number of heads per plant; NB, number of primary branches; DTF, days to 50% flowering.

*Denotes correlation between same trait.
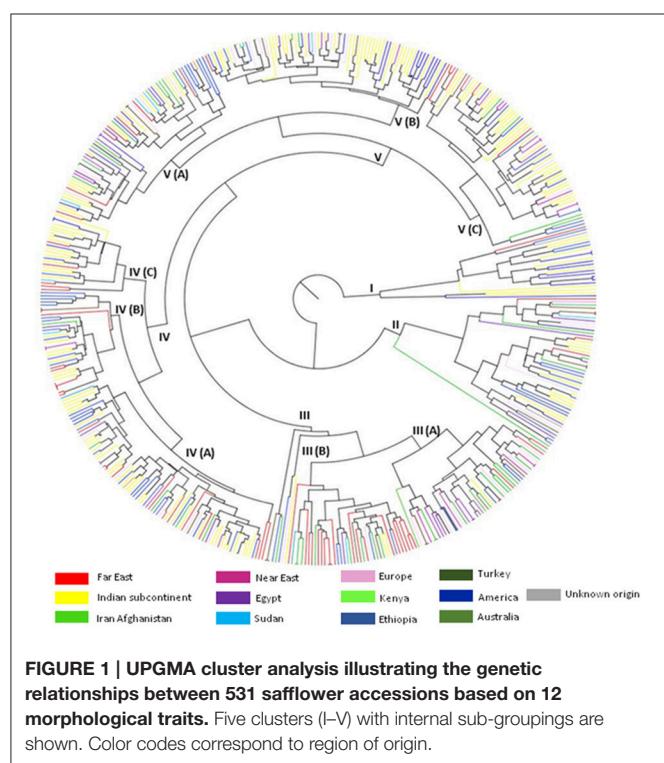
FIGURE 1 | UPGMA cluster analysis illustrating the genetic relationships between 531 safflower accessions based on 12 morphological traits. Five clusters (I–V) with internal sub-groupings are shown. Color codes correspond to region of origin.

I–CL V. Information on distribution of accessions in different clusters is given in **Table 3**. CL V is the largest cluster with 215 accessions. All clusters, except Cluster III, were dominated by accessions from the Indian subcontinent and America. Cluster III had significant representation of accessions from Iran-Afghanistan, Far-East, and Europe.

In principal coordinate analysis (PCoA), coordinate axes 1 and 2 captured 42.5 and 22.4%, respectively of the total existing variation among the accessions (**Figure 2**). Accessions from Indian subcontinent were mainly present in quadrants III and IV with minor representation in quadrants I and II. Accessions from American region were homogenously distributed among all the quadrants of PCoA obtained using phenotypic data. Accessions from Iran-Afghanistan region were mainly found in quadrants

I and II with a few accessions in quadrants III and IV. Far East accessions were restricted to quadrants I and IV while accessions from the European region were limited to quadrant I and II. Accessions from the Near East region were found to be part of quadrants I and II while Sudanese accessions were distributed in all the four quadrants. Accessions from Turkey and Egypt were predominantly found in quadrants I and II.

Analysis of Seasonal Variations and Development of Core Collections Using POWERCORE and MSTRAT

MANOVA analysis indicated significant seasonal effects as well as significant interaction effect between seasons and accession effects by considering all quantitative traits together (**Table 4**). Therefore, phenotypic data for both the seasons (2011–12 and 2012–13) and molecular marker data were treated independently for development of core collections. Usage of the two maximization (M) strategy based programs resulted in the generation of six core collections (CC1-CC6).

In our earlier work, molecular profiling of the 531 accessions identified 157 polymorphic AFLP markers (Kumar et al., 2015). Core collections were developed with these AFLP markers using POWERCORE and MSTRAT and designated as CC1 and CC4, respectively. CC1 included 14 accessions (2.6% of the entire collection) belonging to six out of 10 regional gene pools while CC4 comprised 26 accessions (4.9% of the entire collection) belonging to seven regional gene pools (**Table 5**). Phenotypic data of seasons 2011–12 and 2012–13 was used to develop core collections CC2 and CC3, respectively using POWERCORE. CC2 consisted of 26 accessions (4.9% of the entire collection) from six regional gene pools (**Table 5**) and regions of secondary introduction (Australia and America). CC3 consisted of 27 accessions (5.1% of the entire collection) from six regional gene pools of safflower. Core collections CC5 and CC6, were developed using phenotypic data of season 2011–12 and 2012–13, respectively using MSTRAT. CC5 consisted of 47 accessions (8.8% of the entire collection) from seven regional gene pools and regions of secondary introduction (America and Australia). CC6, comprising 54 accessions (10% of the entire collection) had representation from eight regional gene pools along with regions of secondary introduction.

The ranges, means, and variances for all the quantitative traits were calculated for core collections developed using phenotypic data (CC2, CC3, CC5, and CC6) and compared with corresponding values for the entire collection (Supplementary Tables 1, 2). MD% displays the difference in averages between the core and the entire collection and should be <20% for a representative core collection. MD% ranged from 6.36 to 15.45% for the four core collections (**Table 6**). VD% indicates the variance captured by a core collection and ranged from 36.4 to 59% in the current analysis. The coefficient of variance (VR%) captured in the core collection should have a value higher than 100%. CC5 and CC6 had high VR% above 105% while CC2 and CC3 showed a value of ~96.1% (**Table 6**). The range distribution of traits in a core collection in comparison to entire collection is measured by CR% whose value should be greater than 80%. All

TABLE 3 | Distribution of accessions from different regional gene pools of safflower in clusters of UPGMA dendrogram constructed using phenotypic data.

Clusters / Regional gene pools	CL I	CL II	CL III		CL IV			CL V		
			III A	III B	IV A	IV B	IV C	V A	V B	V C
1. Far East (FE)	1	3	3	20	12	7	5	7	1	7
2. Indian subcontinent (IS)	8	12	1	3	34	8	14	15	41	30
3. Iran-Afghanistan (IA)	–	5	15	9	6	1	–	11	4	4
4. Near East (NE)	–	1	8	–	3	2	–	2	3	10
5. Turkey (TU)	–	6	7	4	1	2	–	2	3	1
6. Egypt (EG)	–	5	6	1	2	1	–	3	2	1
7. Sudan (SU)	–	1	–	2	–	1	1	3	2	–
8. Kenya (KE)	–	–	–	–	–	–	–	1	–	–
9. Ethiopia (ET)	–	–	–	1	1	–	–	–	–	
10. Europe (EU)	–	4	8	2	8	–	–	6	7	2
11. America (US)	13	12	4	6	17	11	5	17	16	11
12. Australia (AUS)	–	–	–	–	1	–	–	1	1	
13. Unknown origin (UN)	1	1	–	–	–	–	–	–	–	1
Total number of genotypes	23	50	52	48	85	33	25	68	80	67
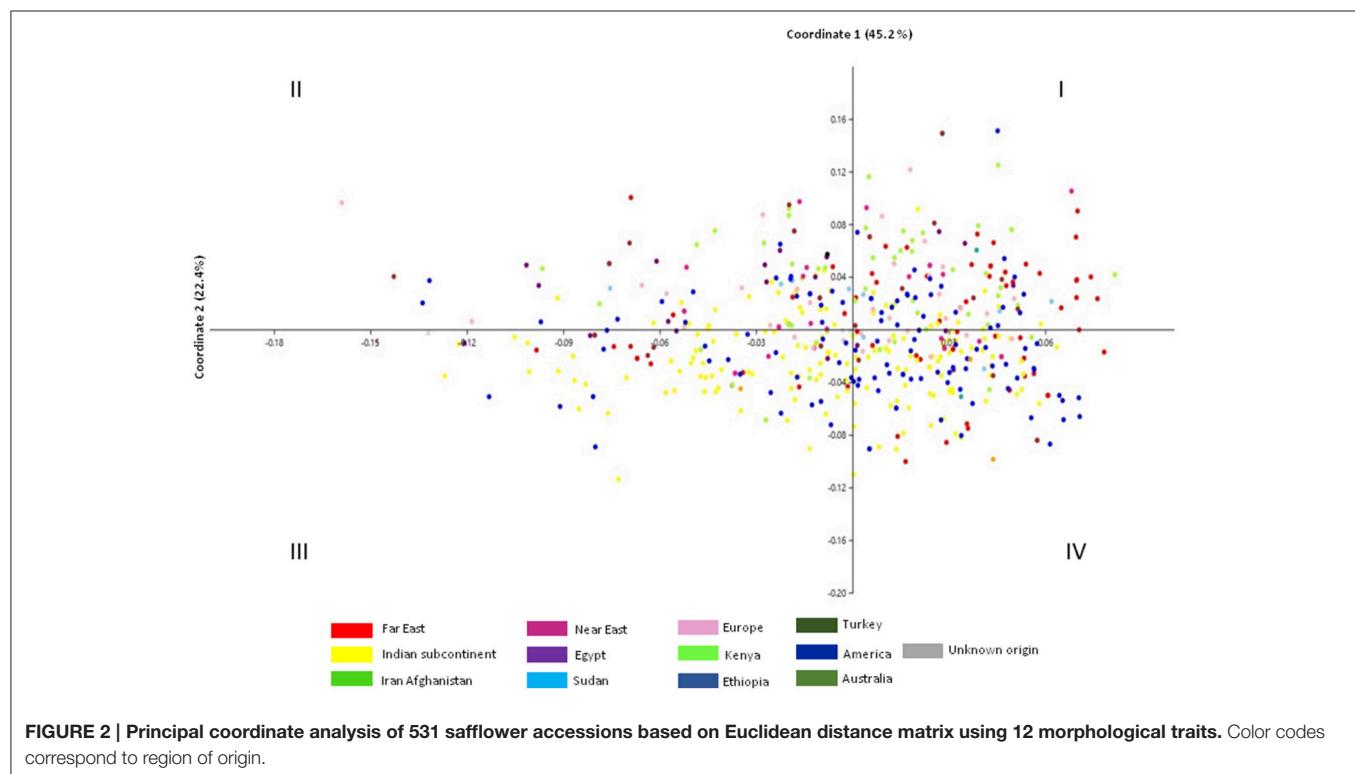
FIGURE 2 | Principal coordinate analysis of 531 safflower accessions based on Euclidean distance matrix using 12 morphological traits. Color codes correspond to region of origin.

the analyzed core collections displayed high CR% value ranging from 94.25 to 143.52%. Shannon-Weaver diversity index (I) was calculated for all the core collections and ranged from 0.44 to 0.53. The core collections, CC1 and CC4 derived using molecular marker data, showed highest Shannon-Weaver diversity index with a value of 0.53 and 0.49, respectively which was higher than the corresponding values obtained for core collections derived using phenotypic data (**Table 6**). Nei's genetic diversity (H) for

the six core collections ranged from 0.273 to 0.346. Similar to Shannon's diversity index, the highest value of H was recorded for CC1 (0.346) and CC4 (0.318) (**Table 6**).

Development and Evaluation of Composite Core Collections

Based on the various indices described above, all the core collections developed in our study appeared to represent the

prevalent diversity of the entire collection. However, none of the core collections contained representation from all the 19 clusters derived by Bayesian Analysis of Population Structure (BAPS) (**Table 7**), which captured diverse combinations of alleles and resulted in meaningful genetic stratification of the collection (Kumar et al., 2015). In order to capture the maximum range of allelic diversity/trait state in a core collection and prevent trade-off between two data types when used together, we attempted to combine phenotypic and molecular variability by merging core collections derived from each strategy separately (**Figure 3**). The core collections developed by POWERCORE, i.e., CC1 (14 accessions), CC2 (26 accessions), and CC3 (27 accessions) were combined to form a non-redundant composite core collection referred to as CartC1 (Supplementary File 2). CartC1 comprised 57 accessions (10.7% of initial collection) representing 19 BAPS clusters, eight regional gene pools and two regions of secondary introduction for safflower (**Tables 5**, **7**). Similarly, the core collections derived through MSTRAT, i.e., CC4 (26 accessions), CC5 (47 accessions), and CC6 (54 accessions) were merged resulting in a non-redundant composite core collection referred to as CartC2 (Supplementary File 2). CartC2 consisted of 106 accessions (∼20% of initial collection) including representation from all 19 BAPS clusters, ten regional gene pools and two

regions of secondary introduction for safflower (**Tables 5**, **7**). Forty four accessions were common among the two composite core collections (**Figure 3**).

The ranges, means and variances for all the quantitative traits for CartC1 and CartC2 are provided in **Table 8**. Homogeneity tests were performed to evaluate the difference in means (*t*-test) and variances (*F*-test) of traits between the entire collection and composite core collections (α = 0.05; **Table 8**). For a core collection to be representative of the entire collection, it is expected that the difference in mean should not deviate by more than 20% for the traits (Hu et al., 2000). Difference between the mean of the entire collection and CartC1 was non-significant for oil content, 100 seed weight, plant height, number of heads per plant, number of primary branches per plant, and days to 50% flowering. We observed non-significant differences in variance for three traits (100 seed weight, plant height, number of primary branches per plant) between CartC1 and the entire collection (**Table 8**). *T*-test provided non-significant differences for oil content, 100 seed weight, plant height, number of heads per plant, and days to 50% flowering while *F*-test revealed non-significant variance for only two traits (100 seed weight and plant height) between CartC2 and the entire collection. In "*Coverage*" analysis (Kim et al., 2007), CartC1 and CartC2 showed 100% *coverage* value for different phenotypic and genetic variables under consideration.

The composite core collections were validated for their representativeness of the entire collection through evaluation indices which are given in **Table 6**. The Shannon's diversity index (*I*) and Nei's genetic diversity (*H*) were 0.47 and 0.306, respectively for CartC1 and 0.46 and 0.301, respectively for CartC2. We assessed distribution of accessions of CartC1 and CartC2 in the dendrogram obtained through phenotypic and genetic analysis of entire collection. CartC1 and CartC2 showed balanced distribution in all the clusters of Neighbor Joining (genetic analysis; **Figure 4**) and UPGMA (phenotypic analysis;

TABLE 4 | Multivariate analysis of variance (MANOVA) to study seasonal differences in quantitative traits.

Effect	Value	*F*	Sig.
Accession	0.001	22.807	<0.001
Season	0.209	1860.796	<0.001
Replication	0.966	8.541	<0.001
Accession * Season	0.104	3.306	<0.001

Based on Wilks' Lambda.

TABLE 5 | Representation from different regional gene pools of safflower in developed core collections (CC).

Regional gene pools	Entire collection	POWERCORE			MSTRAT			Composite core collections	
		CC1	CC2	CC3	CC4	CC5	CC6	CartC1	CartC2
Far East (FE)	67	2	4	4	2	7	9	7	14
Indian subcontinent (IS)	167	5	5	8	7	12	13	14	26
Iran-Afghanistan (IA)	54	1	3	0	4	5	7	4	13
Near East (NE)	26	0	2	1	3	2	0	3	5
Turkey (TU)	29	1	1	0	3	0	1	2	4
Egypt (EG)	20	1	2	2	0	4	3	4	5
Sudan (SU)	10	0	0	0	1	0	0	0	1
Kenya (KE)	1	0	0	0	0	0	1	0	1
Ethiopia (ET)	2	0	0	1	0	1	1	1	1
Europe (EU)	37	2	2	3	2	2	5	7	7
America (US)	112	2	6	7	4	13	12	14	27
Australia (AUS)	3	0	1	1	0	1	1	1	1
Unknown origin (UN)	3	0	0	0	0	0	1	0	1
TOTAL	531	14	26	27	26	47	54	57	106

TABLE 6 | Evaluation indices for developed core collections.

Core Collection	MD%	VD%	VR%	CR%	I	H
CC 1	1.38	12.4	106.55	100	0.52	0.346
CC 2	13.49	54.64	96.14	143.52	0.45	0.293
CC 3	11.03	49.81	96.1	136.42	0.42	0.273
CC 4	3.8	18.76	95.5	66.5	0.48	0.318
CC 5	6.36	36.4	116.8	94.25	0.44	0.289
CC 6	15.45	59	147.25	95.75	0.44	0.289
CartC1	11.2	43.7	132.4	93.4	0.47	0.306
CartC2	9.3	58.8	124.6	95.8	0.46	0.301

MD%, mean difference; VD%, variance difference; VR%, variable rate; CR%, coincidence rate of range; I, Shannon's diversity index, and H, Nei's genetic diversity.

Figure 5). Thus, CartC1 and CartC2 provided a more rational and exhaustive representation of all the phenotypic and genetic variability than the independent core collections (CC1–CC6) developed in the present study.

DISCUSSION

A vast collection consisting of 25,179 accessions of safflower is available in 22 gene banks of 15 countries around the world (Zhang and Johnson, 1999). Phenotypic characterization of safflower germplasm in earlier studies demonstrated significant variability for several agronomic traits (Knowles, 1969; Ashri, 1975; Johnson et al., 2001; Amini et al., 2008; Khan et al., 2009). In spite of substantial diversity in its germplasm, yield enhancement in the crop has achieved limited success. Breeding strategies often focus on a limited set of agronomic traits resulting in cultivars with a narrow genetic base. For example, Kumar et al. (2015) showed that the cultivars and breeding lines of safflower from the Indian subcontinent have a narrow genetic base although extensive genetic diversity was present in the regional germplasm. This makes the cultivars highly susceptible to environmental changes and vulnerable to yield penalties. One of the main limitations of earlier approaches has been the over-dependence on morphological and geographical parameters due to lack of information on the genetic structure of safflower germplasm based on molecular markers. The present study attempted to address the above issue by generating two composite core collections in safflower that include data on molecular variability of the crop in addition to phenotypic and geographical parameters.

Phenotypic Diversity of the Crop and Identification of Accessions with Desirable Agronomic Traits

Significant variation was observed among the 531 accessions for 12 agronomic traits. More than 85% of accessions had plant height <155 cm, which is desirable due to ease of mechanical harvesting from shorter plants (Weiss, 1983). Most safflower varieties and genotypes grown around the world have spines on the leaves and bracts of the plant (Dajue and Mündel, 1996). Spiny nature of the crop is one of the factors responsible for

TABLE 7 | Distribution of accessions of developed core collections in different BAPS clusters derived based on AFLP markers by Kumar et al. (2015).

BAPS clusters	Number of accessions							
	CC1	CC2	CC3	CC4	CC5	CC6	CartC1	CartC2
Bacl 1	—	1	1	—	3	2	2	4
Bacl 2	2	1	1	2	1	3	4	5
Bacl 3	—	1	1	3	2	6	1	7
Bacl 4	1	1	1	3	2	2	2	6
Bacl 5	1	3	2	—	5	5	5	7
Bacl 6	1	—	—	2	—	2	1	4
Bacl 7	1	1	3	2	5	3	6	8
Bacl 8	1	1	—	1	—	1	2	2
Bacl 9	1	2	2	1	3	3	2	6
Bacl 10	—	1	2	—	4	4	2	6
Bacl 11	1	1	—	2	1	2	2	5
Bacl 12	1	3	—	3	4	2	4	8
Bacl 13	—	2	2	—	1	3	4	7
Bacl 14	1	3	4	2	5	5	8	12
Bacl 15	—	1	1	—	1	2	1	2
Bacl 16	1	1	4	—	2	4	5	4
Bacl 17	—	1	1	2	1	1	1	4
Bacl 18	1	1	1	2	3	4	2	8
Bacl 19	1	1	1	1	—	—	3	1

reluctance of farmers to grow safflower, especially in countries like India where harvesting is done manually. Spiny types were widely represented in our collection of 531 accessions. It was hypothesized that non-spiny varieties are generally low in yield and oil content (Dajue and Mündel, 1996). However, we did not observe a significant association between presence of spines and seed oil content. We identified 15 spiny accessions with high seed oil content (Table 1) and several spiny accessions with low seed oil content in the representative collection. The genetics of oil content and spines needs to be investigated further in order to design effective breeding strategies involving these traits.

Traits such as number of primary branches and heads per plant influence seed yield (Ashri et al., 1974; Patil et al., 1994; Dajue and Mündel, 1996). We found significant variation in the above traits and accessions with high number of primary branches and increased number of heads were identified. Analysis of seed yield for these accessions is required to identify promising genotypes. Days to 50% flowering varied between the two growing seasons and ranged from 119 to 160 days (in 2011–12) and from 137 to 182 days (in 2012–13). Delayed flowering in the second year was attributable to cooler temperatures in February and March than in the previous year. The average maximum temperature recorded for the months of February and March 2012 was ~29°C while the corresponding value was ~23°C in 2013. (http://www.weatherspark.com). Though temperature fluctuations did affect developmental stages as well as flowering-related events, the early flowering accessions were consistent between the two seasons.
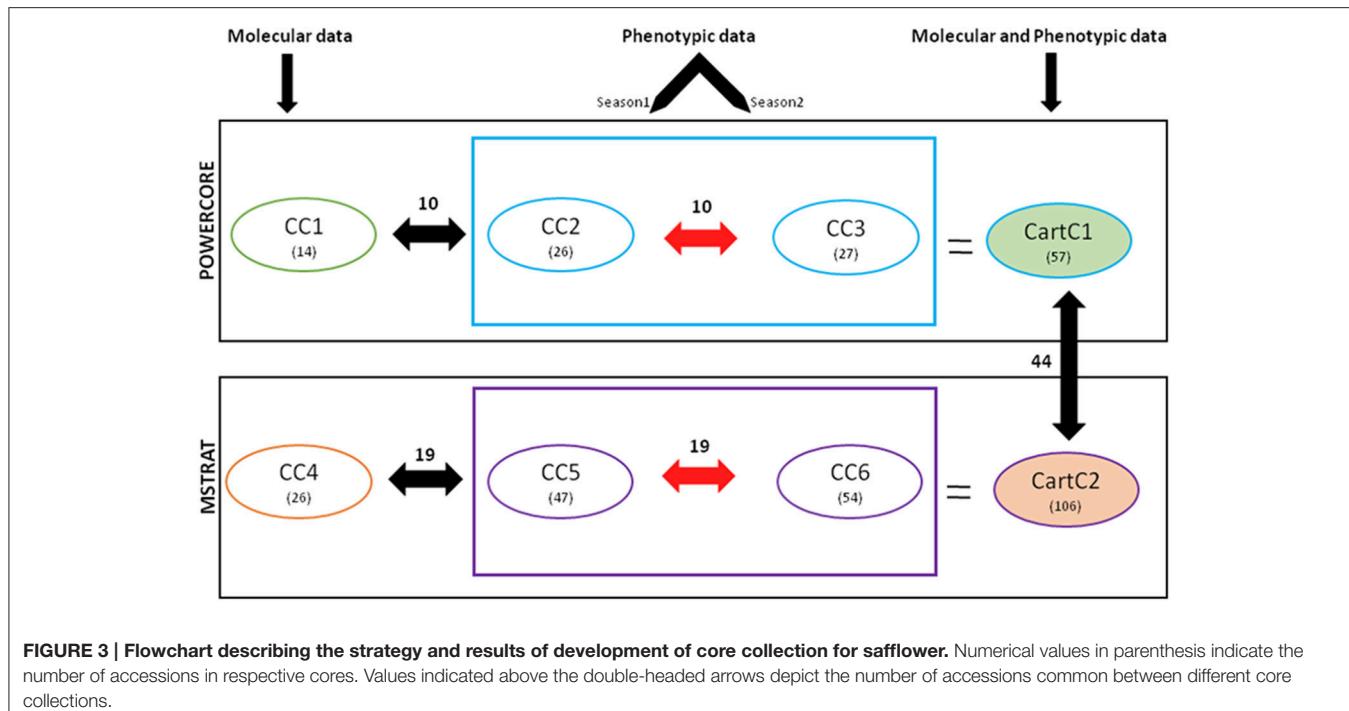
FIGURE 3 | Flowchart describing the strategy and results of development of core collection for safflower. Numerical values in parenthesis indicate the number of accessions in respective cores. Values indicated above the double-headed arrows depict the number of accessions common between different core collections.

TABLE 8 | Ranges, means, and variances for the entire collection and composite core collections, CartC1 and CartC2.

Phenotypic Traits*	Entire collection[#]			CartC1 (POWERCORE)					CartC2 (MSTRAT)				
	Range	Mean	Variance	Range	Mean	Variance	T-test	F-test	Range	Mean	Variance	T-test	F-test
OC (%)	16–50	31	20.2	18–50	31.4	35.3	NS	S	16–50	32	32.7	NS	S
OA (%)	9–82	19	161.2	10–79	28.9	503.1	S	S	10–79	24	359.0	S	S
LA (%)	13–87	71	151.7	13–87	61.7	476.8	S	S	13–87	67	344.3	S	S
SW (g)	1–8	5	2.0	1–8	4.3	2.2	NS	NS	1–8	4	2.4	NS	NS
PH (cm)	94–226	154	578.3	109–226	155	712.1	NS	NS	94–226	156	735.9	NS	NS
NH	11–203	74	1508	16–203	84	3382	NS	S	16–203	79	2330.8	NS	S
NB	5–33	14	25.5	6–33	16	47.5	NS	NS	6–33	16	40.1	S	S
DTF	119–160	141	44.0	120–160	141	75.4	NS	S	119–160	141	91.0	NS	S

*Abbreviations for phenotypic traits provided in footnote of **Table 2**; [#]Data for season 2011–2012 presented.

Identification and use of high oil yielding genotypes is important for increasing oil content in safflower cultivars. Breeding efforts in America led to the development of cultivars with increased seed oil content ranging from 45 to 55% (Bergman et al., 1985; Rubis et al., 2001). However, such improvements are lacking among Indian cultivars which have oil content ranging from 27 to 35%. Evaluation of oil quantity in the 531 accessions by NIRS identified 15 accessions with high oil content (>40%). These would serve as important breeding material in safflower. All the high oil yielding accessions (>40%) had low 100-seed weight (3–4 gm) in our study. This observation is in consonance with earlier reports, which suggest that increased hull thickness enhances seed weight but reduces oil content (Ranga Rao et al., 1977; Dajue and Mündel, 1996). Safflower oil has a desirable fatty acid composition. High linoleic lines of safflower are favored for animal feed and in the paint and varnish industry (Knowles,

1989; Bergman et al., 2001) while high oleic lines are nutritionally desirable because of its hypo-cholesterolemic effect and greater oxidative stability (Fuller et al., 1967). A high oleic line of Indian origin (Knowles and Bill, 1964) was effectively utilized in various safflower breeding programs in the USA (Mündel and Bergman, 2009). The safflower collection used in this study contained 17 accessions with high oleic acid content and 15 accessions with high linoleic acid content (**Table 1**). Accessions with desirable traits identified in the present study could be incorporated in breeding programs for crop improvement.

Assessment of Regional Gene Pools Based on UPGMA Analysis of Phenotypic Data

Accessions from the Indian Subcontinent and American region were distributed in all the five clusters (**Figure 1**) suggesting that they harbor maximum phenotypic diversity for the studied
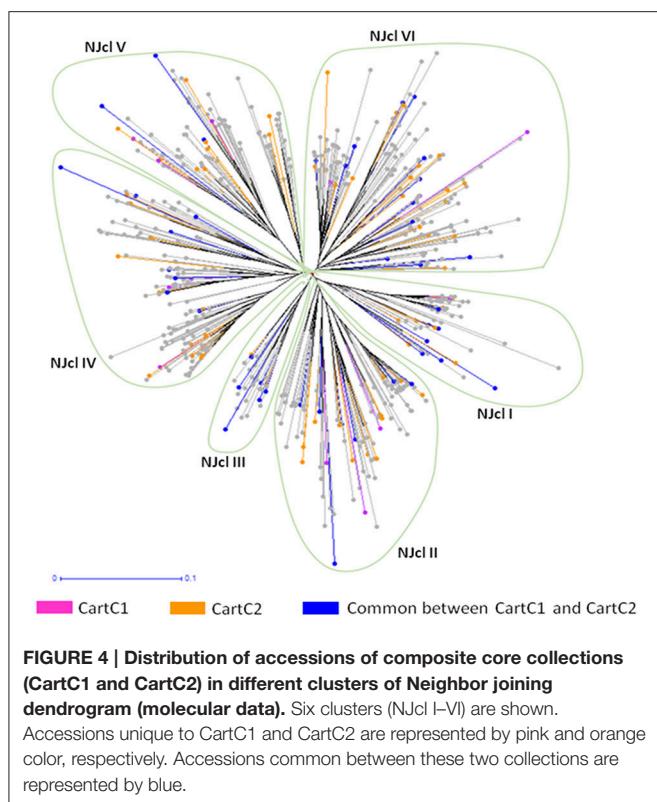
FIGURE 4 | Distribution of accessions of composite core collections (CartC1 and CartC2) in different clusters of Neighbor joining dendrogram (molecular data). Six clusters (NJcl I–VI) are shown. Accessions unique to CartC1 and CartC2 are represented by pink and orange color, respectively. Accessions common between these two collections are represented by blue.

traits. Knowles (1969), based on morphological analysis of accessions from the Indian subcontinent, reported them as a uniform assemblage resulting from a single introduction. In contrast, our assessment indicates that accessions from the Indian subcontinent are phenotypically diverse. Morphological diversity among accessions from the Indian subcontinent was reported in earlier studies (Kupsow, 1932; Chavan, 1961; Hanelt, 1961). Indian accessions have also been shown to harbor significant genetic diversity (Kumar et al., 2015). The American germplasm was found to be phenotypically diverse in the current study but was genetically conserved (Kumar et al., 2015). Near East and Iran-Afghanistan accessions clustered together based on phenotypic data, supporting our earlier proposal of considering them as a single gene pool (Kumar et al., 2015). Accessions from European region were distributed in several clusters based on phenotypic data similar to the observation obtained through molecular data analysis. Interestingly, accessions from Far East, Turkey and Egyptian region were present in all the clusters although they exhibited low genetic diversity (Kumar et al., 2015). These results indicate that UPGMA analysis based on phenotypic data alone is unable to accurately define the genetic relationships among safflower accessions.

Composite Core Collections Effectively Capture the Global Molecular, Phenotypic, and Geographical Variability of the Crop

In recent years, increased availability of molecular resources has enabled their utilization in development of core collections

in crop species (Belaj et al., 2012; El Bakkali et al., 2013) but until now, no such attempts have been made in safflower. Use of molecular markers for development of core collections is advantageous as they reflect diversity at the DNA level as opposed to morphological markers wherein different genotypes might show similar phenotypic traits due to environmental effects. Additionally, molecular markers are more effective in identifying and minimizing redundancy. Several studies have emphasized on use of maximization (M) strategy for development of highly robust core collections (Bataillon et al., 1996; McKhann et al., 2004). The M strategy retains maximum number of alleles at each locus and is considered as the most powerful approach for maintaining diverse alleles (Schoen and Brown, 1993). MSTRAT and POWERCORE programs have been successfully used for construction of core collection in various plant species such as grapes, olive and sesame (Le Cunff et al., 2008; Belaj et al., 2012; Zhang et al., 2012). A combination of molecular markers and maximization (M) strategy has been utilized for the first time in our study for construction of a core collection in safflower.

Earlier studies reported significant GE interactions in safflower and emphasized on multi-location and multi-seasonal trials to evaluate heritability of characters for their effective utilization in breeding programs (Singh et al., 2004; Mahasi et al., 2006). In our study, MANOVA analysis indicated prominent GE interactions (**Table 4**). Therefore, seasonal datasets were treated independently for developing core collections. The six core collections thus generated, efficiently captured the entire range of trait variability but failed to include complete genetic diversity represented in 19 clusters derived earlier (Kumar et al., 2015) through Bayesian analysis. Additionally, many accessions were common between different core collections. For example, in core collections developed using POWERCORE, 10 accessions were common between CC1 (marker-based) and CC2/CC3 (phenotype-based). Only 4 accessions were unique to CC1 while 16 and 17 accessions were unique to CC2 and CC3, respectively. In MSTRAT-derived core collections, 19 accessions were common between CC4 (marker-based) and CC5/CC6 (phenotype-based). The number of accessions unique to CC4, CC5, and CC6 were 7, 28, and 35, respectively. The presence of common accessions between core collections derived using different types of data indicates an overlap in genetic and phenotypic components of the studied accessions. These accessions represent a subset of genotypes that are highly diverse at both molecular and phenotypic level.

The core collections developed using each program were merged to derive a more robust and non-redundant composite core collection (CartC1 by POWERCORE and CartC2 by MSTRAT). The vast phenotypic diversity of the initial collection was retained in both collections. Accessions with desirable agronomic traits and extreme phenotypes, which were present in very low numbers in the entire collection and scattered in the initial core collections were captured in the composite core collections (**Table 8**). Both the composite core collections provided comprehensive coverage of allelic diversity and had representation from all the 19 BAPS clusters identified earlier for safflower (Kumar et al., 2015; **Table 7**). Evaluation indices (MD%, VD%, VR%, CR%, *I*, *H*) for CartC1 and CartC2 were
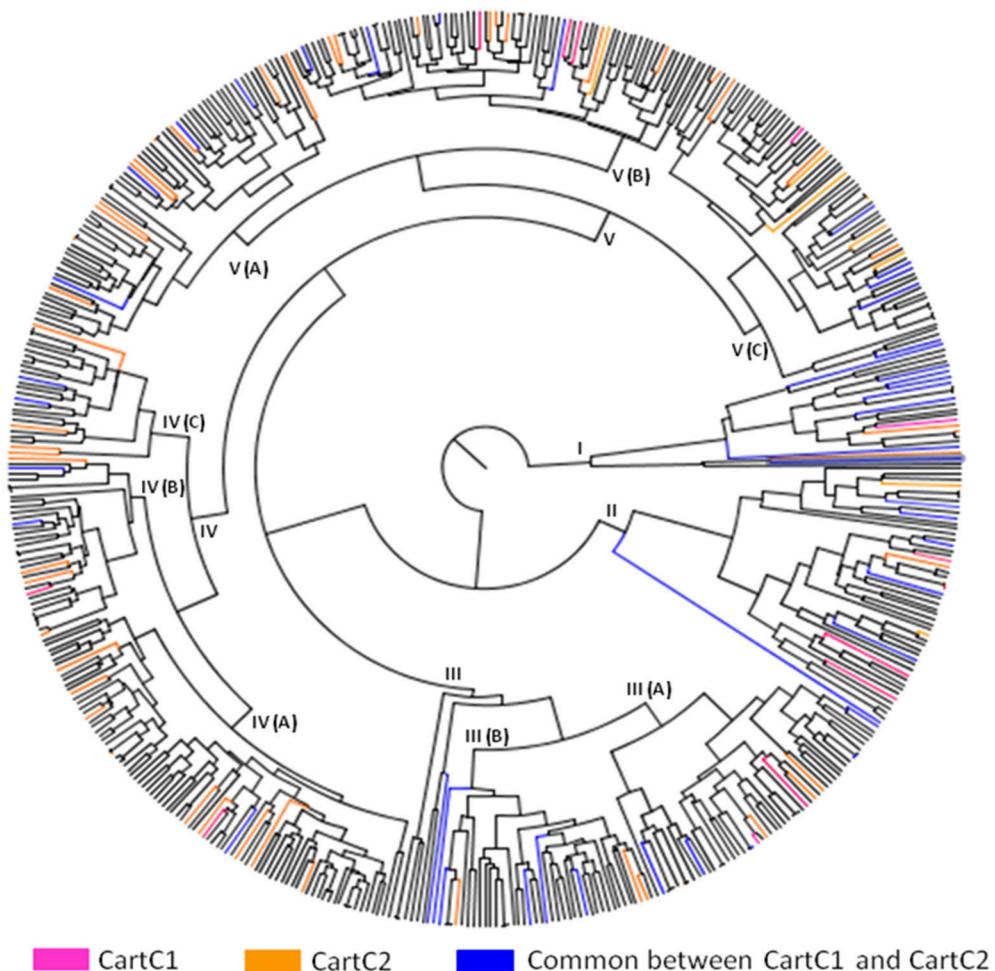
FIGURE 5 | Distribution of accessions of composite core collections (CartC1 and CartC2) in different clusters of UPGMA dendrogram (morphological data). Accessions unique to CartC1 and CartC2 are represented by pink and orange color, respectively. Accessions common between these two collections are represented by blue.

comparable and reflect their effectiveness in capturing diversity of the crop (**Table 6**). Our approach of deriving independent core collections from molecular and phenotypic data and their subsequent merger to create composite core collections avoided trade-off between the diversity captured using the molecular and phenotypic data sets.

Geographical distribution influences the extent of genetic variability of a species. The effect is more prominently seen in case of in-breeding species (Rao and Hodgkin, 2002). Geographical patterning is evident in safflower which is highly self-pollinating in nature and is grown in different agro-climatic regions across the world (Knowles, 1969; Ashri, 1975; Chapman et al., 2010). The two composite core collections showed minor variations in representation of the 10 regional gene pools. CartC1 included 8 regional pools excluding Sudan and Kenya while CartC2 contained representation from all the 10 regional gene pools (**Table 5**). Similar to the entire collection, both CartC1 and CartC2 showed predominance of accessions from Indian subcontinent and America accounting for ~50% of the total

entries. In contrast, the earlier core collection developed by Johnson et al. (1993) had a major proportion of accessions (~46%) from the Mediterranean region and South-West Asia while the core collection derived by Dwivedi et al. (2005) consisted of ~78% accessions from South and South-East Asia.

The number of accessions in a core collection is an important factor determining its effective utilization (Brown and Spillane, 1999). The core collections developed earlier for safflower consisted of 210 accessions (Johnson et al., 1993) and 570 accessions (Dwivedi et al., 2005) while the composite core collections developed in the present study are comparatively smaller with 57 (CartC1) and 106 (CartC2) accessions. The larger size of the core collections developed in earlier studies could be due to the larger number of accessions in their initial germplasm collection. However, the advantage of the present study is that the initial collection used for development of composite core collections has been characterized extensively for both molecular and phenotypic diversity and the generated core collections have therefore effectively captured the global genetic and phenotypic

diversity of the crop. Additionally, CartC1 has better utility value in comparison to CartC2 due to its smaller size and comparable diversity.

The present study is the first attempt where molecular diversity data has been used in conjunction with phenotypic data and geographical distribution to develop core collections in safflower. The small size of the composite core collections would be advantageous for field studies and association mapping. These collections will provide access to genetically diverse and agronomically important germplasm that would be useful in widening the genetic base of the crop and facilitate characterization of genetic determinants of trait variability. This information can be used to design more effective breeding programs to increase the global utility of safflower as an oilseed crop.

AUTHOR CONTRIBUTIONS

AJ, SG conceived and designed the experiments. SK, HA performed the experiments, analyzed the data. MV helped in collection of phenotypic data. AR provided necessary support in the statistical analysis. SK, HA, SG, AJ wrote the manuscript. MA, AK, AJ, SG provided facilities for completion of experiments and reviewed the manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01554

REFERENCES

Amini, F., Saeidi, G., and Arzani, A. (2008). Study of genetic diversity in safflower genotypes using agro-morphological traits and RAPD markers. *Euphytica* 163, 21–30. doi: 10.1007/s10681-007-9556-6

Ashri, A. (1975). Evaluation of the germ plasm collection of safflower, *Carthamus tinctorius* L. V. Distribution and regional divergence for morphological characters. *Euphytica* 24, 651–659. doi: 10.1007/BF00132903

Ashri, A., Knowles, P. F., Urie, A. L., Zimmer, D. E., Cahaner, A., and Marani, A. (1977). Evaluation of the germplasm collection of Safflower, *Charthamus tinctorius* L. III. Oil content and iodine value and their associations with other characters. *Econ. Bot.* 31, 38–46. doi: 10.1007/BF02860650

Ashri, A., Zimmer, D. E., Urie, A. L., Cahaner, A., and Marani, A. (1974). Evaluation of the world collection of safflower, *Carthamus tinctorius* L. IV. Yield and yield components and their relationships. *Crop Sci.* 14, 799–802. doi: 10.2135/cropsci1974.0011183X001400060006x

Bataillon, T. M., David, J. L., and Schoen, D. J. (1996). Neutral genetic markers and conservation genetics: simulated germplasm collections. *Genetics* 144, 409–417.

Belaj, A., del Carmen Dominguez-García, M., Atienza, S. G., Urdíroz, N. M., De la Rosa, R., Satovic, Z., et al. (2012). Developing a core collection of olive (*Olea europaea* L.) based on molecular markers (DArTs, SSRs, SNPs) and agronomic traits. *Tree Genet. Genomes* 8, 365–378. doi: 10.1007/s11295-011-0447-6

Bergman, J. W., Carlson, G., Kushnak, G., Riveland, N. R., and Stallknecht, G. (1985). Registration of 'Oker' safflower. *Crop Sci.* 25, 1127–1128. doi: 10.2135/cropsci1985.0011183X002500060063x

Bergman, J. W., Flynn, C. R., Kott, R. W., Hatfield, P. G., Wagoner, H. V., Boles, J. A., et al. (2001). "Feedlot performance, carcass composition and muscle and fat conjugated linoleic acid concentrations of lambs fed diets supplemented with high linoleic safflower," in *Proceedings of the 5th International Safflower Conference, Williston, North Dakota and Sidney, Montana, USA, 23–27 July, 2001. Safflower: A Multipurpose Species with Unexploited Potential and World Adaptability* (Department of Plant Pathology, North Dakota State University), 15–22.

Bhattacharjee, R., Khairwal, I. S., Bramel, P. J., and Reddy, K. N. (2007). Establishment of a pearl millet [*Pennisetum glaucum* (L.) R. Br.] core collection based on geographical distribution and quantitative traits. *Euphytica* 155, 35–45. doi: 10.1007/s10681-006-9298-x

Brown, A. H. D. (1989a). "The case for core collections," in *The Use of Plant Genetic Resources*, eds A. H. D. Brown, O. H. Frankel, D. R. Marshall, and J. T. Williams (Cambridge: Cambridge University Press), 136–156.

Brown, A. H. D. (1989b). Core collection: a practical approach to genetic resources management. *Genome* 31, 818–824. doi: 10.1139/g89-144

Brown, A. H. D., and Spillane, C. (1999). "Implementing core collections–principles, procedures, progress, problems and promise," in *Core Collections for Today and Tomorrow*, eds R. C. Johnson and T. Hodgkin (Madison, WI: Crop Science Society of America), 1–10.

Carlsson, A. S., Zhu, L. H., Andersson, M., and Hofvander, P. (2014). Platform crops amenable to genetic engineering – a requirement for successful production of bio-industrial oils through genetic engineering. *Biocatal. Agric. Biotechnol.* 3, 58–64. doi: 10.1016/j.bcab.2013.12.007

Chapman, M. A., Hvala, J., Strever, J., and Burke, J. M. (2010). Population genetic analysis of safflower (*Carthamus tinctorius* L.; Asteraceae) reveals a Near Eastern origin and five centers of diversity. *Am. J. Bot.* 97, 831–840 doi: 10.3732/ajb.0900137

Chavan, V. M. (1961). *Niger and Safflower*. Hyderabad: Indian Central Oilseeds Committee.

Dajue, L., and Mündel, H. H. (1996). *Safflower Carthamus tinctorius L. Promoting the Conservation and Use of Underutilized and Neglected Crops 7*. Gatersleben; Rome: Institute of Plant Genetics and Crop Plant Research; International Plant Genetic Resources Institute.

Díez, C. M., Imperato, A., Rallo, L., Barranco, D., and Trujillo, I. (2012). Worldwide core collection of olive cultivars based on simple sequence repeat and morphological markers. *Crop Sci.* 52, 211–221. doi: 10.2135/cropsci2011.02.0110

Dwivedi, S. L., Upadhyaya, H. D., and Hegde, D. M. (2005). Development of core collection using geographic information and morphological descriptors in safflower (*Carthamus tinctorius* L.) germplasm. *Genet. Resour. Crop Evol.* 52, 821–830. doi: 10.1007/s10722-003-6111-8

Ebana, K., Kojima, Y., Fukuoka, S., Nagamine, T., and Kawase, M. (2008). Development of mini core collection of Japanese rice landrace. *Breed. Sci.* 58, 281–291. doi: 10.1270/jsbbs.58.281

El Bakkali, A., Haouane, H., Moukhli, A., Costes, E., Van Damme, P., and Khadari, B. (2013). Construction of core collections suitable for association mapping to optimize use of Mediterranean olive (*Olea europaea* L.) genetic resources. *PLoS ONE* 8:e61265. doi: 10.1371/journal.pone.0061265

Frankel, O. H. (1984). "Genetic perspectives of germplasm conservation," in *Genetic Manipulation: Impact on Man and Society*, eds W. K. Arber, K.

Llimensee, and W. J. Peacock (Cambridge: Cambridge University Press), 161–170.

Fuller, G., Diamond, M. J., and Applewhite, T. H. (1967). High-oleic safflower oil. Stability and chemical modification. *J. Am. Oil Chem. Soc.* 44, 264–266. doi: 10.1007/BF02639272

Gouesnard, B., Bataillon, T. M., Decoux, G., Rozale, C., Schoen, D. J., and David, J. L. (2001). MSTRAT: an algorithm for building germplasm core collections by maximizing allelic or phenotypic richness. *J. Hered.* 92, 93–94. doi: 10.1093/jhered/92.1.93

Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 4:9.

Hanelt, P. (1961). Zur Kenntnis von *Carthamus tinctorius* L. *Kulturpflanze* 9, 114–145. (In German). doi: 10.1007/BF02095747

Hu, J., Zhu, J., and Xu, H. M. (2000). Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. *Theor. Appl. Genet.* 101, 264–268. doi: 10.1007/s001220 051478

Huaman, Z., Aguilar, C., and Ortiz, R. (1999). Selecting a Peruvian sweetpotato core collection on the basis of morphological, eco-geographical, and disease and pest reaction data. *Theor. Appl. Genet.* 98, 840–844. doi: 10.1007/s001220051142

Johnson, R. C., Ghorpade, P. B., Bradley, V. L., Bergman, J. W., and Henning Mündel, H. (2001). "Evaluation of the USDA core safflower collection for seven quantitative traits," in *Proceedings of the 5th International Safflower Conference, Williston, North Dakota and Sidney, Montana, USA, 23–27 July, 2001. Safflower: A Multipurpose Species with Unexploited Potential and World Adaptability* (Department of Plant Pathology, North Dakota State University), 149–152.

Johnson, R. C., Stout, D. M., and Bradley, V. L. (1993). "The U.S. collection: a rich source of safflower germplasm," in *Proceedings of the Third International Safflower Conference*, eds L. Dajue and H. Yuanzhou (Beijing: Botanical Garden; Institute of Botany; Chinese Academy of Sciences), 202–208.

Khan, M. A., von Witzke-Ehbrecht, S., Maass, B. L., and Becker, H. C. (2009). Relationships among different geographical groups, agro-morphology, fatty acid composition and RAPD marker diversity in safflower (*Carthamus tinctorius* L.). *Genet. Resour. Crop Evol.* 56, 19–30. doi: 10.1007/s10722-008-9338-6

Kim, K. W., Chung, H. K., Cho, G. T., Ma, K. H., Chandrabalan, D., Gwag, J. G., et al. (2007). PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. *Bioinformatics* 23, 2155–2162. doi: 10.1093/bioinformatics/btm313

Knowles, P. F. (1969). Centers of plant diversity and conservation of crop germ plasm: safflower. *Econ. Bot.* 23, 324–329. doi: 10.1007/BF02 860678

Knowles, P. F., and Bill, A. B. (1964). Inheritance of fatty acid content in the seed oil of a safflower introduction from Iran. *Crop Sci.* 4, 406–409. doi: 10.2135/cropsci1964.0011183X000400040023x

Knowles, P. F. (1989). "Safflower," in *Oil Crops of the World*, eds R. K. Downey, G. Robblen, and A. Ashri (New York, NY: Mcgraw-Hill), 363–374.

Kumar, S., Ambreen, H., Murali, T. V., Bali, S., Agarwal, M., Kumar, A., et al. (2015). Assessment of genetic diversity and population structure in a global reference collection of 531 accessions of *Carthamus tinctorius* L. (Safflower) using AFLP markers. *Plant Mol. Biol. Rep.* 33, 1299–1313. doi: 10.1007/s11105-014-0828-8

Kupsow, A. J. (1932). The geographical variability of the species *Carthamus tinctorius* L. *Bull. Appl. Bot. Genet. Plant Breed.* 1, 99–181.

Le Cunff, L., Fournier-Level, A., Laucou, V., Vezzulli, S., Lacombe, T., Adam-Blondon, A.-F., et al. (2008). Construction of nested genetic core collections to optimize the exploitation of natural diversity in *Vitis vinifera* L. subsp. sativa. *BMC Plant Biol.* 8:31. doi: 10.1186/1471-2229-8-31

Li, Y., Shi, Y., Cao, Y., and Wang, T. (2005). Establishment of a core collection for maize germplasm preserved in Chinese National Genebank using geographic distribution and characterization data. *Genet. Resour. Crop Evol.* 51, 845–852. doi: 10.1007/s10722-005-8313-8

Liu, W., Shahid, M. Q., Bai, L., Lu, Z., Chen, Y., Jiang, L., et al. (2015). Evaluation of genetic diversity and development of a core collection of wild rice (*Oryza rufipogon* Griff.) populations in China. *PLoS ONE* 10:e0145990. doi: 10.1371/journal.pone.0145990

Mahalakshmi, V., Ng, Q., Lawson, M., and Ortiz, R. (2007). Cowpea [*Vigna unguiculata* (L.) Walp.] core collection defined by geographical, agronomical and botanical descriptors. *Plant Genet. Resour.* 5, 113–119. doi: 10.1017/S1479262107837166

Mahasi, M. J., Pathak, R. S., Wachira, F. N., Riungu, T. C., Kinyua, M. G., and Waweru, J. K. (2006). Genotype by environment (GxE) interaction and stability in safflower (*Carthamus tinctorious* L.). *Asian J. Plant Sci.* 5, 1017–1021. doi: 10.3923/ajps.2006.1017.1021

McKhann, H. I., Camilleri, C., Bérard, A., Bataillon, T., David, J. L., Reboud, X., et al. (2004). Nested core collections maximizing genetic diversity in *Arabidopsis thaliana*. *Plant J.* 38, 193–202. doi: 10.1111/j.1365-313X.2004.02034.x

McPherson, M. A., Yang, R. C., Good, A. G., Nielson, R. L., and Hall, L. M. (2009). Potential for seed-mediated gene flow in agroecosystems from transgenic safflower (*Carthamus tinctorius* L.) intended for plant molecular farming. *Transgenic Res.* 18, 281–299. doi: 10.1007/s11248-00 8-9217-0

Mündel, H. H., and Bergman, J. W. (2009). "Safflower," in *Oil Crops*, eds J. Vollmann and I. Rajcan (New York, NY: Springer), 423–447.

Nimbkar, N. (2008). "Issues in safflower production in India," in *Safflower: Unexploited Potential and World Adaptability. Proceedings of the Seventh International Safflower Conference* (Wagga Wagga, NSW).

Noirot, M., Hamon, S., and Anthony, F. (1996). The principal component scoring: a new method of constituting a core collection using quantitative data. *Genet. Resour. Crop Evol.* 43, 1–6. doi: 10.1007/BF00126934

Patil, V. D., Reddy, M. V., and Nerkar, Y. S. (1994). Efficiency of early generation selections for yield and related characters in safflower (*Carthamus tinctorius* L.). *Theor. Appl. Genet.* 89, 293–296. doi: 10.1007/bf00225157

Rao, V. R., and Hodgkin, T. (2002). Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell Tissue Organ Cult.* 68, 1–19. doi: 10.1023/A:1013359015812

Ranga Rao, V., Ramachandram, M., and Arunachalam, V. (1977). An analysis of association of components of yield and oil in safflower (*Carthamus tinctorius* L.). *Theor. Appl. Genet.* 50, 185–191.

Rubis, D. D., Bergman, J. W., and Henning Mündel, H. (2001). "Developing new characteristics during 50 years of safflower breeding," in *Proceedings of the 5th International Safflower Conference, Williston, North Dakota and Sidney, Montana, USA, 23–27 July, 2001. Safflower: A Multipurpose Species with Unexploited Potential and World Adaptability* (Department of Plant Pathology, North Dakota State University), 109–111.

Schoen, D. J., and Brown, A. H. (1993). Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. *Proc. Natl. Acad. Sci. U.S.A.* 90, 10623–10627. doi: 10.1073/pnas.90.22.10623

Shehzad, T., Okuizumi, H., Kawase, M., and Okuno, K. (2009). Development of SSR-based sorghum (*Sorghum bicolor* (L.) Moench) diversity research set of germplasm and its evaluation by morphological traits. *Genet. Resour. Crop Evol.* 56, 809–827. doi: 10.1007/s10722-008-9403-1

Singh, V., Deshpande, M. B., Choudhari, S. V., and Nimbkar, N. (2004). Correlation and path coefficient analysis in safflower (*Carthamus tinctorius* L.). *Seasame Safflower Newsl.* 19, 77–81.

SPSS Inc. Released (2009). *PASW Statistics for Windows, Version 18.0*. Chicago: SPSS Inc.

Tai, P. Y. P., and Miller, J. D. (2001). A core collection for L. from the World Collection of Sugarcane. *Crop Sci.* 41, 879–885. doi: 10.2135/cropsci2001.413879x

Upadhyaya, H. D., and Ortiz, R. (2001). A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theor. Appl. Genet.* 102, 1292–1298. doi: 10.1007/s00122-001-0556-y

Upadhyaya, H. D., Ortiz, R., Bramel, P. J., and Singh, S. (2003). Development of a groundnut core collection using taxonomical, geographical and morphological descriptors. *Genet. Resour. Crop Evol.* 50, 139–148. doi: 10.1023/A:1022945715628

Upadhyaya, H. D., Pundir, R. P. S., Dwivedi, S. L., Gowda, C. L. L., Reddy, V. G., and Singh, S. (2009). Developing a mini core collection of sorghum

for diversified utilization of germplasm. *Crop Sci.* 49, 1769–1780. doi: 10.2135/cropsci2009.01.0014

van Hintum, T. J., Brown, A. H. D., Spillane, C., and Hodgkin, T. (2000). "Core collections of plant genetic resources," in *IPGRI Technical Bulletin No. 3* (Rome: International Plant Genetic Resources Institute).

Wang, L., Guan, Y., Guan, R., Li, Y., Ma, Y., Dong, Z., et al. (2006). Establishment of Chinese soybean *Glycine max* core collections with agronomic traits and SSR markers. *Euphytica* 151, 215–223. doi: 10.1007/s10681-006-9142-3

Weiss, E. A. (1983). "Safflower," in *Oilseed Crops,* ed E. A. Weiss (London: Longman), 216–281.

Yeh, F. C., Yang, R. C., Boyle, T., Ye, Z. H., and Mao, J. X. (1999). *POPGENE, Version 1.32: The User Friendly Software for Population Genetic Analysis.* Edmonton, AB: Molecular Biology and Biotechnology Centre, University of Alberta.

Zhang, C. Y., Chen, X. S., Zhang, Y. M., Yuan, Z. H., Liu, Z. C., Wang, Y. L., et al. (2009). A method for constructing core collection of *Malus sieversii* using molecular markers. *Sci. Agric. Sin.* 42, 597–604.

Zhang, Y., Zhang, X., Che, Z., Wang, L., Wei, W., and Li, D. (2012). Genetic diversity assessment of sesame core collection in China by phenotype and molecular markers and extraction of a mini-core collection. *BMC Genet.* 13:102. doi: 10.1186/1471-2156-13-102

Zhang, Z., and Johnson, R. C. (1999). *Safflower Germplasm Collection Directory.* Available online at: http://www.safflower.wsu.edu/saf_dir.pdf

Trichome-Related Mutants Provide a New Perspective on Multicellular Trichome Initiation and Development in Cucumber (*Cucumis sativus* L)

Xingwang Liu [1,2], Ezra Bartholomew [1,2], Yanling Cai [2] and Huazhong Ren [1,2]*

[1] College of Horticulture, China Agricultural University, Beijing, China, [2] Beijing Key Laboratory of Growth and Developmental Regulation for Protected Vegetable Crops, China Agricultural University, Beijing, China

Trichomes are specialized epidermal cells located in aerial parts of plants that function in plant defense against biotic and abiotic stresses. The simple unicellular trichomes of *Arabidopsis* serve as an excellent model to study the molecular mechanism of cell differentiation and pattern formation in plants. Loss-of-function mutations in *Arabidopsis thaliana* have suggested that the core genes *GL1* (which encodes a MYB transcription factor) and *TTG1* (which encodes a WD40 repeat-containing protein) are important for the initiation and spacing of leaf trichomes, while for normal trichome initiation, the genes *GL3*, and *EGL3* (which encode a bHLH protein) are needed. However, the positive regulatory genes involved in multicellular trichrome development in cucumber remain unclear. This review focuses on the phenotype of mutants (*csgl3*, *tril*, *tbh*, *mict*, and *csgl1*) with disturbed trichomes in cucumber and then infers which gene(s) play key roles in trichome initiation and development in those mutants. Evidence indicates that *MICT*, *TBH*, and *CsGL1* are allelic with alternative splicing. *CsGL3* and *TRIL* are allelic and override the effect of *TBH*, *MICT*, and *CsGL1* on the regulation of multicellular trichome development; and affect trichome initiation. *CsGL3*, *TRIL*, *MICT*, *TBH*, and *CsGL1* encode HD-Zip proteins with different subfamilies. Genetic and molecular analyses have revealed that *CsGL3*, *TRIL*, *MICT*, *TBH*, and *CsGL1* are responsible for the differentiation of epidermal cells and the development of trichomes. Based on current knowledge, a positive regulator pathway model for trichome development in cucumber was proposed and compared to a model in *Arabidopsis*. These data suggest that trichome development in cucumber may differ from that in Arabidopsis.

Keywords: unicellular, *arabidopsis*, multicellular, cucumber, mutants, trichome-related genes, regulator pathway

INTRODUCTION

Plant trichomes are highly specialized epidermal protrusions that are located on the surfaces of leaves, stems, petioles, sepals, seed coats, and other aerial organs. Their diversity is almost as great as the number of species on which they are found. Morphologically, they can be unicellular or multicellular as well as secretory glandular or non-glandular (Hülskamp et al., 1998; Hülskamp, 2004; Tissier, 2012; Chen et al., 2014). In the model plant *Arabidopsis thaliana*,

extensive studies have been performed on unicellular trichome development, especially on leaves (Hülskamp, 2004; Ishida et al., 2008; Pesch and Hülskamp, 2009). Classical molecular genetic approaches have identified several regulators that work in distinct developmental processes, such as trichome initiation/formation, endo-reduplication, and branch construction, and growth orientation (Schwab et al., 2000; Szymanski et al., 2000; Chen et al., 2014). The regulatory pathways of unicellular trichome development comprise both positive (mutants develop fewer trichomes) and negative (mutants develop more and/or clusters of trichomes) transcription factors (Ishida et al., 2008; Balkunde et al., 2010; Grebe, 2012). The crucial positive transcription factors belong to three protein classes: one WD40-repeat protein TRANSPARENT TESTA GLABRA1 (TTG1) (Galway et al., 1994; Walker et al., 1999); four basic helix-loop-helix (bHLH) proteins: GLABRA3 (GL3)m ENHANCER OF GLABRA3 (EGL3) (Payne et al., 2000; Zhang et al., 2003), TRANSPARENT TESTA (TT8) (Zhang et al., 2003), and MYC-1 (Zhao et al., 2012); and three R2-R3 type-MYB transcription factors: GLABRA1 (GL1), MYB23 and MYB5 (Oppenheimer et al., 1991; Li et al., 2009; Tominaga-Wada et al., 2012). They bind together to form an MYB-bHLH-WD40 (MBW) trimeric complex that activates the downstream gene GL2 (*GLABRA2*), which initiates trichome differentiation (Pesch and Hülskamp, 2009). Moreover, several small, single-repeat MYB-negative regulatory proteins, such as TRIPTYCHON (TRY), CAPRICE (CPC), ENHANCER OF TRYAND CPC1 (ETC1), ECT2, ETC3, CAPRICE-LIKE MYB3, and TRICHOMELESS1, and TRICHOMELESS2 (TCL1, and TCL2, respectively), have been shown to act in a non-cell-autonomous manner (Wada et al., 1997, 2002; Esch et al., 2004; Kirik et al, 2004; Pesch and Hülskamp, 2009; Wester et al., 2009; Edgar et al., 2014; Hauser, 2014; Wang and Chen, 2014). They compete with the R2R3 MYB protein GL1 and bind to bHLH proteins, including GL3/EGL3, to suppress trichome initiation in adjacent cells (Wang et al., 2014).

Cucumber (*Cucumis sativus* L.), as one of the most important vegetable crops, is also covered with trichomes ranging from the stems and leaves to the flowers, branches, fruits, and tendrils. During early fruit development, deep ridges along the length of the fruit cover the fruit surface, and densely spaced fruit trichomes are randomly scattered relative to the ridges (Liu et al., 2011; Chen et al., 2014). Trichomes on cucumber fruit are called spines. In cucumber, the fruit spine combines with tubercles to form the warty fruit trait, which is a very important fruit quality trait (Chen et al., 2014; Li et al., 2015). Compared to warty fruit, smooth fruit, with no spines, or tubercles, is very important for the breeding of freshly eaten cucumber types, as they are easier to clean, package, transport and store (Zhang et al., 2010; Yang et al., 2014). Moreover, smooth fruit is becoming increasingly popular due to its attractive

and distinctly shiny appearance (Li et al., 2015; Pan et al., 2015; Cui et al., 2016). Despite the importance of fruit spines in cucumber breeding for external quality, there are limited reports of the regulation of fruit spine development and few detailed characterizations of cucumber genes with disturbed trichome development. To understand trichome development in cucumber, we must determine the crucial regulators for its initiation and formation based on cucumber trichome-related mutants.

We begin this review by focusing on cucumber trichome-related mutants and the role that they play in trichome formation in plants with multicellular trichomes. Several mutants, such as *trichome-less* (*tril*) (Wang et al., 2016), *glabrous 3* (*csgl3*) (Pan et al., 2015; Cui et al., 2016), *tiny branched hair* (*tbh*) (Chen et al., 2014), *micro-trichome* (*mict*) (Zhao et al., 2015a), and *glabrous 1* (*csgl1*) (Li et al., 2015), inhibit trichome development via different mechanisms, but all of these mutants cause a reduction in one type of spine. The possible mechanisms involved in modulating trichome development in cucumber mutants are summarized herein.

MUTANTS WITH DISTURBED TRICHOME IN CUCUMBER

Five trichome-related mutants have been reported in cucumber (Chen et al., 2014; Li et al., 2015; Pan et al., 2015; Zhao et al., 2015a; Cui et al., 2016; Wang et al., 2016). Wild-type cucumber fruits have two types of trichomes, both of which are multicellular (**Figures 1**, **2**) (Chen et al., 2014). Type I trichomes are small, glandular trichomes, with a 3- to 5-cell base topped with a 4- to 8-cell head. Type II trichomes, which predominate, are much larger, non-glandular trichomes that are composed of a base and stalk (**Figure 2**). Compared to wild-type cucumber plants, all mutants appeared to be glabrous with no noticeable trichomes on the leaves, stems, tendrils, floral organs, or fruits (**Table 1**, **Figures 1**, **2**). Each mutant has a trichome type except for the mutants *tril* and *csgl3* (**Figure 2**). Trichomes have been reclassified into three morphologies according to their shape on leaves and fruit surfaces. Type I trichomes, which are found in the *tbh, mict*, and *csgl1* mutants, have a small papillar-shaped head (**Figures 2A,D,F**). Type II trichomes, which exist in the *tbh* and *mict* mutants, consist of one to five rounded cells without the pyramid-shaped head or pie-shaped base (**Figures 2B,E**). However, type III trichomes occur only in *tbh* mutants; this type of trichome was not previously described. Here, we classified these trichomes into the new type III according to the branching at the top of the trichome (**Figure 2C**). Trichome phenotypes from *tbh*, *mict* and *csgl1* indicated that these three mutants are all involved in trichome development and not in trichome initiation. Interestingly, the *tril* and *csgl3* mutants showed a completely glabrous phenotype on the shoot and fruit epidermis (**Figures 2G,H**), suggesting that *TRIL* and *CsGL3* may be upstream positive regulators of *TBH*, *MICT*, and *CsGL1* for the regulation of multicellular trichome development and may affect epidermal cell initiation.

Abbreviations: TRIL, trichome-less; TBH, tiny branched hair; MICT, micro-trichome; TTG1, transparent testa glabra1; bHLH, basic helix-loop-helix; GL3, glabra3; EGL3, enhancer of glabra3; MBW, MYB-bHLH-WD40; TRY, triptychon; CPC, caprice; ETC1, enhancer of tryand caprice1; HD-Zip, homeodomain-leucine zipper; PDF2, protodermal factor2; ATML1, *Arabidopsis thaliana* meristem layer1; MADS, M̲CM1-A̲GAMOUS-D̲EFICIENS-S̲RF4.

FIGURE 1 | Phenotypic characterization of mutants and the wild type (WT). (A) WT and *tril* plants; **(B)** Leaves; **(C)** Fruits; **(D)**. Fruit 0 days after anthesis; **(E)** Stems; **(F)** Tendril. Picture **(A)** is from Wang et al. (2016); Pictures **(B,D)** are from Chen et al. (2014); and Pictures **(C,E)** are from Li et al. (2015); Picture **(F)** cited form Pan et al. (2015).

Compared to the multicellular trichomes found on aerial organs, such as leaves and fruits, the trichomes found on underground organs, such as root hairs, are characterized as single-celled, unbranched, elongated, and soft-structured with small tumors attached. Based on the published results, there was no difference between the wild-type and the *tbh*, *mict*, and *csgl1* mutants (Chen et al., 2014; Li et al., 2015; Zhao et al., 2015a), indicating that the *TBH*, *MICT*, and *CsGL1* mutants may not

be involved in root hair formation. In contrast, the root length and number of branches of the *tril* mutant increased, suggesting that root hair formation in cucumber might be regulated by *TRIL* (Wang et al., 2016).

Other remarkable differences (except for the common phenotype) may exist among each mutant, such as dwarfism, branching, leaf curvature, and petal opening rates. In *mict* and *tbh*, rounded-head trichomes were found on the hypocotyls,
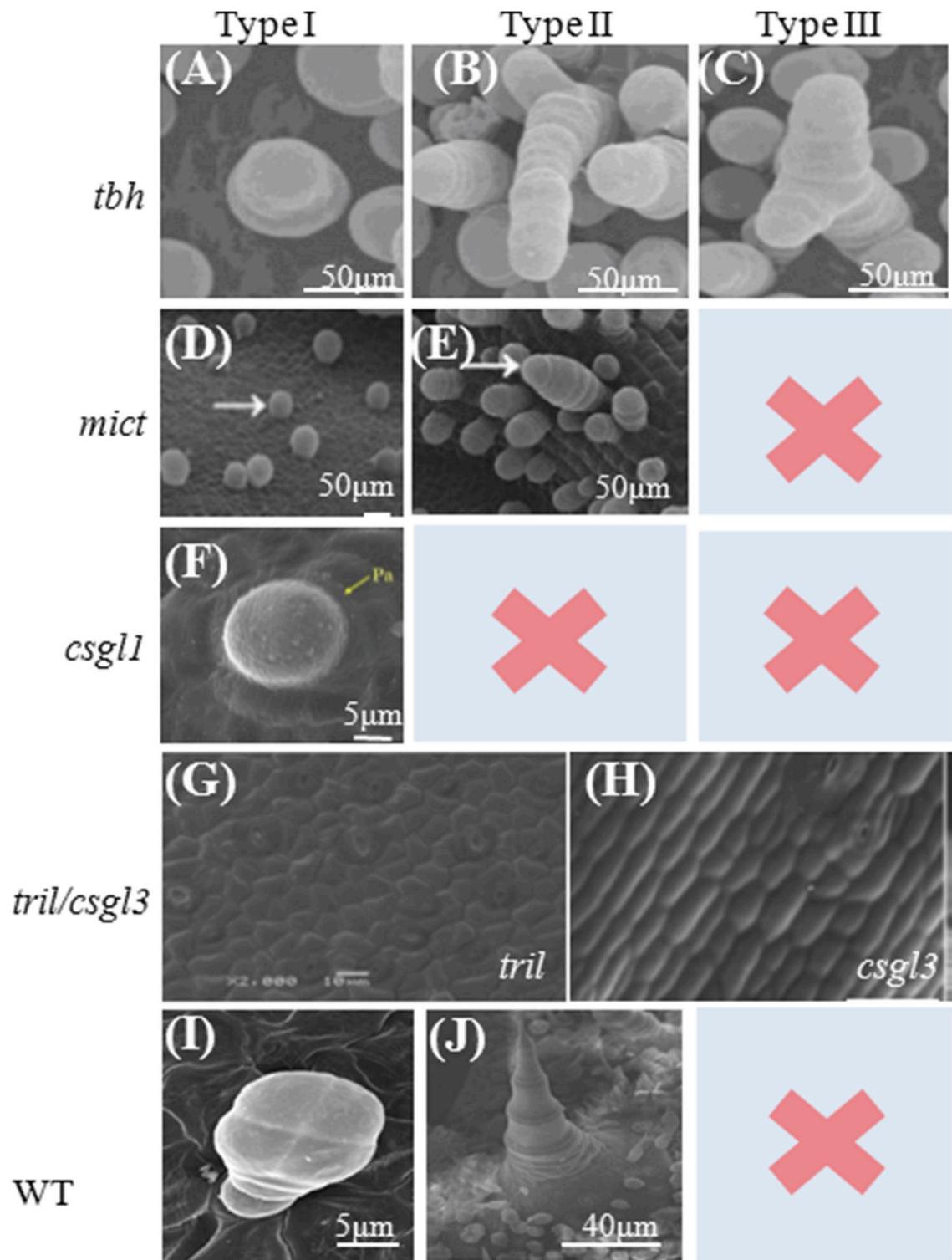
FIGURE 2 | Scanning Electron Microscopy images of trichomes of cucumber mutants. (A–C) SEM images of fruit spines **(A,B)** and leaf trichomes **(C)** in *tbh*, three morphologies from left to right; **(D–E)**. Micro-trichomes on a *mict* leaf. Arrows indicate the two morphologies: type I (right) and type II (left), respectively; **(F)**. Trichomes on the epidermis of leaves from *csgl1*; **(G,H)**. The *tril/csgl3* mutant has a completely glabrous phenotype, without trichome material in fruit epidermis. Trichome on the Fruit from the Wild type as control **(I,J)**. Pictures **(A–C)** are from (Chen et al., 2014); Pictures **(D,E,H)** are from (Zhao et al., 2015a); Picture **(F)** is from (Li et al., 2015), and Picture **(G)** is from (Wang et al., 2016).

TABLE 1 | Previous morphological studies of trichome in *tbh*, *csgl1*, *mict*, *tril* and *csgl3*.

Mutant name	Origin	Types of trichome described	References
tbh	Northern China	Type I: tiny 3- to 5-cell base topped with a 4- to 8-cell head Type II: non-glandular and branchless	Chen et al., 2014
mict	Northern China	Type I: only a small papillar-shaped head Type II: 1-5 cells without the pyramid shaped head	Zhao et al., 2015a,b
csgl1	Northern China	No trichomes on leave except glandular trichomes	Li et al., 2015
tril	European greenhouse	No any types trichome	Wang et al., 2016
csgl3	European greenhouse	No any types trichome	Pan et al., 2015

but none were found in *csgl1*, suggesting trichome distribution is specific (Chen et al., 2014; Li et al., 2015; Zhao et al., 2015a).

TBH, MICT, AND *CSGL1* ARE THE ALLELIC WITH ALTERNATIVE SPLICING

To decipher the molecular defects in the *mict*, *csgl1*, and *tbh* mutants, a map-based cloning approach was undertaken by three independent groups to isolate these genes. The results showed that *TBH*, *MICT*, and *CsGL1* are allelic and that there was a 2649 bp fragment deletion from -189 to 2460 bp of the start codon in *Csa3M748220* in the *mict*, *csgl1*, and *tbh* mutants (Li et al., 2015; Zhao et al., 2015a) (**Figure 3A**). According to the gene ID that provided in the references, we extracted their proteins from the cucumber genome database (http://www.icugi.org/cgi-bin/ICuGI/genome/search.cgi) and analyzed their protein domains online (https://blast.ncbi.nlm.nih.gov/Blast.cgi). These genes were predicted to encode a class I homeodomain-leucine zipper (HD-Zip) protein consisting of different amino acid residues. For example, CsGL1 and TBH contain 240 amino acids with the conserved HD (65AA-120AA) and Zip domains (121AA-164AA), but the MICT protein consists of 242 amino acid residues with a HALZ motif. Notably, there were two isoforms of this gene in the cucumber genome database (Csa3M748820.1 and Csa3M748820.2); therefore, we infer that the different protein lengths encoded by the same gene suggest that *Csa3M748220* may exist by alternative splicing in cucumber. Subcellular localization showed that the *Csa3M748220* coding sequence was fused to GFP (35S) in the nuclei of tobacco and onion epidermal cells (Zhao et al., 2015a,b). Moreover, a transcriptional activation activity assay in yeast found that *Csa3M748220* had weak activity as a transcriptional activator (Zhao et al., 2015a). Therefore, based on the above results, *Csa3M748220* has the typical features of a transcription factor. HD-Zip proteins are unique to the plant kingdom and can be classified into four groups, I-IV, based on their distinctive traits of DNA-binding specificities, gene structures, and common motifs (Abe et al., 2003; Hülskamp et al., 2005). HD-Zip I genes have been demonstrated to be involved largely in biological processes, such as abiotic stress responses, meristem regulation,and trichome development (Hanson et al., 2001; Himmelbach et al., 2002; Hjellström et al., 2003; Saddicl et al., 2006; Zhao et al., 2015a,b).

TRIL, CSGL3 ARE THE ALLELIC WHICH OVERRIDE THE EFFECT OF *TBH/MICT/CSGL1* ON THE REGULATION OF MULTICELLULAR TRICHOME INITIATION AND DEVELOPMENT

Thereare few reports on the regulatory genes that control multicellular trichome development in cucumber. This presents a good opportunity to use these mutants to analyze their regulatory mechanisms in trichomes of cucumber. As mentioned above, *tril* and *csgl3* show a completely different glabrous morphology from that of the other three trichome-development-related mutants, and only the epidermal cells, including stomata and encircling guard cells, were visible (**Figure 2G**). This indicates that *TRIL* and *CsGL3* function in trichome cell fate determination. A transcriptome profiling analysis among wild-type, *tril*, *tbh*, *mict*, and *csgl1* revealed that *TBH*, *MICT*, and *CsGL1* were not expressed in the *tril* mutant (baseMean 0) but were highly expressed in the wild-type (baseMean 518.25) (Zhao et al., 2015a). *TRIL* was mapped to *Cas6M514870*, a member of the class IV HD-Zip family that shares 66.7% identity with *PROTODERMAL FACTOR2* (*PDF2*, *At4g04890.1*), a shoot epidermal cell differentiation-related gene, and 35% identity with GL2 (*At1g79840.2*), a gene that initiates trichome differentiation in *Arabidopsis*. In the *tril* mutant, there is a 5008 bp insertion fragment after the first exon (**Figure 3B**). *CsGL3* was also mapped to *Csa6M514870* by Pan and Cui (Pan et al., 2015; Cui et al., 2016). In Pan's study, the loss of function of the *CsGL3* was due to the insertion of a 5-kb-long terminal repeat (LTR) retrotransposon in the 4th exon of *CsGL3* (**Figure 3B**). The insertion location is different from that of *tril*. Cui used three markers (InDel-19, dCAPs-2, and dCAPs-11) designed from the sequence of *Csa6M514870*, which co-segregated with the trait. In addition, *Csa6M514870* was found to harbor 3 single-base substitutions: T→C (611 bp), G→A (820 bp), and G→A (865 bp) at the fourth exon, resulting in a change in the amino acid sequence. We still do not know how the *CsGL3* changed in its corresponding mutant *csgl3*.

In the *tbh*, *mict*, and *csgl1* mutants, *TRIL/CsGL3* was over-expressed at different levels due to the different stages at which the samples were collected (Chen et al., 2014; Li et al., 2015; Zhao et al., 2015a). Moreover, the trichome phenotype analyzed in the F2 population between the *tril* and *mict* mutants suggests that the *TRIL* gene has a significant influence on trichome initiation and in determining the fate of epidermal cells,
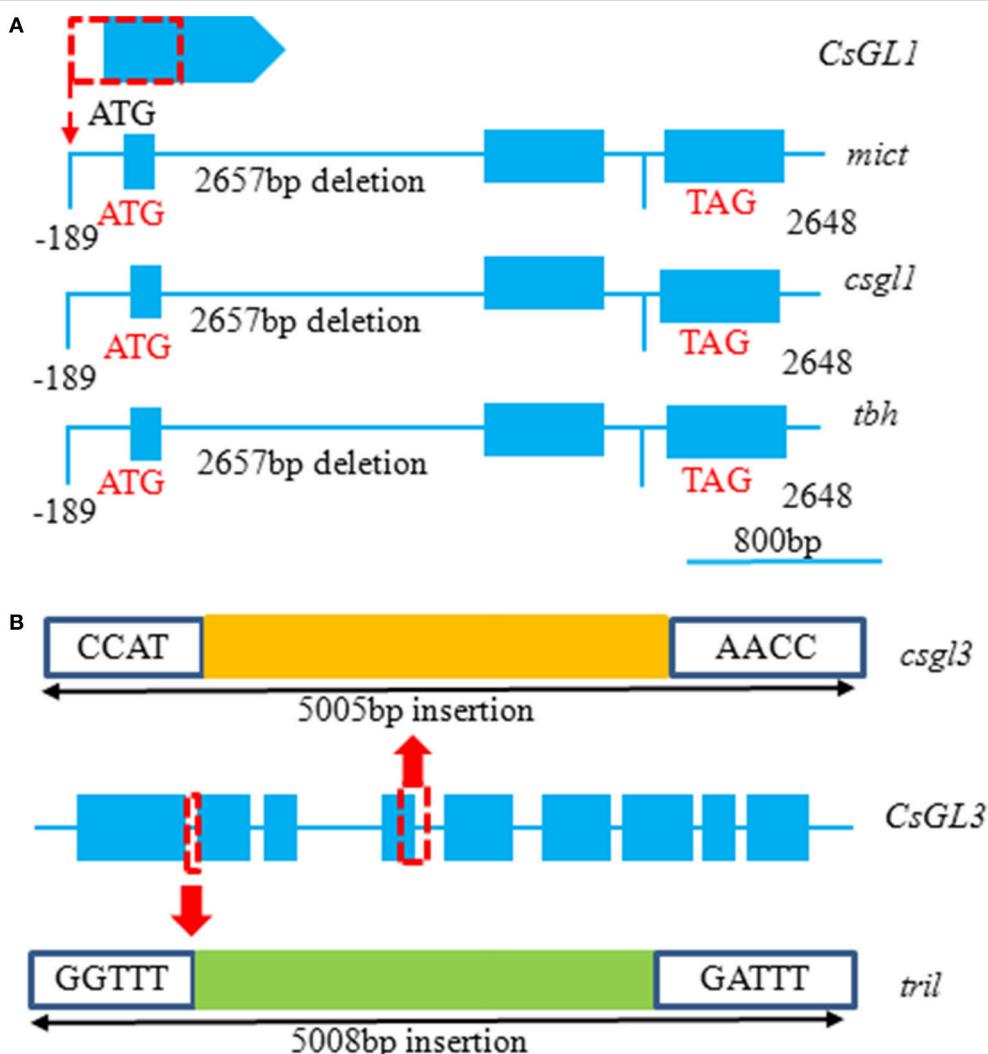
FIGURE 3 | Gene structure of *CsGL1* and *CsGL3* in the WT and mutants. (A) A 2649 bp deletion of the start codon in *CsGL1* in the *mict*, *csgl1 and tbh* mutants compared to the WT. Boxes represent exons, and lines represent introns and intergenic regions. ATG and TAG indicate start and stop codons, respectively, and numbers represent nucleotide deletion length in mutants. *Mict* is from Zhao et al. (2015a); *csgl1* is from Li et al. (2015). **(B)** A 5005 bp insertion at the 4th exon in *csgl3* and a 5008 bp insertion after the first exon in the *tril* mutant. The red box indicates insertion sites, and yellow and green boxes represent insertion length. Different insertions in this gene showed the same phenotype in cucumber. *tril* is from Wang et al. (2016); *csgl3* is from Pan et al. (2015).

whereas the *MICT/TBH/CsGL1* gene only influences trichome development at the shoot. Genetically, based on the double-mutant phenotype in *csgl3csgl1*, the *TRIL/GsGL3* gene is assumed to act upstream of the *MICT/TBH/CsGL1* gene and control the expression of *MICT/TBH/ CsGL1* (Zhao et al., 2015a; Wang et al., 2016).

MODELS FOR TRICHOME PATTERNING IN CUCUMBER MAY DIFFER FROM MODELS IN *ARABIDOPSIS*

In the model plant *Arabidopsis*, the activator-inhibitor model has guided intuitive modeling and experimental design for a long time because it offers a reasonable explanation for the apparently paradoxical situation in that trichome-promoting and trichome-inhibiting genes are both expressed strongly in trichomes (Pesch and Hülskamp, 2009). Here, we focus on a positive regulatory model for trichome development between *Arabidopsis* and cucumber. The main reason is that all of the mutants in cucumber reveal fewer or no trichomes visually, suggesting that they encode a positive regulator of trichome development. Solid evidence of the genetic basis of trichome initiation has identified genes that (a) control the entry into the trichome pathway and (b) control the spacing of initiation events. Loss-of-function mutations in *Arabidopsis thaliana* have suggested that the genes *GL1* (which encodes an MYB transcription factor) and *TTG1* (which encodes a WD40 repeat-containing protein) are important for the initiation and spacing of leaf trichomes (Galway et al., 1994; Walker et al., 1999), while for normal

trichome initiation, the genes GL3, and EGL3 (Payne et al., 2000; Zhang et al., 2003), which encode helix-loop-helix (bHLH) proteins, are needed (**Figure 4A**).

In cucumber, the *TRIL/CsGL3* gene encodes a HD-Zip IV protein. The loss-of-function mutant showed that *tril/csgl3* has a trichome-less phenotype, indicating that this gene is a positive regulator of trichome development. Its homologous genes PDF2 and *ARABIDOPSIS THALIANA MERISTEM LAYER1* (*ATML1*) in *Arabidopsis* also encode HD-Zip IV protein family members (Nakamura et al., 2006). The *pdf2atml1* double mutant displays striking defects in shoot epidermal cells (Abe et al., 2003). In tomato, another typical multicellular trichome plant, *Wo* RNAi transgenic tomato plants showed similar shoot and root epidermal cell formation as that of the *tril* mutant. *Wo* shares an amino acid sequence identical to that of *TRIL/CsGL3* and belongs to the same HD-Zip IV protein family. These results indicate that the role of *TRIL/CsGL3* in the initiation and control

of multicellular trichrome pathways. Based on previous results, a proposed model explaining trichome patterning in the positive regulator pathway in cucumber was built (**Figure 4B**). In this model, HD-Zip transcription factors may bind to other types of transcription factors to generate a specific complex to control cucumber multicellular trichome formation.

Based on transcriptional data from all of the mutants, several candidate genes should be focused on extensively (Chen et al., 2014; Zhao et al., 2015a,b). For example, the *CsMYB6*, *CsWIN1*, and *CsGL2* genes were down-regulated not only in *tril/csgl3* but also *in tbh*, *mict*, and *csgl*, indicating that those genes were involved in multicellular trichome development in cucumber.

DISCUSSIONS

In the past decade, substantial progress has been made in delineating the genes that control trichome development in
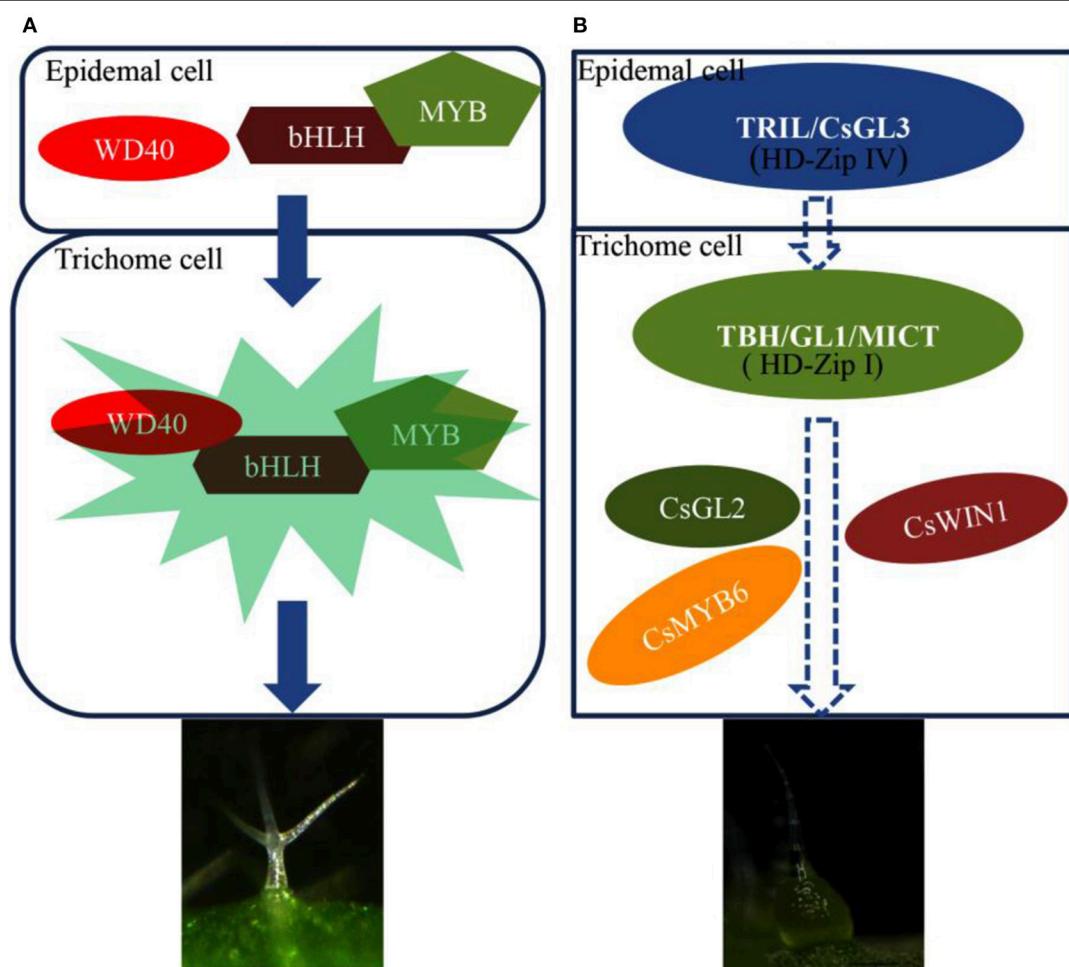
FIGURE 4 | Proposed model to explain trichome patterning in the positive regulator pathway in cucumber and in *Arabidopsis*. (A) *TRIL/CsGL3* encodes the HD-Zip IV protein and may directly interact with the HD-Zip I protein (MICT/TBH/CsGL1) to regulate trichome development in cucumber. *CsMYB6*, *CsWIN1*, and *CsGL2* may be located downstream of *MICT/TBH/CsGL1* and positively participate in trichome development. **(B)** In unicellular trichomes, WD40, Bhlh, and MYB form a complex to positively regulate trichome development. The core complex in MBW **(A)** may differ from the HD-Zip complex **(B)**. The bold arrows in the boxes indicate developmental progression.

cucumber. Several groups provide evidence to suggest that a number of transcriptional activators, such as HD-Zip I and HD-Zip IV, play a role in fine tuning the spatial and temporal distribution of trichomes. Researchers have tried to demonstrate the possible mechanisms of these transcriptional activators.

However, much remains unknown and needs to be elucidated in future research. The functions of the HD-Zip IV and HD-Zip I genes require further investigation through genetic transformation in cucumber plants. Our current knowledge about the gene regulatory networks is largely limited to the unicellular trichomes in the model plant *Arabidopsis*. However, little is known about the regulatory network that controls the development of multicellular trichomes in cucumber. Evidence indicates that the common regulatory mechanisms of unicellular trichomes in *Arabidopsis* or of multicellular trichomes in cotton involve plant-specific genes that function distinctively. An analysis of the differential expression data generated by RNA-seq can offer new information for identifying putative key multicellular development transcription factors in cucumber. This is why we focused on the critical transcription factor genes. A new set of more than 42 transcription factor genes, including Homeodomain, MCMI-AGAMOUS-DEFICIENS-SRF4 (MADS), and WRKY domains, has been identified in a transcriptome analysis of all cucumber trichome-related mutants (Zhao et al., 2015a). These transcription factor genes are unique to plants and are involved in a range of activities; many are associated with multicellular trichrome development and other species-specific development processes. The study of the interaction among those transcription factors is an emerging area of research because these factors probably share many biological functions for trichome development. All of this information warrants further investigation.

The aim of this review is to provide readers with a summary of the progress of cucumber trichome development and to encourage plant scientists to further investigate the mechanism of trichrome initiation and development and their regulatory network. We have summarized the mutants related to fruit trichomes, the key genes that control trichome initiation and development, gene relationships and a possible model for fruit trichome positive regulatory mechanisms that is different from *Arabidopsis* in core transcriptional factor numbers. All of this may help us to better understand the advances in the study of cucumber trichomes. At the same time, more investigations must be conducted.

Because cucumber is a horticultural crop of worldwide importance and fruit spines directly affect its commercial quality, an extensive characterization of cucumber trichomes will not only help us to understand the underlying molecular mechanisms involved in multicellular trichome development but will also pave the way for creating new cucumber varieties with desired trichome growth and density. Moreover, cucumber trichomes may serve as a model system for studying the development of multicellular trichomes.

AUTHOR CONTRIBUTIONS

HR and XL organized the review, wrote the first draft and generated **Figures 1–4**. EB and YC contributed to a second draft. All of the authors revised the manuscript multiple times. HR and XL performed the final revision of the manuscript, which was read and approved by all authors.

ACKNOWLEDGMENTS

REFERENCES

Abe, M., Katsumata, H., Komeda, Y., and Takahashi, T. (2003). Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in Arabidopsis. *Development* 130, 635–643. doi: 10.1242/dev.00292

Balkunde, R., Pesch, M., and Hülskamp, M. (2010). Trichome patterning in *Arabidopsis thaliana*: from genetic to molecular models. *Curr. Top. Dev. Biol.* 91, 299–321. doi: 10.1016/S0070-2153(10)91010-7

Chen, C., Liu, M., Jiang, L., Liu, X., Zhao, J., Yan, S., et al. (2014). Transcriptome profiling reveals roles of meristem regulators and polarity genes during fruit trichome development in cucumber (*Cucumis sativus* L.). *J. Exp. Bot.* 65, 4943–4958. doi: 10.1093/jxb/eru258

Cui, J., Miao, H., Ding, L., Wehner, T. C., Liu, P., Wang, Y., et al. (2016). A new glabrous Gene (csgl3) identified in trichome development in cucumber (*Cucumis sativus* L.). *PLoS ONE* 11:e0148422. doi: 10.1371/journal.pone.0148422

Edgar, B. A., Zielke, N., and Gutierrez, C. (2014). Endocycles: a recurrent evolutionary innovation for post-mitotic cell growth. *Nat. Rev. Mol. Cell Biol.* 15, 197–210. doi: 10.1038/nrm3756

Esch, J. J., Chen, M., Hillestad, M., and David, M. M. (2004). Comparison of *TRY* and the closely related *At1g01380* gene in controlling Arabidopsis trichome patterning. *Plant J.* 40, 860–869. doi: 10.1111/j.1365-313X.2004.02259.x

Galway, M. E., Masucci, J. D., Lloyd, A. M., Walbot, V., Davis, R. W., and Schiefelbein, J. W. (1994). The *TTG* gene is required to specify epidermal cell fate and cell patterning in the Arabidopsis root. *Dev. Biol.* 166, 740–754. doi: 10.1006/dbio.1994.1352

Grebe, M. (2012). The patterning of epidermal hairs in Arabidopsis—updated. *Curr. Opin. Plant Biol.* 15, 31–37. doi: 10.1016/j.pbi.2011.10.010

Hanson, J., Johannesson, H., and Engström, P. (2001). Sugar-dependent alterations in cotyledon and leaf development in transgenic plants expressing the HDZhdip gene ATHB13. *Plant Mol. Biol.* 45, 247–262. doi: 10.1023/A:1006464907710

Hauser, M. T. (2014). Molecular basis of natural variation and environmental control of trichome patterning. *Front. Plant Sci.* 5:320. doi: 10.3389/fpls.2014.00320

Himmelbach, A., Hoffmann, T., Leube, M., Höhener, B., and Grill, E. (2002). Homeodomain protein ATHB6 is a target of the protein phosphatase ABI1 and regulates hormone responses in Arabidopsis. *EMBO J.* 21, 3029–3038. doi: 10.1093/emboj/cdf316

Hjellström, M., Olsson, A. S., Engström, P., and Söderman, E. (2003). Constitutive expression of the water deficit-inducible homeobox gene *ATHB7* in transgenic Arabidopsis causes a suppression of stem elongation growth. *Plant Cell Environ.* 26, 1127–1136. doi: 10.1046/j.1365-3040.2003.01037.x

Hülskamp, J. A., Walker, A. R., Timmis, J. N., and Orford, S. J. (2005). Two WD-repeat genes from cotton are functional homologues of the *Arabidopsis thaliana*

TRANSPARENT TESTA GLABRA1 (*TTG1*) gene. *Plant Mol. Biol.* 57, 67–81. doi: 10.1007/s11103-004-6768-1

Hülskamp, M. (2004). Plant trichomes: a model for cell differentiation. *Nat. Rev. Mol. Cell Biol.* 5, 471–480. doi: 10.1038/nrm1404

Hülskamp, M., Schnittger, A., and Folkers, U. (1998). Pattern formation and cell differentiation: trichomes in Arabidopsis as a genetic model system. *Int. Rev. Cytol.* 186, 147–178. doi: 10.1016/S0074-7696(08)61053-0

Ishida, T., Kurata, T., Okada, K., and Wada, T. (2008). A genetic regulatory network in the development of trichomes and root hairs. *Annu. Rev. Plant Biol.* 59, 365–386. doi: 10.1146/annurev.arplant.59.032607.092949

Kirik, V., Simon, M., Wester, K., Schiefelbein, J., and Hulskamp, M. (2004). ENHANCER of TRYand CPC2 (ETC2) reveals redundancy in the region-specific control of trichome development of Arabidopsis. *Plant Mol. Biol.* 55, 389–398. doi: 10.1007/s11103-004-0893-8

Li, Q., Cao, C., Zhang, C., Zheng, S., Wang, Z., Wang, L., et al. (2015). The identification of *Cucumis sativus* Glabrous 1 (*CsGL1*) required for the formation of trichomes uncovers a novel function for the homeodomain-leucine zipper I gene. *J. Exp. Bot.* 66, 2515–2526. doi: 10.1093/jxb/erv046

Li, S., Milliken, O. N., Pham, H., Seyit, R., Napoli, R., Preston, J., et al. (2009). The Arabidopsis *MYB5* transcription factor regulates mucilage synthesis, seed coat development, and trichome morphogenesis. *Plant Cell* 21, 72–89. doi: 10.1105/tpc.108.063503

Liu, X., Wang, L., Liu, L., Guo, Y., and Ren, H. (2011). Alleviating effect of exogenous nitric oxide in cucumber seedling against chilling stress. *Afr. J. Biotechnol.* 10, 4380–4386. doi: 10.5897/AJB10.812

Nakamura, M., Katsumata, H., Abe, M., Yabe, N., Komeda, Y., Yamamoto, K. T., et al. (2006). Characterization of the class IV homeodomain-leucine zipper gene family in Arabidopsis. *Plant Physiol.* 141, 1363–1375. doi: 10.1104/pp.106.077388

Oppenheimer, D. G., Herman, P. L., Sivakumaran, S., Esch, J., and Marks, M. D. (1991). A *myb* gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. *Cell* 67, 483–493. doi: 10.1016/0092-8674(91)90523-2

Pan, Y., Bo, K., Cheng, Z., and Weng, Y. (2015). The loss-of-function GLABROUS 3 mutation in cucumber is due to LTR-retrotransposon insertion in a class IV HD-ZIP transcription factor gene *CsGL3* that is epistatic over *CsGL1*. *BMC Plant Biol.* 15:302. doi: 10.1186/s12870-015-0693-0

Payne, C. T., Zhang, F., and Lloyd, A. M. (2000). *GL3* encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. *Genetics* 156, 1349–1362.

Pesch, M., and Hülskamp, M. (2009). One, two, three…models for trichome patterning in Arabidopsis? *Curr. Opin. Plant Biol.* 12, 587–592. doi: 10.1016/j.pbi.2009.07.015

Saddicl, A., Huvermann, B., Bezhani, S., Su, Y., Winter, C. M., et al. (2006). The LEAFY target LMI1 is a meristem identity regulator and acts together with LEAFY to regulate expression of CAULIFLOWER. *Development* 133, 1673–1682. doi: 10.1242/dev.02331

Schwab, B., Folkers, U., Ilgenfritz, H., and Hülskamp, M. (2000). Trichome morphogenesis in Arabidopsis. *Philos Trans. Soc. Lond. B Biol. Sci.* 355, 879–883. doi: 10.1098/rstb.2000.0623

Szymanski, D. B., Lloyd, A. M., and Marks, M. D. (2000). Progress in the molecular genetic analysis of trichome initiation and morphogenesis in Arabidopsis. *Trends Plant Sci.* 5, 214–219. doi: 10.1016/S1360-1385(00)01597-1

Tissier, A. (2012). Glandular trichomes: what comes after expressed sequence tags? *Plant J.* 70, 51–68. doi: 10.1111/j.1365-313X.2012.04913.x

Tominaga-Wada, R., Nukumizu, Y., Sato, S., Kato, T., Tabata, S., and Wada, T. (2012). Functional divergence of *MYB*-related genes, *WEREWOLF* and *AtMYB23* in Arabidopsis. *Biosci. Biotechnol. Biochem.* 76, 883–887. doi: 10.1271/bbb.110811

Wada, T., Kurata, T., Tominaga, R., Koshino-Kimura, Y., Tachibana, T., Goto, K., et al. (2002). Role of a positive regulator of root hair development, CAPRICE, in Arabidopsis root epidermal cell differentiation. *Development* 129, 5409–5419. doi: 10.1242/dev.00111

Wada, T., Tachibana, T., Shimura, Y., and Okada, K. (1997). Epidermal cell differentiation in Arabidopsis determined by a Myb homolog, *CPC*. *Science* 277, 1113–1116. doi: 10.1126/science.277.5329.1113

Walker, A. R., Davison, P. A., Bolognesi-Winfield, A. C., James, C. M., Srinivasan, N., Blundell, T. L., et al. (1999). The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *Plant Cell* 11, 1337–1349. doi: 10.1105/tpc.11.7.1337

Wang, S., and Chen, J. G. (2014). Regulation of cell fate determination by single-repeat R3 MYB transcription factors in Arabidopsis. *Front. Plant Sci.* 5:133. doi: 10.3389/fpls.2014.00133

Wang, S., Wang, J. W., Yu, N., Li, C. H., Luo, B., et al. (2014). Control of plant trichome development by a cotton fiber MYB gene. *Plant Cell* 16, 2323–2334. doi: 10.1105/tpc.104.024844

Wang, Y., Nie, J., Chen, H., Guo, C., Pan, J., He, H., et al. (2016). Identification and mapping of Tril, a homeodomain-leucine zipper gene involved in multicellular trichome initiation in *Cucumis sativus*. *Theor. Appl. Genet.* 129, 305–316. doi: 10.1007/s00122-015-2628-4

Wester, K., Digiuni, S., Geier, F., Timmer, J., Fleck, C., and Hülskamp, M. (2009). Functional diversity of R3 single-repeat genes in trichome development. *Development* 136, 1487–1496. doi: 10.1242/dev.021733

Yang, X., Zhang, W., He, H., Nie, J., Bie, B., Zhao, J., et al. (2014). Tuberculate fruit gene Tu encodes a C2H2 zinc finger protein that is required for the warty fruit phenotype in cucumber (*Cucumis sativus* L.). *Plant J.* 78, 1034–1046. doi: 10.1111/tpj.12531

Zhang, F., Gonzalez, A., Zhao, M., and Payne, C. T. (2003). Lloyd A. A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. *Development* 130, 4859–4869. doi: 10.1242/dev.00681

Zhang, W., He, H., Guan, Y., Du, H., Yuan, L., Li, Z., et al. (2010). Identification and mapping of molecular markers linked to the tuberculate fruit gene in the cucumber (*Cucumis sativus* L.). *Theor. Appl. Genet.* 120, 645–654. doi: 10.1007/s00122-009-1182-3

Zhao, H., Wang, X., Zhu, D., Cui, S., Li, X., Cao, Y., et al. (2012). A single amino acid substitution in IIIf subfamily of basic helix-loop-helix transcription factor *AtMYC1* leads to trichome and root hair patterning defects by abolishing its interaction with partner proteins in Arabidopsis. *J. Biol. Chem.* 287, 14109–14121. doi: 10.1074/jbc.M111.280735

Zhao, J., Pan, J., Guan, Y., Nie, J., Yang, J., Qu, M., et al. (2015b). Transcriptome analysis in *Cucumis sativus* identifies genes involved in multicellular trichome development. *Genomics* 105, 296–303. doi: 10.1016/j.ygeno.2015.01.010

Zhao, J., Pan, J., Guan, Y., Zhang, W., Bie, B., Wang, Y., et al. (2015a). Micro-trichome as a class I homeodomain-leucine zipper gene regulates multicellular trichome development in *Cucumis sativus*. *J. Integr. Plant Biol.* 57, 925–935. doi: 10.1111/jipb.12345

Interactive Effects of Elevated [CO$_2$] and Water Stress on Physiological Traits and Gene Expression during Vegetative Growth in Four Durum Wheat Genotypes

Susan Medina [1,2], Rubén Vicente [1]*, Amaya Amador [3] and José Luis Araus [1]

[1] Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, Barcelona, Spain, [2] Crop Physiology Laboratory, International Crops Research Institute for Semi-Arid Tropics, Patancheru, India, [3] Unitat de Genòmica, Centres Científics i Tecnològics, Universitat de Barcelona, Barcelona, Spain

The interaction of elevated [CO$_2$] and water stress will have an effect on the adaptation of durum wheat to future climate scenarios. For the Mediterranean basin these scenarios include the rising occurrence of water stress during the first part of the crop cycle. In this study, we evaluated the interactive effects of elevated [CO$_2$] and moderate to severe water stress during the first part of the growth cycle on physiological traits and gene expression in four modern durum wheat genotypes. Physiological data showed that elevated [CO$_2$] promoted plant growth but reduced N content. This was related to a down-regulation of Rubisco and N assimilation genes and up-regulation of genes that take part in C-N remobilization, which might suggest a higher N efficiency. Water restriction limited the stimulation of plant biomass under elevated [CO$_2$], especially at severe water stress, while stomatal conductance and carbon isotope signature revealed a water saving strategy. Transcript profiles under water stress suggested an inhibition of primary C fixation and N assimilation. Nevertheless, the interactive effects of elevated [CO$_2$] and water stress depended on the genotype and the severity of the water stress, especially for the expression of drought stress-responsive genes such as dehydrins, catalase, and superoxide dismutase. The network analysis of physiological traits and transcript levels showed coordinated shifts between both categories of parameters and between C and N metabolism at the transcript level, indicating potential genes and traits that could be used as markers for early vigor in durum wheat under future climate change scenarios. Overall the results showed that greater plant growth was linked to an increase in N content and expression of N metabolism-related genes and down-regulation of genes related to the antioxidant system. The combination of elevated [CO$_2$] and severe water stress was highly dependent on the genotypic variability, suggesting specific genotypic adaptation strategies to environmental conditions.

Keywords: climate change, durum wheat, elevated [CO$_2$], genotypic variability, stable isotopes, transcript levels, vegetative growth, water stress

INTRODUCTION

Food security is facing new challenges nowadays due to the increase in the world population and the impacts of climate change on agriculture and food supply. Wheat is a very important crop for the human diet, ranking in fourth position in terms of the world's most important crops by production quantity after sugarcane, maize, and rice (FAO, 2013). Although bread wheat dominates global wheat production, durum wheat is an economically and culturally important staple crop in the Mediterranean region, used for the production of pasta, bread, burghul, couscous, and freekeh (Habash et al., 2009). In the second half of the twentieth century, local durum wheat landraces were replaced by improved semi-dwarf cultivars, which showed higher yield and harvest index (Soriano et al., 2016). In the early 1970s, introduction of germplasm from CIMMYT (International Maize and Wheat Improvement Centre) increased grain yield (Sanchez-Garcia et al., 2013). Improvement in wheat yield per unit area constitutes one of the largest challenges to be addressed by breeding programs, covering numerous research areas (McKersie, 2015). Projections of wheat production assume that the growth rate will be lower than the historical growth rates reported in the second half of the twentieth century (Bort et al., 2014; Nakhforoosh et al., 2015), with insignificantly higher yields in modern wheat genotypes released in recent years (Sanchez-Garcia et al., 2013). It is unlikely that any improvements will support the increase in world population or mitigate against future extreme weather events (Araus et al., 2002; Alexandratos and Bruinsma, 2012; Trnka et al., 2014).

Observations of the climate system confirm that Earth's mean surface temperature is increasing rapidly as a consequence of the anthropogenic emissions of CO$_2$ and other greenhouse gases (IPCC, 2013). The atmospheric concentration of CO$_2$ ([CO$_2$]) has increased by more than 40% since the beginning of the industrial revolution and is expected to double by the end of this century (IPCC, 2013). As atmospheric [CO$_2$] is currently a limiting factor for C$_3$ photosynthesis, the primary effect of a short-term exposure to elevated [CO$_2$] includes an initial stimulation of photosynthesis due to both enrichment of substrate for ribulose bisphosphate carboxylase oxygenase (Rubisco) carboxylation and inhibition of competitive Rubisco oxygenation which may eventually contribute to a higher biomass (Stitt and Krapp, 1999; Long et al., 2006). High [CO$_2$] also induces a stomatal closure leading to a better leaf water status. However, growth over the long-term under elevated [CO$_2$] leads to a down-regulation of photosynthetic capacity, which has been related to a decline in Rubisco protein content and activity, together with a higher carbohydrate accumulation and a decline in N concentration and protein content in wheat (Aranjuelo et al., 2011, 2013; Vicente et al., 2015a,b). This phenomenon suggests that regulatory mechanisms may occur in the plant, e.g., end-product inhibition, carbon sink limitation, biomass dilution effects, or a decline in nutrient uptake and/or assimilation (Stitt and Krapp, 1999; Vicente et al., 2015a). Moreover, elevated [CO$_2$] leads to an altered expression pattern of genes involved in the photosynthetic apparatus, the distribution of C, respiration, and N metabolism in durum wheat (Vicente et al., 2015b).

Increasing greenhouse gas emissions may cause further warming together with rainfall reduction in the next decades, which will increase the frequency and intensity of drought in the Mediterranean basin (Habash et al., 2009; IPCC, 2013; McKersie, 2015). For the Iberian Peninsula it is predicted that drought stress can occur at any growth stage of wheat (Russo et al., 2015), with the grain-filling phase being the most studied. However, the number of studies focusing on drought stress during early growth is limited. Although rainfall has been traditionally most abundant and evapotranspiration the lowest during winter, the occurrence of drought in winter months during the early stages of the crop cycle has been reported in recent times (Russo et al., 2015). This can further constrain wheat growth and thus final grain yield, mostly through a decrease in the ear density and number of kernels per unit crop area (Araus et al., 2008; Rebolledo et al., 2013). In addition, a constitutive (i.e., in absence of water stress) rapid development of wheat plants (early vigor) could be a positive trait and relevant for further avoiding drought stress-related consequences at both early and late growth stages. Early vigor could benefit plant growth and yield by increasing resource acquisition, shading the soil, preventing evaporation from it, and suppressing weeds (Maydup et al., 2012; Bort et al., 2014; Pang et al., 2014). As a consequence, differences in early growth (tillering and further stem elongation) will affect the number of fertile stems (and thus the ear density) and the size of the ears (and thus the potential number of grains per ear), which are the main contributors determining grain yield (Guo et al., 2016).

Plant responses to water stress define a complex and sophisticated regulatory network comprising physiological, biochemical, and molecular mechanisms. In wheat, some of these responses include inhibition of plant growth and photosynthetic capacity, together with a wide range of physiological responses, including changes in stomatal closure and decreases in transpiration, Rubisco efficiency, and chlorophyll content as well as an increase in oxidative stress among other responses (Budak et al., 2013; Nezhadahmadi et al., 2013). Such responses are modulated by stress severity. Cessation of watering showed a progressive reduction in leaf relative water content, water potential and photosynthesis in durum wheat (Habash et al., 2014). Liu et al. (2016) reported a progressive inhibition of photosynthetic activity as water stress is more severe in field-grown bread wheat, probably due to non-stomatal limitations, which led to lower grain yields even at moderate water stress. Furthermore, water stress in wheat leads to complex changes in the expression of some genes, including those involved in photosynthesis, respiration, N metabolism, lipid metabolism, transcription factors, signal transducers, and synthesis of protective proteins (Habash et al., 2009, 2014; Budak et al., 2013; Yousfi et al., 2016). These changes in gene expression occurred mainly in the early phases of the stress (Habash et al., 2014).

Plant responses to elevated [CO$_2$] or water stress are influenced by the duration and level of the environmental factor, the growth stage, and the genetic variability. Studies carried out with different durum wheat genotypes demonstrated that the responsiveness to elevated [CO$_2$] (Aranjuelo et al., 2013), water stress (De Leonardis et al., 2007; Aprile et al., 2013; Habash et al., 2014), and the combination of both (Erice et al., 2014) is

genotype specific. Moreover, the growth stage greatly influences the response of durum wheat to elevated [CO$_2$] (Aranjuelo et al., 2011; Vicente et al., 2015a) and drought (Liu et al., 2016). In addition, the interactive effects of environmental conditions and genotypic variability cannot be anticipated from the individual effects of these treatments (Ceccarelli et al., 1991). Some studies have shown positive effects of elevated [CO$_2$] on water stress tolerance of different bread wheat varieties (Harnos et al., 2002; Wall et al., 2006; Robredo et al., 2011; Bencze et al., 2014). A positive synergistic effect of elevated [CO$_2$] and water stress has been reported to decrease g$_s$, and thus leads to an improvement in water use efficiency at the stomatal and whole plant level (Bencze et al., 2014; Pazzagli et al., 2016). The decrease in photosynthesis under water stress is often mitigated by elevated [CO$_2$] (Bencze et al., 2014), resulting in increased levels of carbohydrates for the development of new tissues or filling grain (Wall et al., 2006). However, such positive effects of elevated [CO$_2$] in improving stress tolerance are not always achieved (Hudak et al., 1999; Pleijel et al., 2000). Bencze et al. (2014) reported that drought at elevated [CO$_2$] led to a stimulation of the antioxidant enzyme system in bread wheat, which suggests a high level of oxidative stress. Erice et al. (2014) showed that the stimulation of plant growth by elevated [CO$_2$] was only found in durum wheat genotypes with high harvest indices and optimal water supply. Therefore, additional efforts are still necessary to deepen our understanding of the interactive effect of [CO$_2$] and water regime in durum wheat.

The aim of this work was to determine the physiological and molecular mechanisms involved in the adaptive response of four semi-dwarf (i.e., post-Green Revolution) durum wheat cultivars to different [CO$_2$] and water regimes. Durum wheat genotypes were grown under controlled conditions at ambient and elevated [CO$_2$] and two different water regimes (fully irrigated and moderate/severe water stress). We assessed plant growth, physiological traits, stable C and N isotopic signatures, and transcript levels for stress-responsive genes that could be good indicators of durum wheat's adaptation to future climate conditions at vegetative growth stages. The genes selected corresponded to key enzymes in the metabolism of C (the Rubisco large and small subunits, RBCL and RBCS, respectively, and phosphoenolpyruvate carboxylase, PEPC) and N (the cytosolic and plastidial glutamine synthetases, GS1 and GS2, respectively), as well as proteins involved in stress responses (dehydrins 11, DHN11, and 16, DHN16, catalase, CAT, and superoxide dismutase, SOD). Rubisco is the key enzyme for photosynthetic CO$_2$ assimilation, and its activity is highly responsive to atmospheric [CO$_2$] (Vicente et al., 2011; Carmo-Silva et al., 2015). PEPC is a cytosolic enzyme that catalyzes the β-carboxylation of phosphoenolpyruvate to produce oxaloacetate, which is involved in anaplerotic functions. GS1 and GS2 play a central role in N metabolism: the former is thought to be involved in the primary assimilation of ammonium from nitrate reduction and photorespiration, while the latter is mainly involved in the transport of N through the plant and N recycling from catabolic processes. The function of the dehydrin family is not completely understood, but these proteins are involved in conferring stress tolerance (Kosová et al., 2014). Catalases and

superoxide dismutases are primary antioxidant enzymes involved in the elimination of reactive oxygen species (ROS) such as the cytotoxic H$_2$O$_2$ produced by photorespiration (Luna et al., 2005) and the superoxide generated during photosynthetic electron transport (Xu et al., 2010; Huseynova et al., 2014). Thus, our study combines the effects of genotypic variability and future environmental conditions, integrating plant performance with gene expression, and aims to identify traits associated with better performance during vegetative growth.

MATERIALS AND METHODS

Plant Material and Growth Conditions

The experiment was conducted with four semi-dwarf durum wheat [*Triticum turgidum* L. ssp. *durum* (Desf.)] genotypes: Mexa (year of commercial release: 1977), Regallo (1988), Burgos (1997) and Ramirez (2006). These cultivars represent high-yield genotypes released in the last forty years that are (or were) widely cultivated in the Mediterranean regions of Spain. The study of these genotypes could provide information about the adaptation of modern cultivars to climate change and whether there are differences between them associated with the year they were released. The experiment was conducted from May to July 2015 in two controlled environment chambers (Conviron E15; Controlled Environments, Winnipeg, MB, Canada) in the Experimental Facilities of the Faculty of Biology at the University of Barcelona. A total of 96 durum wheat plants (24 for each genotype) were sown in 2 L pots containing a mixture of standard substrate:perlite (1:1, v/v) and were grown with a long light period of 16 h, a photosynthetic photon flux density (PPFD) of 350 μmol m^{-2} s^{-1}, a day/night temperature of 23/17°C and a relative humidity of 60%. During the entire experiment, half of the pots were cultivated under atmospheric [CO$_2$] (400 μmol mol^{-1}) in one chamber, while the other half grew under elevated [CO$_2$] (790 μmol mol^{-1}) in the other chamber with injection of CO$_2$ from an external bottle (Carburos Metálicos S.A., Barcelona, Spain). The temperature, relative humidity and [CO$_2$] within each chamber were continuously monitored by Conviron series controllers (CMP3243 Controlled Environments Ltd., Winnipeg, MB, Canada). The technical staff of the Experimental Facilities of the Faculty of Biology tested the growth conditions of each chamber periodically with external sensors: an HMP75 humidity and temperature probe and a GMP222 CO$_2$ probe for use with an MI70 series hand-held indicator (Vaisala, Vantaa, Finland). Similarly, the PPFD was periodically verified with an LI-188B quantum/radiometer/photometer (LI-COR Inc., Lincoln, NB, USA).

The plants were uniformly irrigated every 2 days with 50% Hoagland's nutrient solution over a 25 day period. After that (Zadoks 21), the water stress was imposed; one half of the plants of each genotype and [CO$_2$] were maintained under well-watered conditions (100% pot capacity, PC) until the end of the experiment, while the other half were subjected to water stress conditions. The maximum soil volumetric water content of each pot was evaluated at the beginning of the experiment as the difference between pot weight after watering with the excess water drained and the pot dry weight. Thus, pots were

watered by direct measurements of the pot weight and the water supply was adjusted to the pot water conditions established for each water regime. In the water-stressed plants the watering was progressively restricted by 10% PC every 2 days. First, after 8 days the water-stressed plants received a 60% PC (moderate water stress) and this irrigation regime was strictly maintained for 10 days (see **Figure 1** for a schematic representation of the experimental design). At the end of this period (Zadoks 26), equal numbers (48) of well-watered and water-stressed plants were sampled. The youngest fully expanded leaf was collected, immediately frozen in liquid nitrogen and stored at −80°C for gene expression, C and N content and stable isotope analyses. After that, the whole plant was harvested and dried in an oven at 60°C for 72 h for biomass analysis. Second, in the remaining half of the plants (48), the progressive water limitation continued for 8 more days until water-stressed plants received a 30% PC (severe water stress). As in the moderate water stress, the irrigation conditions in well-watered and water-stressed plants were maintained for 10 days. Later, these 51-day-old plants (Zadoks 28–32) were collected following the procedure described above. The moderate and severe water stresses were defined in this experiment based on similar reductions in irrigation and stomatal conductance used in other studies (Galmes et al., 2007; Liu et al., 2016). The pots were rotated three times a week to avoid edge effects in the growth chambers over the course of the experiment. We used a rotatory randomized complete block design with three replicates (one plant per pot) per factor combination ([CO$_2$], water level and genotype) at each sampling.

Physiological Traits

Prior to harvest a hand-held portable spectroradiometer (GreenSeeker, NTech Industries, Ukiah, CA, USA) was used to estimate the normalized difference vegetation index (NDVI) of each plant (only at the second sampling date). Relative chlorophyll content was measured with a Minolta SPAD-502 chlorophyll meter (Spectrum Technologies, Plainfield, IL, USA). Stomatal conductance (g$_s$) was measured using a Decagon SC-1 Leaf Porometer (Decagon Device, Inc., Pullman, WA, USA). Both chlorophyll content and stomatal conductance of the adaxial surface were recorded in the central segment of the same youngest fully expanded leaf between 3 and 5 h after the start of the photoperiod. In addition, plants were collected to determine the leaf, shoot, root, and plant dry weights as indicated above, while the roots were washed in tap water until all substrate was removed. The number of tillers and the root to shoot dry weight ratio (root/shoot) were then determined.

C and N Content and Stable Isotope Signatures

A fraction of the youngest fully expanded leaf was finely powdered and then 1 mg of this leaf material was used for the measurements of total C and N content (as a percentage of

FIGURE 1 | Schematic representation of the experimental design.
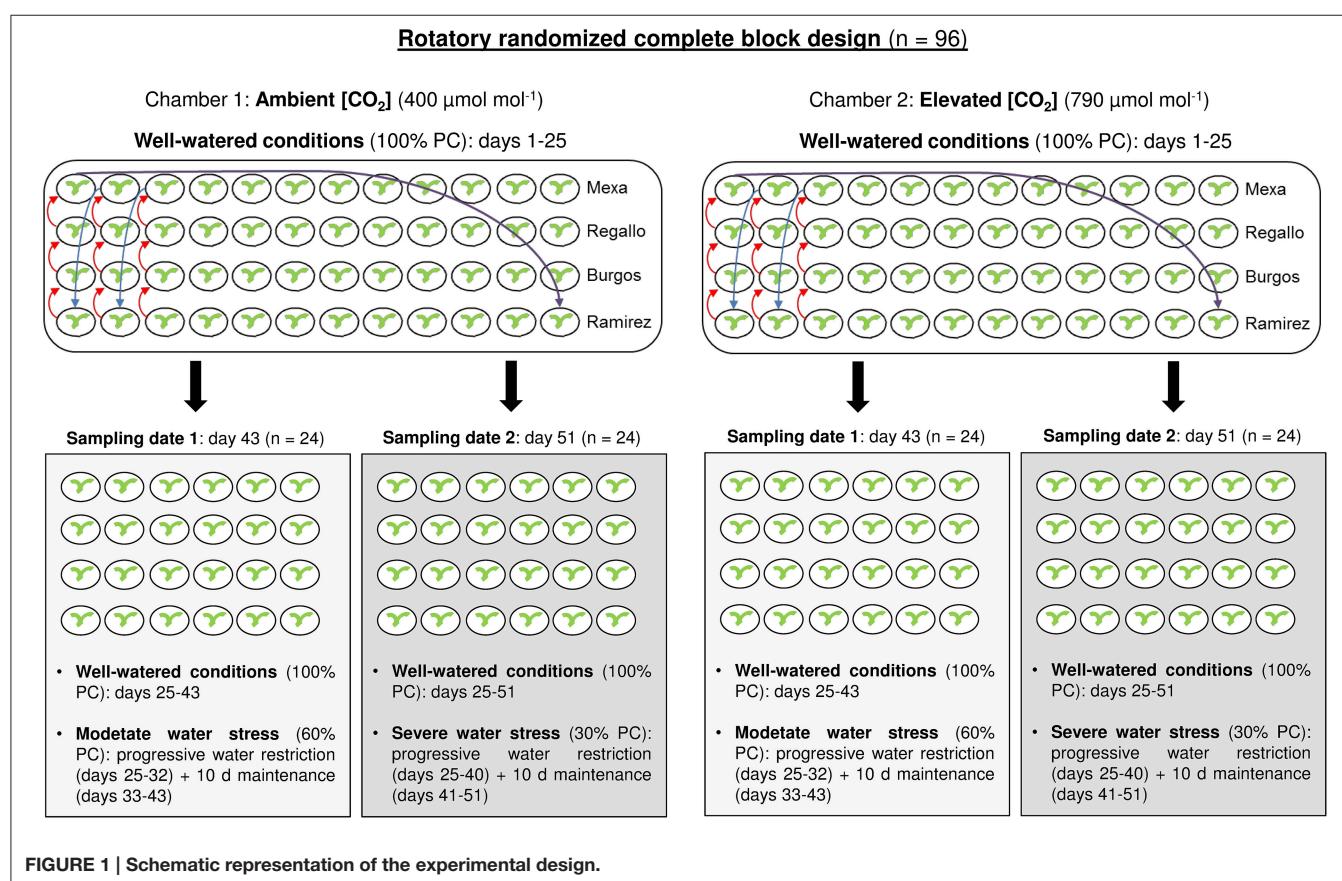
leaf dry weight) and the stable C ($^{13}C/^{12}C$) and N ($^{15}N/^{14}N$) isotope ratios. Measurements were carried out using an elemental analyzer (Flash 1112 EA; ThermoFinnigan, Bremen, Germany) coupled with an isotope ratio mass spectrometer (Delta C IRMS; ThermoFinnigan), operating in continuous flow mode, at the Scientific Facilities of the University of Barcelona. As has been described previously (Bort et al., 2014; Yousfi et al., 2016), the $^{13}C/^{12}C$ ratio was expressed in δ notation: $\delta^{13}C$ (‰) = [($^{13}C/^{12}C$)$_{sample}$/($^{13}C/^{12}C$)$_{standard}$ − 1] × 1000. The standard refers to international secondary standards of known $^{13}C/^{12}C$ ratios (IAEA CH7 polyethylene foil, IAEA CH6 sucrose, and USGS 40 L-glutamic acid) calibrated against Vienna Pee Dee Belemnite calcium carbonate. The same δ notation was used for the $^{15}N/^{14}N$ ratio ($\delta^{15}N$) using N_2 in air as standard.

Quantitative Reverse Transcriptase PCR Amplification

Frozen leaf samples were ground with liquid nitrogen and subsequently RNA was isolated from 100 mg of this material with Ribozol RNA Extraction Reagents (Amresco, Solon, OH, USA) according to the manufacturer's instructions. RNA quantity and quality was measured using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). RNA integrity was checked by 1.5% (w/v) agarose gel electrophoresis. Total RNA (1 μg) was treated with PerfeCTa DNase I RNase-free (Quanta Biosciences, Gaithersburg, MD, USA) to eliminate residual genomic DNA. cDNA was synthesized using a qScript cDNA Synthesis Kit (Quanta Biosciences) following the manufacturer's instructions. The qRT-PCR assays were performed in optical 384-well-plates with the LightCycler 480 System (Roche Applied Science, Penzberg, Germany) in the Centres Científics i Tecnològics de la Universitat de Barcelona (CCiTUB), in a reaction volume of 10 μL: 5 μL of PerfeCTa SYBR Green FastMix (Quanta Biosciences), 200 nM of each gene-specific primer and 1 μL of diluted cDNA (1:10). The thermal profile was as follows: initial denaturation for 30 s at 95°C, PCR cycling (45 cycles) for 5 s at 95°C, 15 s at 60°C, and 10 s at 72°C, and a final step of 95°C for 5 s and 60°C for 60 s to obtain the dissociation curve. Two technical replicates were analyzed per biological replicate. Specific primers for genes encoding the Rubisco large subunit (NC_021762), phosphoenolpyruvate carboxylase (Y15897), plastidial glutamine synthetase (DQ124212), dehydrin 11 (AJ890140), and superoxide dismutase (KP696754) were designed in Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) using the following criteria: Tm = 60 ± 1°C, primer length of 18–25 bases, GC content of 30–70% and product size of 60–150 bases. The specificity of PCR amplification was confirmed by the presence of unique amplicons of the expected length on 3.5% (w/v) agarose gels. The genes encoding the ADP-ribosylation factor and the RNase L inhibitor-like protein, previously identified as potential reference genes (Vicente et al., 2015b), were used to normalize qRT-PCR data after the evaluation of their expression stability in this study. All primers used for gene expression analysis and their symbols are listed in Supplementary Table S1. The values of the cycle threshold (C_t) were calculated using the LightCycler 1.5 software (Roche Applied Science). The quantification of the relative gene expression was analyzed using the comparative C_t method $2^{-\Delta\Delta Ct}$ (Schmittgen and Livak, 2008), and the data were presented as the $\log_2$ fold change.

Data Analysis

The effects of [CO_2] (ambient and elevated), water regime (well-watered and water stressed), genotype (Mexa, Regallo, Burgos, and Ramirez), and their interaction on plant growth, chlorophyll content, g_s, and C and N contents and isotope composition were determined through a three-factor (2 CO_2 × 2 water regimes × 4 genotypes) analysis of variance (ANOVA) for each sampling date (moderate and severe waters stress; see Supplementary Table S2) with GenStat 6.2 (VSN International Ltd, Hemel Hempstead, UK). Further, and given the implicit complexity of the design, each genotype was analyzed through a two-factor ANOVA (2 CO_2 × 2 water regimes) for both sampling dates. All factors were treated as fixed independent variables. When the F-ratio was significant ($P < 0.05$), the least significant difference (LSD) test was used to assess differences between treatment means. Clustered heat maps of relative gene expression were built in the R statistics environment (R Development Core Team, 2008) to study the effects of elevated [CO_2] and water stress on transcript levels. A correlation matrix was generated in R for evaluating the relationships between all parameters analyzed. Visualization of significant correlations was performed using Cytoscape software (Shannon et al., 2003).

RESULTS

Effect of [CO_2], Water Regime, and Genotype on Plant Growth

Total biomass of the plant and its different fractions (leaves, shoot and root), the root/shoot ratio, and the number of tillers were analyzed through two-factor ANOVA ([CO_2] × water regime) for each genotype (**Tables 1**, **2**). Moderate and severe water stress were established with reductions of 40 and 70% in the water supplied to the pots and average decreases of 34 and 57% in g_s, respectively, compared to well-watered plants (data not shown). Growth under elevated [CO_2] led to significant increases in biomass compared to ambient [CO_2] (**Tables 1**, **2**). At the first sampling date, elevated [CO_2] increased root biomass in Mexa, Regallo and Ramirez (and also in Burgos, $P = 0.074$), but only increased plant biomass in Regallo. The root/shoot ratio also increased in Regallo and Burgos under elevated [CO_2]. At the second sampling date, elevated [CO_2] increased plant biomass due to higher shoot and root biomass compared to ambient [CO_2], with larger increases in plant biomass in Mexa and Regallo under well-watered conditions in comparison to water stressed conditions. As a consequence of the increases in both shoot and root dry weights by elevated [CO_2], the root/shoot ratio was not altered, except in Regallo. Moreover, in this genotype an increase in the tillers per plant was also observed under elevated [CO_2] but only in well-watered conditions.

Moderate water stress did not lead to statistical differences in biomass, the root/shoot ratio, or the number of tillers between

TABLE 1 | Total leaf (LDW), shoot (SDW), root (RDW) and plant (PDW) dry weight, root/shoot ratio, and number of tillers per plant in four durum wheat genotypes grown under ambient or elevated [CO_2] and well-watered or moderate water stress conditions (100 vs. 60% pot capacity).

	Genotype	Ambient [CO_2]		Elevated [CO_2]		P_C	P_W	$P_{C \times W}$
		Well-watered	Water stressed	Well-watered	Water stressed			
LDW (g)	Mexa	1.84	1.29	1.45	1.23	0.527	0.304	0.656
	Regallo	1.26[a]	1.37[ab]	1.59[b]	1.18[a]	0.402	0.105	**0.013**
	Burgos	1.32	2.36	1.83	1.28	0.562	0.618	0.128
	Ramirez	1.98	1.41	1.70	1.15	0.417	0.109	0.964
SDW (g)	Mexa	3.12	1.85	2.17	2.27	0.732	0.463	0.387
	Regallo	1.75	1.97	2.14	1.90	0.441	0.962	0.258
	Burgos	4.18	3.78	2.47	1.91	0.147	0.679	0.946
	Ramirez	3.27	2.36	2.39	1.64	0.339	0.320	0.924
RDW (g)	Mexa	0.65	0.47	1.24	0.82	**0.023**	0.109	0.512
	Regallo	0.65	0.65	0.96	1.26	**<0.001**	0.080	0.074
	Burgos	0.85	0.49	0.90	1.00	0.074	0.374	0.125
	Ramirez	0.72	0.62	1.10	1.00	**0.022**	0.459	0.981
PDW (g)	Mexa	3.77	2.32	3.41	3.10	0.803	0.298	0.490
	Regallo	2.41	2.62	3.10	3.16	**0.002**	0.351	0.576
	Burgos	5.03	4.27	3.36	2.91	0.211	0.602	0.895
	Ramirez	3.99	2.98	3.50	2.64	0.577	0.232	0.921
Root/shoot	Mexa	0.30	0.26	0.56	0.36	0.055	0.184	0.353
	Regallo	0.38	0.37	0.45	0.69	**0.045**	0.214	0.180
	Burgos	0.28	0.19	0.37	0.52	**0.025**	0.711	0.146
	Ramirez	0.32	0.34	0.47	0.61	0.095	0.484	0.588
Tiller/plant	Mexa	9.3	7.0	8.0	6.7	0.543	0.200	0.713
	Regallo	7.0	8.3	8.3	8.3	0.567	0.567	0.567
	Burgos	8.7	11.0	9.0	9.0	0.620	0.491	0.491
	Ramirez	11.7	8.7	7.0	6.3	**0.045**	0.248	0.451

Significant effects for elevated [CO_2] (C), water stress (W) and their interaction (C × W) were determined by two-factor ANOVA (P). Values with the same letter are not significantly different for the interaction [CO_2] × water level. Significant P values are marked in bold (P < 0.05).

well-watered and water-stressed plants (**Table 1**). However, severe water stress led to significant changes in these parameters, while NDVI was also affected (**Table 2**). Plant biomass generally decreased under severe water stress compared to well-watered conditions and was associated with decreases in leaf, shoot, and root dry weights. Water restriction decreased the number of tillers per plant in Burgos under severe water stress, while this reduction was not significant in the other genotypes. Additionally, the NDVI values were lower in water-stressed plants compared to well-watered plants, irrespective of the [CO_2] and the genotype (**Table 2**). In general, at severe water stress the interaction [CO_2] × water regime × genotype showed that the root/shoot ratio strongly increased in Regallo, especially under ambient [CO_2] and well-watered conditions (Supplementary Table S2). The Burgos and Mexa cultivars had higher shoot dry weight than Ramirez and Regallo, while root dry weight was higher in Regallo (Supplementary Table S2). Furthermore, significant [CO_2] × genotype interaction showed that Burgos and Regallo under elevated [CO_2]

increased tiller production, whereas Ramirez and Regallo plants under ambient [CO_2] had lower tillering (Supplementary Table S2).

Effect of [CO_2], Water Regime, and Genotype on Chlorophyll Content, g_S, C and N Content and C and N Isotope Composition

The interactive effects of [CO_2] and water regime on chlorophyll content, g_s, and C and N contents and isotope composition were analyzed in the youngest fully expanded leaf through two-factor ANOVA for each genotype during vegetative growth under moderate (**Table 3**) and severe water stress (**Table 4**). At moderate water stress, elevated [CO_2] compared to ambient [CO_2] decreased N content in Mexa and Regallo, and δ^{13}C regardless of the genotype, while it increased chlorophyll content in Mexa, g_s, and C content in Regallo, and δ^{15}N in all genotypes except in Regallo. Water stress reduced g_s and increased

TABLE 2 | Total leaf (LDW), shoot (SDW), root (RDW) and plant (PDW) dry weight, root/shoot ratio, number of tillers per plant, and normalized difference vegetation index (NDVI) in four durum wheat genotypes grown under ambient or elevated [CO$_2$] and well-watered or severe water stress conditions (100 vs. 30% pot capacity).

	Genotype	Ambient [CO$_2$]		Elevated [CO$_2$]		P_C	P_W	$P_{C \times W}$
		Well-watered	Water stressed	Well-watered	Water stressed			
LDW (g)	Mexa	2.40[a]	2.69[a]	4.73[b]	1.97[a]	0.127	**0.031**	**0.012**
	Regallo	1.09[a]	1.75[b]	4.56[c]	2.23[b]	**<0.001**	**0.001**	**<0.001**
	Burgos	2.77	1.70	4.67	1.80	0.082	**0.004**	0.112
	Ramirez	2.27	1.86	3.28	1.98	0.178	0.057	0.282
SDW (g)	Mexa	4.26[a]	4.96[a]	8.35[b]	4.97[a]	**0.039**	0.147	**0.040**
	Regallo	1.49[a]	3.71[b]	6.24[c]	3.20[b]	**0.001**	0.361	**<0.001**
	Burgos	4.78	4.21	8.43	5.43	0.058	0.143	0.301
	Ramirez	3.59	3.53	6.98	3.73	**0.035**	**0.048**	0.054
RDW (g)	Mexa	1.31[a]	1.14[a]	2.65[b]	1.47[a]	**0.002**	**0.008**	**0.030**
	Regallo	2.42	1.92	3.54	2.45	**0.002**	**0.002**	0.139
	Burgos	1.54	1.33	3.01	1.75	**0.003**	**0.012**	0.053
	Ramirez	1.59	1.37	2.33	1.93	**0.013**	0.166	0.689
PDW (g)	Mexa	5.56[a]	6.10[a]	10.99[b]	6.44[a]	**0.018**	0.074	**0.031**
	Regallo	3.91[a]	5.63[b]	9.78[c]	5.65[b]	**<0.001**	**0.025**	**<0.001**
	Burgos	6.32	5.54	11.43	7.19	**0.011**	**0.039**	0.128
	Ramirez	5.18	4.89	9.32	5.66	**0.012**	**0.032**	0.058
Root/shoot	Mexa	0.33	0.23	0.32	0.30	0.400	0.123	0.267
	Regallo	1.65[b]	0.58[a]	0.57[a]	0.76[a]	**0.011**	**0.013**	**0.002**
	Burgos	0.40	0.48	0.36	0.34	0.563	0.850	0.718
	Ramirez	0.58	0.39	0.34	0.53	0.662	0.994	0.136
Tiller/plant	Mexa	9.3	9.3	11.0	7.0	0.852	0.279	0.279
	Regallo	6.3[a]	8.0[a]	15.7[b]	10.0[a]	**0.002**	0.138	**0.017**
	Burgos	12.3	7.0	13.0	8.7	0.377	**0.005**	0.699
	Ramirez	8.0	7.7	6.7	6.7	0.773	0.174	0.267
NDVI	Mexa	0.29	0.18	0.36	0.13	0.678	**<0.001**	0.104
	Regallo	0.29	0.18	0.34	0.15	0.610	**<0.001**	0.067
	Burgos	0.25	0.17	0.30	0.16	0.524	**0.006**	0.347
	Ramirez	0.29	0.18	0.25	0.16	0.287	**0.003**	0.651

Significant effects for elevated [CO$_2$] (C), water stress (W) and their interaction (C × W) were determined by two-factor ANOVA (P). Values with the same letter are not significantly different for the interaction [CO$_2$] × water level. Significant P values are marked in bold (P < 0.05).

δ^{15}N in Regallo, and increased δ^{13}C in Burgos and Ramirez. Three-factor ANOVA showed significant interactions for δ^{15}N (Supplementary Table S2). The [CO$_2$] × genotype interaction mainly showed that δ^{15}N was higher in Ramirez and Burgos at elevated [CO$_2$] and in Regallo at both [CO$_2$], whereas the lowest values were observed in Ramirez at ambient [CO$_2$]. The water regime × genotype interaction indicated that δ^{15}N was higher in Mexa and Regallo under water stress than in the other genotypes.

At severe water stress, elevated [CO$_2$] relative to ambient [CO$_2$] decreased the N content in Regallo and Burgos and δ^{13}C regardless of the genotype, while it increased δ^{15}N in Burgos (**Table 4**). Furthermore, g$_s$ in Burgos and N content in Regallo decreased under severe water stress compared to well-watered

conditions. In addition, under well-watered conditions g$_s$ was higher in Burgos than in other genotypes (Supplementary Table S2). Chlorophyll content was lower in Regallo, whereas δ^{13}C was higher in Burgos, as compared to other genotypes (Supplementary Table S2).

Effect of [CO$_2$] and Water Regime on Gene Expression for Each Durum Wheat Genotype

Treatment effects on transcript levels were evaluated for each genotype using nine genes that encode enzymes of primary C and N metabolism and stress-responsive proteins (Supplementary Table S1). Elevated [CO$_2$] and water stress led to changes in

TABLE 3 | Chlorophyll content, stomatal conductance (g$_S$), N and C content, and N and C isotope composition (δ^{15}N and δ^{13}C, respectively) in four durum wheat genotypes grown under ambient or elevated [CO$_2$] and well-watered or moderate water stress conditions (100 vs. 60% pot capacity).

	Genotype	Ambient [CO$_2$]		Elevated [CO$_2$]		P_C	P_W	$P_{C \times W}$
		Well-watered	Water stressed	Well-watered	Water stressed			
Chlorophyll (SPAD units)	Mexa	46.0	51.3	52.3	52.5	**0.049**	0.122	0.156
	Regallo	45.8	52.2	50.5	47.7	0.962	0.472	0.093
	Burgos	45.0	49.4	51.3	50.8	0.178	0.470	0.369
	Ramirez	47.2	41.3	49.3	48.5	0.099	0.216	0.337
g$_S$ (mmol m^{-2} s^{-1})	Mexa	260.1	58.5	246.9	201.2	0.466	0.183	0.384
	Regallo	248.2	100.9	315.0	262.4	**0.003**	**0.006**	0.114
	Burgos	255.1	105.4	180.3	144.8	0.680	0.056	0.206
	Ramirez	248.6	140.0	166.1	253.1	0.772	0.839	0.093
N (%)	Mexa	5.04	4.79	3.85	4.34	**0.016**	0.675	0.212
	Regallo	5.14	4.72	3.81	3.84	**0.007**	0.534	0.477
	Burgos	5.05	5.16	3.63	4.78	0.174	0.323	0.410
	Ramirez	4.92	4.58	4.52	3.63	0.069	0.092	0.413
δ^{15}N (‰)	Mexa	2.43	3.18	3.81	3.78	**0.024**	0.344	0.301
	Regallo	2.35	4.43	3.15	3.34	0.741	**0.029**	0.058
	Burgos	2.19	2.26	3.92	3.18	**<0.001**	0.152	0.087
	Ramirez	2.01	1.98	3.38	3.56	**0.002**	0.831	0.744
C (%)	Mexa	40.6	41.3	41.5	41.0	0.527	0.842	0.158
	Regallo	38.9	39.3	41.0	40.8	**0.015**	0.821	0.577
	Burgos	40.1	40.1	40.3	45.2	0.229	0.268	0.268
	Ramirez	42.1	40.6	40.3	36.8	0.185	0.227	0.620
δ^{13}C (‰)	Mexa	−32.8	−29.7	−52.7	−55.5	**<0.001**	0.981	0.483
	Regallo	−33.4	−33.9	−53.9	−60.7	**<0.001**	0.205	0.265
	Burgos	−32.7	−29.0	−57.9	−50.4	**<0.001**	**0.016**	0.337
	Ramirez	−32.5	−29.8	−60.5	−52.8	**<0.001**	**0.047**	0.286

Significant effects for elevated [CO$_2$] (C), water stress (W) and their interaction (C × W) were determined by two-factor ANOVA (P). Significant P values are marked in bold (P < 0.05).

gene expression depending on genotype and the level of water restriction (**Table 5**; Supplementary Figure S1). At moderate water stress, elevated [CO$_2$] decreased transcript levels of RBCL, RBCS, and GS2 relative to control conditions (ambient [CO$_2$] and well-watered conditions), particularly in the Mexa and Regallo genotypes. Under water stress the transcript levels for these enzymes markedly increased in Ramirez. Elevated [CO$_2$] caused a generalized increase in the transcript levels of PEPC and GS1, particularly when it was combined with moderate water stress. Transcript abundances of the dehydrins, DHN11 and DHN16, were generally higher under elevated [CO$_2$] and well-watered conditions, but lower under ambient [CO$_2$] and water stress relative to control conditions. However, *DHN11* and *DHN16* showed opposite expression patterns under elevated [CO$_2$] and water stress. [CO$_2$] enrichment and moderate water stress decreased transcript levels of CAT and SOD in Mexa, Regallo, and Burgos compared with control conditions, whereas in Ramirez they did not change significantly.

Gene expression analysis indicated greater genotype-specific differences under severe water stress than under moderate

water stress (**Table 5**; Supplementary Figure S1). In Mexa under elevated [CO$_2$] and well-watered conditions there were higher transcript levels of RBCL, RBCS, PEPC, GS1, GS2, and CAT and lower levels of DHN16, relative to control conditions. Severe water stress did not substantially alter gene expression. In Regallo most of the transcripts studied were lower in all treatment combinations than in control conditions. However, DHN16 and SOD transcripts increased under ambient [CO$_2$] and water stress, and these together with CAT and GS1 also increased under elevated [CO$_2$] and water stress. In the case of Burgos, elevated [CO$_2$], water stress and their combination strongly reduced transcript levels in comparison to control conditions, especially for GS1 and DHN16, while SOD transcripts increased under water stress and elevated [CO$_2$] × water stress as observed in Regallo. In Ramirez elevated [CO$_2$] led to a reduction in the transcript levels of RBCL, RBCS, and SOD and an increase in PEPC and GS1 compared to control conditions. Water stress increased PEPC and DHN16 transcript levels relative to control conditions, while under the combination of elevated [CO$_2$] and water stress

TABLE 4 | Chlorophyll content, stomatal conductance (g$_s$), N and C content, and N and C isotope composition (δ^{15}N and δ^{13}C, respectively) in four durum wheat genotypes grown under ambient or elevated [CO$_2$] and well-watered or severe water stress conditions (100 vs. 30% pot capacity).

	Genotype	Ambient [CO$_2$]		Elevated [CO$_2$]		P_C	P_W	$P_{C \times W}$
		Well-watered	Water stressed	Well-watered	Water stressed			
Chlorophyll (SPAD units)	Mexa	55.8	55.0	52.7	52.7	0.100	0.790	0.807
	Regallo	50.6	43.1	46.6	45.8	0.877	0.351	0.439
	Burgos	56.6	53.2	53.3	58.4	0.661	0.683	0.076
	Ramirez	52.7	51.6	52.8	49.0	0.532	0.230	0.500
g$_s$(mmol m^{-2} s^{-1})	Mexa	81.9	55.0	223.5	91.5	0.112	0.150	0.323
	Regallo	101.8	82.3	64.4	32.2	0.265	0.499	0.866
	Burgos	245.0	96.7	247.0	30.7	0.390	**<0.001**	0.362
	Ramirez	44.0	52.5	151.3	57.2	0.160	0.270	0.194
N (%)	Mexa	4.39	4.67	4.10	4.54	0.491	0.257	0.785
	Regallo	4.68	4.33	4.37	3.78	**0.017**	**0.011**	0.424
	Burgos	4.89	4.66	4.15	3.97	**0.013**	0.398	0.918
	Ramirez	3.81	4.22	4.16	4.43	0.521	0.437	0.872
δ^{15}N (‰)	Mexa	2.71	2.74	3.52	3.59	0.060	0.895	0.950
	Regallo	3.78	3.27	3.55	3.58	0.947	0.685	0.646
	Burgos	2.45	2.67	3.69	3.47	**0.018**	0.997	0.547
	Ramirez	2.69	2.99	3.23	2.73	0.508	0.641	0.084
C (%)	Mexa	41.8	41.8	41.3	41.9	0.490	0.368	0.439
	Regallo	39.9	38.0	40.4	41.1	0.064	0.476	0.165
	Burgos	41.9	41.4	41.8	41.3	0.877	0.590	0.968
	Ramirez	41.4	40.7	40.9	42.1	0.416	0.658	0.115
δ^{13}C (‰)	Mexa	−32.3	−31.4	−56.9	−56.4	**<0.001**	0.546	0.851
	Regallo	−33.5	−32.2	−54.4	−54.0	**0.001**	0.850	0.927
	Burgos	−32.8	−32.4	−46.1	−46.3	**0.004**	0.979	0.940
	Ramirez	−31.9	−32.6	−59.1	−56.4	**<0.001**	0.752	0.594

Significant effects for elevated [CO$_2$] (C), water stress (W) and their interaction (C × W) were determined by two-factor ANOVA (P). Significant P values are marked in bold (P < 0.05).

greater transcript abundances were observed for most of the genes.

Correlation Network of Physiological Traits and Gene Expression

A Pearson correlation matrix was generated using the mean values for each treatment combination, genotype and sampling date ($n = 32$) of the physiological traits and transcript levels (Supplementary Table S3), excluding NDVI, which was only measured at severe water stress, and δ^{13}C, which was influenced by C composition of the CO$_2$ bottles used in the elevated [CO$_2$] chamber (Aljazairi et al., 2015). Of the 190 correlations between parameters, there were 28 positive and 19 negative significant correlations ($P < 0.05$) that are represented in an association network (**Figure 2**). Most of the significant correlations were observed between physiological traits and transcript levels independently. Positive correlations were found among leaf, shoot, and plant dry weights, between the leaf and shoot dry weights with the number of tillers, and between root and plant dry weights. The root/shoot ratio was positively correlated with root dry weight and negatively correlated with leaf and shoot dry weights, the number of tillers and N content. Furthermore, δ^{15}N was also negatively correlated with N content and the number of tillers. Chlorophyll content was correlated positively with leaf, shoot, root, and plant dry weights, and negatively with g$_s$. On the other hand, positive correlations were found between N content with leaf and shoot dry weights and the number of tillers, and negative correlations between N content with root dry weight, and between g$_s$ with root and plant dry weights. In the case of transcript levels, RBCL was correlated with RBCS, GS1, GS2, and PEPC, whereas RBCS correlated with GS2 and DHN11, GS2 with DHN11, and PEPC with GS1. Furthermore, some relationships were found between physiological traits and gene expression (**Figure 2**). Positive correlations appeared between DHN16 with plant biomass (leaf, shoot, root, and plant dry weights), CAT with g$_s$ and SOD with C content. Moreover, negative correlations were found between chlorophyll content with RBCL, RBCS, GS2, and CAT, also between root dry weight with RBCL, RBCS, GS2, and DHN11, and finally plant dry weight with RBCL.

TABLE 5 | Transcript changes in four durum wheat genotypes grown under ambient or elevated [CO$_2$] and well-watered or water stressed conditions: (A) moderate and (B) severe water stress.

Genotype	[CO$_2$]	Water supply	RBCL	RBCS	PEPC	GS1	GS2	DHN11	DHN16	CAT	SOD
(A) MODERATE WATER STRESS (100 vs. 60% POT CAPACITY)											
Mexa	Ambient [CO$_2$]	Water stressed	0.15	0.26	−1.01	−0.88	−0.06	−1.02	−2.83	−2.02	−0.33
Mexa	Elevated [CO$_2$]	Well-watered	−0.96	−4.81	0.76	1.93	−3.43	−0.98	1.79	−1.01	−1.06
Mexa	Elevated [CO$_2$]	Water stressed	−1.69	−1.66	−2.25	1.65	−1.81	−1.28	−1.02	−2.42	−1.05
Regallo	Ambient [CO$_2$]	Water stressed	−0.25	−0.62	0.2	−0.53	−0.27	−0.74	−0.66	−2.42	−0.85
Regallo	Elevated [CO$_2$]	Well-watered	−2.59	−1.71	−1.14	−0.35	−1.26	1.26	1.8	−2.31	−0.42
Regallo	Elevated [CO$_2$]	Water stressed	−0.59	−0.55	1.52	0.68	−1.29	−0.95	4.81	−1.62	−1.91
Burgos	Ambient [CO$_2$]	Water stressed	0.06	0.07	0.19	−0.21	0.93	−0.65	3.27	−2.65	−1.99
Burgos	Elevated [CO$_2$]	Well-watered	−2.45	−1.35	0.1	−0.92	−0.87	0.89	1.8	−3.18	−0.35
Burgos	Elevated [CO$_2$]	Water stressed	0.47	0.08	2.14	0.67	−0.37	−0.83	2.48	−2.75	−1.76
Ramirez	Ambient [CO$_2$]	Water stressed	1.7	1.76	0.63	0.34	1.24	−0.03	−2.1	−0.51	0.49
Ramirez	Elevated [CO$_2$]	Well-watered	−0.42	−0.87	0.67	1.03	−0.7	0.38	0.62	−0.47	−0.68
Ramirez	Elevated [CO$_2$]	Water stressed	0.56	−1.73	1.99	2.43	−1.21	−1.12	2.49	0.18	0.39
(B) SEVERE WATER STRESS (100 vs. 30% POT CAPACITY)											
Mexa	Ambient [CO$_2$]	Water stressed	0.57	0.52	0.31	0.19	0.08	−0.37	0.40	0.60	0.25
Mexa	Elevated [CO$_2$]	Well-watered	1.84	1.49	0.84	1.09	1.27	−0.23	−1.22	1.56	−0.03
Mexa	Elevated [CO$_2$]	Water stressed	−0.34	0.16	−0.42	−0.29	0.13	−0.58	−0.34	1.30	0.04
Regallo	Ambient [CO$_2$]	Water stressed	−1.35	−0.59	0.03	−0.19	−0.84	−1.14	1.02	−0.26	0.95
Regallo	Elevated [CO$_2$]	Well-watered	−1.38	−0.90	−2.38	−0.61	−0.77	−1.35	−1.69	−0.40	−0.17
Regallo	Elevated [CO$_2$]	Water stressed	−0.30	−0.05	−0.59	2.15	−0.44	0.17	3.31	1.80	0.57
Burgos	Ambient [CO$_2$]	Water stressed	−0.93	−0.83	−1.98	−3.21	−0.81	−1.35	−6.30	−2.70	0.71
Burgos	Elevated [CO$_2$]	Well-watered	−2.13	−1.70	−2.18	−2.75	−1.38	0.19	−7.46	−1.85	−0.02
Burgos	Elevated [CO$_2$]	Water stressed	−1.81	−1.73	−2.08	−2.92	−1.05	−0.46	−5.56	−1.91	0.80
Ramirez	Ambient [CO$_2$]	Water stressed	−0.44	−0.45	2.34	0.29	0.15	−0.57	1.68	−0.62	0.32
Ramirez	Elevated [CO$_2$]	Well-watered	−2.71	−2.34	0.55	1.11	−1.00	−0.04	0.04	−0.18	−1.18
Ramirez	Elevated [CO$_2$]	Water stressed	0.51	0.64	0.98	0.25	0.98	1.01	−0.47	0.81	−0.40

White indicates no change, blue up-regulation, and red down-regulation in each treatment relative to the treatment under ambient [CO$_2$] and optimal water supply for each genotype, as shown in the color bar for a log$_2$ scale. RBCL, Rubisco large subunit; RBCS, Rubisco small subunit; PEPC, phosphoenolpyruvate carboxylase; GS1, cytosolic glutamine synthetase; GS2, plastidial glutamine synthetase; DHN11, dehydrin 11; DHN16, dehydrin 16; CAT, catalase; SOD, superoxide dismutase.

−3   0   3

DISCUSSION

Although, substantial efforts have been made in recent years to identify traits associated with wheat performance during early growth (Maydup et al., 2012; Rebolledo et al., 2013; Bort et al., 2014; Pang et al., 2014; Wilson et al., 2015), little attention has been paid to the effect of interactions between elevated [CO$_2$] and water stress in durum wheat. The effects of water restriction on crop growth have been mostly studied with the view of improving drought impacts at late growth stages in Mediterranean environments. However, projections of future climate change in the Iberian Peninsula predict major rainfall limitations and higher evapotranspiration during winter months (Russo et al., 2015) and therefore early-season drought is a matter of concern. In this context, we describe the effects of elevated [CO$_2$] and water stress during the first part of the growth cycle in four durum wheat genotypes on physiological traits and expression of nine genes that respond to changes in [CO$_2$] and water levels (Ali-Benali et al., 2005; Budak et al., 2013;

Vicente et al., 2015b; Yousfi et al., 2016). The coordination of these parameters under the different combinations of factors is discussed.

Changes in Physiological Traits of Durum Wheat Genotypes under Different Water Regimes and [CO$_2$] Levels

A moderate water stress in 43-day-old plants did not significantly alter plant growth (**Table 1**). Long-term exposure to elevated [CO$_2$] led to higher root biomass relative to ambient [CO$_2$] independently of genotypic variability, in concordance with reports from other crop species (Madhu and Hatfield, 2013). This increment was associated with higher plant growth in Regallo and higher root/shoot ratios in Regallo, Burgos, and Mexa (**Table 1**). In fact, under elevated [CO$_2$] root growth is often more stimulated than the aerial part of the plant, although it depends on genotype × environment variation (Stitt and Krapp, 1999; Madhu and Hatfield, 2013). A severe water stress
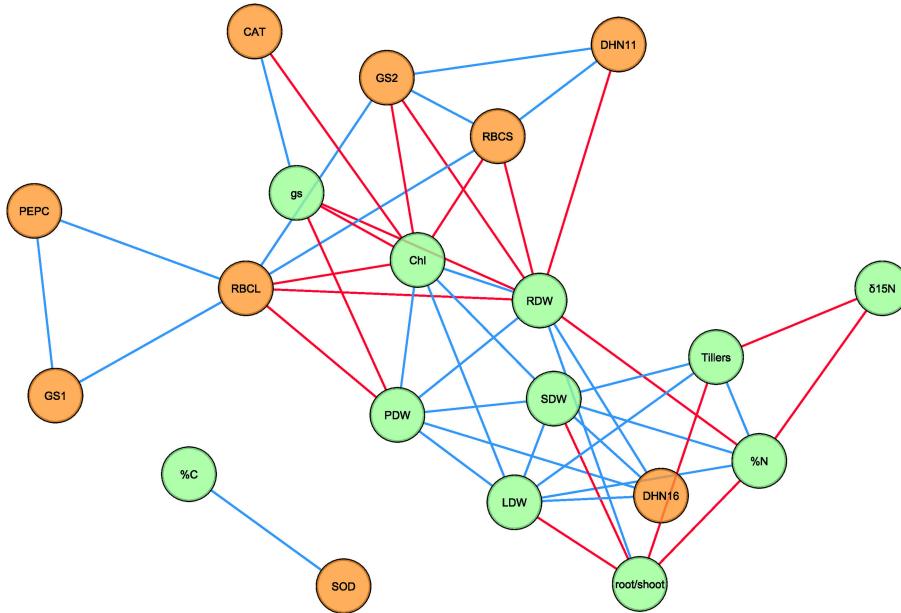
FIGURE 2 | Network analysis of physiological traits and transcript levels under different [CO₂] levels, water regimes and sampling dates in four durum wheat genotypes. The network consists of 20 nodes and 47 edges. Green and orange nodes represent physiological traits and transcript levels, respectively. Blue and red edges represent positive and negative correlations, respectively, based on Pearson's correlation coefficients. For transcript description see the legend of **Table 5**. Chl, chlorophyll; g$_s$, stomatal conductance; LDW, leaf dry weight; PDW, plant dry weight; SDW, shoot dry weight; RDW, root dry weight; %C, carbon content; %N, nitrogen content.

in 51-day-old plants showed greater effects on plant growth than moderate water stress (**Table 2**). [CO₂] enrichment generally led to an increase in plant biomass by increasing root and shoot biomass and tillering, particularly under optimal water supply. This could be due to the effects of [CO₂] fertilization on the net photosynthetic rate (Long et al., 2006; Vicente et al., 2015b), especially in genotypes with large harvest indices such as post-Green Revolution cultivars (Aranjuelo et al., 2013). It could also be caused by carbohydrate accumulation, which may lead to increases in the number of tillers (Stitt and Krapp, 1999). On the other hand, severe water stress constrained plant growth (dry matter and NDVI), in agreement with earlier studies in durum wheat (Erice et al., 2014; Nakhforoosh et al., 2015; Yousfi et al., 2016), with Ramirez and Burgos being the genotypes most affected. According to Marti et al. (2007), we suggest that progressive water restriction during the vegetative stage constrained the photosynthetic area, which may cause negative effects on final biomass and yield.

Chlorophyll content, g$_s$, and N and C contents and isotope compositions at moderate and severe water stress did not reveal statistical significance for the interactions [CO₂] × water regime and [CO₂] × water regime × genotype (**Tables 3, 4**; Supplementary Table S2). Stomatal conductance (g$_s$) generally decreases under elevated [CO₂] and drought stress due to an increase in internal [CO₂] and as a water saving strategy, respectively (Long et al., 2006; Nakhforoosh et al., 2015; Vicente et al., 2015b; Pazzagli et al., 2016). The average g$_s$ values decreased under water restriction at moderate and severe water stress, but it

was only significantly decreased in some genotypes (**Tables 3, 4**). On the other hand, elevated [CO₂] did not alter g$_s$ at this growth stage, except for an increase in g$_s$ under moderate water stress in Regallo, which could favor CO₂ assimilation and consequently biomass accumulation under this water regime (**Tables 1, 3**). Earlier studies have shown a decrease in g$_s$ under water stress (Peremarti et al., 2014; Pazzagli et al., 2016), while negligible changes have been reported under elevated [CO₂] in tomato and durum wheat, and increases have even been recorded for Regallo (Vicente et al., 2015a; Pazzagli et al., 2016). Therefore, the growth stage and the severity of the water stress influenced stomatal closure, while elevated [CO₂] had minor effects on g$_s$ during vegetative growth.

Elevated [CO₂] generally decreased N content in the present study (**Tables 3, 4**), which has been observed in C₃ plants through shifts in N uptake and/or assimilation (which agrees with the changes in transcript levels of N-metabolism enzymes; see below) together with other uncertain mechanisms, e.g., the biomass dilution effect, increased N loss, and sink limitation (Stitt and Krapp, 1999; Aranjuelo et al., 2011; Vicente et al., 2015a,b). N content was also diminished by severe water stress in Regallo, in agreement with previous studies in durum wheat (Yousfi et al., 2012, 2016). Chlorophyll content only increased under elevated [CO₂] in Mexa at the first sampling date, but the effect disappeared at the second sampling (**Table 3**). [CO₂] enrichment and water stress did not modify C content in leaves, suggesting that the decrease in N content was not simply due to N dilution caused by rapid growth (Taub and Wang, 2008). Overall, our

data showed that the decrease in N content in plants grown under elevated [CO_2] and water stress during vegetative growth is genotypically dependent.

The δ^{13}C and δ^{15}N have been used as potential physiological tracers in plants under elevated [CO_2] and water limitation (Aranjuelo et al., 2011; Yousfi et al., 2012, 2016; Araus et al., 2013; Bort et al., 2014). Elevated [CO_2] and water stress caused an increase in δ^{15}N, although these effects depended on the genotype and were attenuated or disappeared in severe water stress relative to the moderate stress treatment (**Tables 3, 4**; Supplementary Table S2). Variations in δ^{15}N in response to the growth conditions, together with N content, could indicate shifts in N metabolism (Bort et al., 2014), although δ^{15}N is determined by many processes that are not completely understood (Ariz et al., 2015). Nevertheless, the higher δ^{15}N could suggest lower N availability, because N absorption and assimilation cannot fractionate between the ^{14}N and ^{15}N isotopologues under such environmental factors (Lopes and Araus, 2006; Tcherkez, 2011). Additionally, this could reflect a decrease in N translocation from the root to the shoot (Lopes and Araus, 2006). Moreover, δ^{13}C increased in some genotypes under moderate water stress, regardless of the [CO_2] considered, but this increment, also observed under severe water stress, did not reach statistical significance (**Tables 3, 4**). Elazab et al. (2012) and Bort et al. (2014) also showed a δ^{13}C increase in flag leaves of different durum wheat genotypes under water stress at later growth stages, which could be associated with higher water-use efficiency (Araus et al., 2008, 2013; Tardieu, 2013; Bort et al., 2014). A stronger water stress does not always lead to larger changes in δ^{13}C, particularly when analyzed in dry matter, as noted in previous studies in rice (Kano-Nakata et al., 2014) and *Pinus tabuliformis* (Ma et al., 2014). In addition, δ^{13}C was strongly reduced at high [CO_2] because of the very negative δ^{13}C of the CO_2 used to increase the [CO_2] within the growth chamber (Aljazairi et al., 2015).

Expression of Stress-Responsive Genes in Durum Wheat Genotypes under Different Water Regimes and [CO₂] Levels

Strong differences in gene expression were observed between treatments and among the different genotypes studied (**Table 5**; Supplementary Figure S1). In our study, RBCL and RBCS showed a common expression pattern (**Table 5**), confirming the coordinated expression of both subunits necessary for the assembly of the Rubisco holoenzyme (Suzuki and Makino, 2012). At the first sampling date, gene expression of *RBCL* and *RBCS* was down-regulated in response to elevated [CO_2] no matter which water regime was considered, in agreement with other wheat studies (Aranjuelo et al., 2013; Habash et al., 2014; Vicente et al., 2015b). This down-regulation was associated with lower N content and higher δ^{15}N in a genotype-dependent manner. The former could be explained by non-selective decreases in N or reallocation of N within the plant under elevated [CO_2] (Aranjuelo et al., 2011; Vicente et al., 2015a). The latter was probably associated with changes in N uptake, assimilation or redistribution within the plant (Araus et al., 2013). At the second

sampling date, elevated [CO_2] decreased the N content in Regallo and Burgos, which was related to down-regulation of transcript levels of Rubisco subunits and N-assimilation enzymes (GS1 and GS2), and higher root and plant biomass. These shifts could indicate that plant biomass might increase under elevated [CO_2] in a genotype-dependent manner even when transcript levels of Rubisco subunits decrease during vegetative growth. This could be due to the remobilization of an N over-investment in Rubisco to reuse it in developing new tissues (Richards, 2000; Vicente et al., 2011; Carmo-Silva et al., 2015). However, the decrease in Rubisco transcript levels under water stress did not indicate the greater photosynthetic efficiency that was hypothesized under elevated [CO_2]. Instead it was associated with lower plant biomass, which might suggest an inhibition of CO_2 assimilation and plant growth in concordance with previous studies (Hayano-Kanashiro et al., 2009; Peremarti et al., 2014).

PEPC is a multifaceted key enzyme that in C₃ plants is linked to the provision of Krebs cycle intermediates, and its overexpression in transgenic wheat improved drought tolerance and grain yield (Qin et al., 2015). *PEPC* expression has not been widely studied during early growth in durum wheat plants. In the current work it was induced under the combination of elevated [CO_2] and moderate water stress in most genotypes, whereas at severe water stress genotypic variation determined its expression pattern (**Table 5**). The induction could be related to its major role in providing C skeletons for amino acid and lipid biosynthesis (González et al., 2003). This may be due to an increase in the enzyme's substrates, such as carbohydrates, typically found under elevated [CO_2] and water stress (Khoshro et al., 2013; Vicente et al., 2015b). These results indicate that further work is necessary to broaden our understanding of the biological role of PEPC and its implication in plant growth, especially in genotypes (i.e., Ramirez) with an up-regulation of gene expression under stress conditions.

At moderate and severe water stress, *GS1* and *PEPC* expression was significantly coordinated, as were the expressions of the *GS2* and Rubisco genes (**Table 5**; Supplementary Table S3). Under severe water stress, *GS1* and *GS2* expression was more influenced by genotypic variability than environmental conditions. Yousfi et al. (2016) also reported genotypic differences in the expression of these genes under drought stress, with a general down-regulation under stress conditions. Lower N contents and transcript abundances for RBCL and RBCS under water stress and especially under elevated [CO_2] were associated with higher repression of the *GS2* gene, indicating a coregulation of primary C and N metabolism (Stitt and Krapp, 1999; Vicente et al., 2015b, 2016). In some treatments, mainly at the first sampling date, opposing gene expression patterns were observed between *GS1* and *GS2*. This fact, together with the coordination of *GS1* with *PEPC*, might indicate a predominant remobilization of C and N compounds and an inhibition of primary N assimilation under water stress and elevated [CO_2]. Thus, the results support a significant coordination between C and N metabolism at the transcript level under conditions of elevated [CO_2] and water stress. In addition, the pattern of gene expression for *GS1* and *GS2* supports the use of these genes as indicators of N metabolism under

water stress conditions, as reported previously (Nagy et al., 2013).

DHN11 and DHN16 encode for two dehydrins that belong to group 2 of late embryogenesis abundant (LEA) proteins (Ali-Benali et al., 2005). The up-regulation of dehydrin genes under water restriction is often associated with stress tolerance, although their specific role as osmotically active compounds is still unknown (Kosová et al., 2014). Moderate and severe water stress reduced *DHN11* gene expression regardless of the [CO$_2$] level compared with control conditions. In the case of the *DHN16* gene, moderate water stress mostly up-regulated its expression, whereas under severe water stress the opposite occurred (**Table 5**). Elevated [CO$_2$] at the first sampling date mostly enhanced *DHN11* and *DHN16* gene expression, while at the second sampling date its combination with severe water stress led to a wide range of changes in transcript levels in a genotype-dependent manner. Our results showed that the pattern of gene expression could differ between dehydrins, in concordance with previous studies (Ali-Benali et al., 2005; Melloul et al., 2013; Kosová et al., 2014). Additionally, the severity of the water stress, [CO$_2$] enrichment and the genotype influenced dehydrin transcript levels.

CAT and SOD enzymes form part of the system responsible for lowering ROS and avoiding oxidative stress. In general, gene expression of *CAT* and *SOD* was repressed under moderate water stress regardless of [CO$_2$] (**Table 5**). Such repression was only maintained for *CAT* at severe water stress in Burgos, while their expression was up-regulated in the other genotypes under elevated [CO$_2$] × severe water stress. This could suggest a higher demand for ROS control, which would indicate a limitation to the transfer of electrons through photosystems to drive C assimilation (Martins et al., 2016). Enzyme activity and *CAT* gene expression have been reported to decrease under elevated [CO$_2$] in wheat, possibly due to the inhibition of photorespiration, while they increased only in response to severe drought (Luna et al., 2005; Xu et al., 2010; Vicente et al., 2015b). The available studies reporting changes in *SOD* gene expression and protein content under such conditions are contradictory, reporting different pattern of changes (Kim et al., 2006; Li et al., 2008; Caruso et al., 2009; Xu et al., 2010). Our results highlighted that water regime and genotype were key factors influencing the expression of genes involved in the antioxidant system, indicating a greater need for protection against oxidative damage under severe water stress.

Coordination between Physiological Traits and Transcript Levels in Durum Wheat Grown under Different Environmental Conditions during Vegetative Growth

The different changes in plant growth parameters indicate that the responsiveness to elevated [CO$_2$] and water stress during early growth depends on (i) the duration of the treatment, because [CO$_2$] enrichment results in greater increases in plant biomass in older plants; (ii) the severity of the water stress, which is more pronounced under severe water stress; (iii) and the genotypic variability. In general, elevated [CO$_2$] stimulated plant growth and reduced N content, which at the transcript

level was related to a down-regulation of Rubisco and N assimilation genes and up-regulation of genes that take part in C-N remobilization. Moderate water stress did not lead to gross changes in physiological traits, but severe water stress restricted plant growth and N content, while changes in g$_s$ and δ^{13}C suggested a water-saving strategy relative to well-watered conditions. The transcript profile suggested an inhibition of primary C fixation and N assimilation, differences between dehydrins and a genotypic variation in gene expression under severe water stress, with an induction of genes involved in antioxidant machinery. The stimulation of plant biomass under elevated [CO$_2$] did not compensate for plant growth limitation under water restriction. Lastly, we observed different genotypic responses to environmental factors, as also reported in barley (Ceccarelli et al., 1991). Regallo showed the lowest plant biomass and chlorophyll and N contents, which was related to a repression of genes for N assimilation and induction for dehydrins, SOD and CAT, while the opposite results were recorded for Burgos (data not shown). Therefore, increased plant growth was linked to up-regulation of N assimilation and down-regulation of stress-responsive genes, suggesting lower oxidative damage.

Considering different environmental conditions predicted for the future climate scenario and genotypic variations, network analysis was used to identify physiological traits, and transcript levels that are correlated during vegetative growth in durum wheat (**Figure 2**). Early growth is a positive trait for improving plant tolerance in water-limited environments that has the potential for larger final plant biomass and yield (Wilson et al., 2015). Plant growth parameters were positively correlated with each other in most cases, suggesting that early plant growth is driven by all plant fractions and tiller production, as reported in other studies (Rebolledo et al., 2013; Wilson et al., 2015). Regardless of genotype, the positive correlation between root and plant biomass was mainly due to the stimulation of root biomass under elevated [CO$_2$], in agreement with previous reports (Madhu and Hatfield, 2013; and citations therein). In contrast, water restriction (mainly severe water stress) limited both root and shoot biomass, which are often diminished under severe drought conditions (Nezhadahmadi et al., 2013). Positive effects of elevated [CO$_2$] on root biomass could mitigate drought effects on plant growth by allowing better exploitation of water and nutrients from deep soil layers (Madhu and Hatfield, 2013).

N content was correlated negatively with the root/shoot ratio and positively with the tillers per plant and shoot biomass, and this was probably due to the typically higher N content observed in shoots relative to roots (Vicente et al., 2015a). Hence, greater vegetative growth in durum wheat requires high amounts of N, which in turn will be conditioned by N availability. δ^{15}N has been proposed as an indicator of responses to stress, such as water stress, N starvation and salinity (Yousfi et al., 2012, 2016; Bort et al., 2014), although it has had little attention for studies of elevated [CO$_2$] (Ariz et al., 2015). Here we observed a negative correlation of δ^{15}N with N content and tillers per plant, with elevated [CO$_2$] being the main factor that increased δ^{15}N in our experiment. Nevertheless, the fractionating processes of N metabolism affecting δ^{15}N under elevated [CO$_2$] and water stress are not fully understood (Tcherkez, 2011).

Leaf chlorophyll content has been extensively used as an indicator of different physiological and agronomical components, particularly at later growth stages (Araus et al., 2008). The network analysis confirmed that chlorophyll content is a positive trait for vegetative growth in durum wheat, and this can be easily implemented in most of studies because this measurement is simple, quick, and non-destructive with modern portable devices. Effects of elevated $[CO_2]$ and water stress on g_s have been widely studied (Long et al., 2006; Pazzagli et al., 2016), including the proposal of g_s as a trait indicator of drought stress tolerance (Nagy et al., 2013). In our study g_s was negatively correlated with chlorophyll content and root and plant biomass. This could highlight that increased vegetative growth was related to stomatal closure, maybe as a water saving strategy or as a direct response to elevated $[CO_2]$.

The positive correlations among the transcript levels of the genes encoding RBCL, RBCS, GS1, GS2, and PEPC supported a balanced coordination between C and N metabolism under elevated $[CO_2]$ and water stress. On the other hand, our results underlined the key role of Rubisco and GS in plant responses to environmental conditions (Nagy et al., 2013; Carmo-Silva et al., 2015; Vicente et al., 2015b; Yousfi et al., 2016). We showed negative associations between transcript levels of Rubisco subunits and GS2 with chlorophyll content and plant biomass. This fact could indicate that a stimulation of plant growth may be associated with a lower investment of resources (mainly N) in Rubisco protein, especially under elevated $[CO_2]$, thus leading to a higher nitrogen efficiency (Pang et al., 2014; Carmo-Silva et al., 2015). The negative correlation between transcript levels of CAT and chlorophyll content highlighted that the up-regulation of *CAT* expression was a response to the high H_2O_2 levels generated under stress conditions (Luna et al., 2005), which could promote chlorophyll degradation (Upadhyaya et al., 2007). Interestingly, transcript levels of CAT were positively correlated with g_s, although a negative correlation should be expected since greater g_s leads to lower photorespiration rates and consequently lower H_2O_2 generation (Luna et al., 2005). We found a positive relationship between C content and transcript accumulation for SOD, not previously reported to our knowledge. Higher *SOD* expression might suggest a better ROS control that triggers an efficient electron transfer and C fixation. In our study, DHN11 transcript accumulation was negatively associated with root biomass, while transcripts for DHN16 were positively linked with plant biomass. These results suggest promising functions for DHN16 in stress tolerance during vegetative growth, as Kosová et al. (2014) proposed in a study examining wheat seed development.

In summary, parameters such as chlorophyll and N content, g_s and $\delta^{15}N$, and the expression of *RBCL*, *RBCS*, *GS2*, *DHN11*, and *DHN16* genes were identified as good indicators for the selection of genotypes with better performance during early plant growth under elevated CO_2 and water stress. Additionally, network analysis underlined the relevance of N metabolism-traits such as N content, $\delta^{15}N$, GS1, and GS2, in the genotypic response of durum wheat to future environmental scenarios in the Mediterranean basin.

CONCLUSION

We conclude that $[CO_2]$ effects on plant growth had greater impacts than moderate or severe water stress during vegetative growth of durum wheat. Whereas, elevated $[CO_2]$ generally led to increases in plant growth, water stress had a negative effect, preferentially as the water stress develops over time. In addition, the interactive effects of both $[CO_2]$ and water regime depends on genotypic variability. Gene expression profiles at moderate water stress were mainly affected by environmental conditions among the different genotypes. However, with further water restriction, genotype-specific differences were found to affect gene expression more than environmental conditions. These facts reflect a wide range of adaptation mechanisms in durum wheat under elevated $[CO_2]$ and water stress during vegetative growth, probably due to the complex regulatory network that takes place with both factors. Moreover, our study did not show a clear trend concerning the genetic advance in response to future climate change scenarios. Our results evidenced for durum wheat the need to take into account the genotypic variability for a greater understanding of plant adaptation to climate change. Moreover, the correlation network demonstrated that the combination of phenotyping and gene expression analysis is a useful approach to identify phenotype-genotype relationships and their behavior in response to different environments during vegetative stages.

AUTHOR CONTRIBUTIONS

SM and JLA conceived and designed the experiments. SM, RV, and AA contributed to the experimental work. SM, RV, and JLA analyzed the data and interpreted the results. RV wrote the paper under the supervision of JLA, and SM and AA revised the manuscript. All authors have read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01738/full#supplementary-material

REFERENCES

Alexandratos, N., and Bruinsma, J. (2012). *World Agriculture Towards 2030/2050: The 2012 Revision*. ESA Working paper No. 12-03, FAO, Rome.

Ali-Benali, M. A., Alary, R., Joudrier, P., and Gautier, M. F. (2005). Comparative expression of five Lea genes during wheat seed development and in response to abiotic stresses by real-time quantitative RT-PCR. *Biochim. Biophys. Acta* 1730, 56–65. doi: 10.1016/j.bbaexp.2005.05.011

Aljazairi, S., Arias, C., and Nogués, S. (2015). Carbon and nitrogen allocation and partitioning in traditional and modern wheat genotypes under pre-industrial and future CO$_2$ conditions. *Plant Biol.* 17, 647–659. doi: 10.1111/plb.12280

Aprile, A., Havlickova, L., Panna, R., Marè, C., Borrelli, G. M., Marone, D., et al. (2013). Different stress responsive strategies to drought and heat in two durum wheat cultivars with contrasting water use efficiency. *BMC Genomics* 14, 1–18. doi: 10.1186/1471-2164-14-821

Aranjuelo, I., Cabrera-Bosquet, L., Morcuende, R., Avice, J. C., Nogués, S., Araus, J. L., et al. (2011). Does ear C sink strength contribute to overcoming photosynthetic acclimation of wheat plants exposed to elevated CO$_2$? *J. Exp. Bot.* 62, 3957–3969. doi: 10.1093/jxb/err095

Aranjuelo, I., Sanz-Sáez, Á., Jauregui, I., Irigoyen, J. J., Araus, J. L., Sánchez-Díaz, M., et al. (2013). Harvest index, a parameter conditioning responsiveness of wheat plants to elevated CO$_2$. *J. Exp. Bot.* 64, 1879–1892. doi: 10.1093/jxb/ert081

Araus, J. L., Cabrera-Bosquet, L., Serret, M. D., Bort, J., and Nieto-Taladriz, M. T. (2013). Comparative performance of δ^{13}C, δ^{18}O and δ^{15}N for phenotyping durum wheat adaptation to a dryland environment. *Funct. Plant Biol.* 40, 595–608. doi: 10.1071/FP12254

Araus, J. L., Slafer, G. A., Reynolds, M. P., and Royo, C. (2002). Plant breeding and drought in C$_3$ cereals: what should we breed for? *Ann. Bot.* 89, 925–940. doi: 10.1093/aob/mcf049

Araus, J. L., Slafer, G. A., Royo, C., and Serret, M. D. (2008). Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant Sci.* 27, 377–412. doi: 10.1080/07352680802467736

Ariz, I., Cruz, C., Neves, T., Irigoyen, J. J., García, C., Nogués, S., et al. (2015). Leaf δ^{15}N as a physiological indicator of the responsiveness of N$_2$-fixing alfalfa plants to elevated [CO$_2$], temperature and low water availability. *Front. Plant Sci.* 6:574. doi: 10.3389/fpls.2015.00574

Bencze, S., Bamberger, Z., Janda, T., Balla, K., Varga, B., Bedö, Z., et al. (2014). Physiological response of wheat varieties to elevated atmospheric CO$_2$ and low water supply levels. *Photosynthetica* 52, 71–82. doi: 10.1007/s11099-014-0008-y

Bort, J., Belhaj, M., Latiri, K., Kehel, Z., and Araus, J. L. (2014). Comparative performance of the stable isotope signatures of carbon, nitrogen and oxygen in assessing early vigour and grain yield in durum wheat. *J. Agric. Sci.* 152, 408–426. doi: 10.1017/S0021859613000269

Budak, H., Kantar, M., and Yucebilgili Kurtoglu, K. (2013). Drought tolerance in modern and wild wheat. *Sci. World J.* 2013:548246. doi: 10.1155/2013/548246

Carmo-Silva, E., Scales, J. C., Madgwick, P. J., and Parry, M. A. J. (2015). Optimizing Rubisco and its regulation for greater resource use efficiency. *Plant Cell Environ.* 38, 1817–1832. doi: 10.1111/pce.12425

Caruso, G., Cavaliere, C., Foglia, P., Gubbiotti, R., Samperi, R., and Laganà, A. (2009). Analysis of drought responsive proteins in wheat (*Triticum durum*) by 2D-PAGE and MALDI-TOF mass spectrometry. *Plant Sci.* 177, 570–576. doi: 10.1016/j.plantsci.2009.08.007

Ceccarelli, S., Acevedo, E., and Grando, S. (1991). Breeding for yield stability in unpredictable environments: single traits, interaction between traits, and architecture of genotypes. *Euphytica* 56, 169–185. doi: 10.1007/BF000 42061

De Leonardis, A. M., Marone, D., Mazzucotelli, E., Neffar, F., Rizza, F., Di Fonzo, N., et al. (2007). Durum wheat genes up-regulated in the early phases of cold stress are modulated by drought in a developmental and genotype dependent manner. *Plant Sci.* 172, 1005–1016. doi: 10.1016/j.plantsci.2007.02.002

Elazab, A., Molero, G., Serret, M. D., and Araus, J. L. (2012). Root traits and δ^{13}C and δ^{18}O of durum wheat under different water regimes. *Funct. Plant Biol.* 39, 379–393. doi: 10.1071/FP11237

Erice, G., Sanz-Sáez, A., Urdiain, A., Araus, J. L., Irigoyen, J. J., and Aranjuelo, I. (2014). Harvest index combined with impaired N availability constrains the responsiveness of durum wheat to elevated CO$_2$ concentration and terminal water stress. *Funct. Plant Biol.* 41, 1138–1147. doi: 10.1071/FP14045

FAO (2013). *Food and Agriculture Organization of the United Nations, Statistics Division*. Available online at: http://faostat3.fao.org (accessed April 10, 2016)

Galmes, J., Medrano, H., and Flexas, J. (2007). Photosynthetic limitations in response to water stress and recovery in Mediterranean plants with different growth forms. *New Phytol.* 175, 81–93. doi: 10.1111/j.1469-8137.2007. 02087.x

González, M. C., Sánchez, R., and Cejudo, F. J. (2003). Abiotic stresses affecting water balance induce phosphoenolpyruvate carboxylase expression in roots of wheat seedlings. *Planta* 216, 985–992. doi: 10.1007/s00425-002-0951-x

Guo, Z., Slafer, G. A., and Schnurbusch, T. (2016). Genotypic variation in spike fertility traits and ovary size as determinants of floret and grain survival rate in wheat. *J. Exp. Bot.* 67, 4221–4230. doi: 10.1093/jxb/erw200

Habash, D. Z., Baudo, M., Hindle, M., Powers, S. J., Defoin-Platel, M., Mitchell, R., et al. (2014). Systems responses to progressive water stress in durum wheat. *PLoS ONE* 9:e108431. doi: 10.1371/journal.pone.0108431

Habash, D. Z., Kehel, Z., and Nachit, M. (2009). Genomic approaches for designing durum wheat ready for climate change with a focus on drought. *J. Exp. Bot.* 60, 2805–2815. doi: 10.1093/jxb/erp211

Harnos, N., Bencze, S., Janda, T., Juhász, A., and Veisz, O. (2002). Interactions between elevated CO$_2$ and water stress in two winter wheat cultivars differing in drought resistance. *Cereal Res. Commun.* 30, 359–366.

Hayano-Kanashiro, C., Calderón-Vázquez, C., Ibarra-Laclette, E., Herrera-Estrella, L., and Simpson, J. (2009). Analysis of gene expression and physiological responses in three Mexican maize landraces under drought stress and recovery irrigation. *PLoS ONE* 4:e7531. doi: 10.1371/journal.pone.0007531

Hudak, C., Bender, J., Weigel, H. J., and Miller, J. (1999). Interactive effects of elevated CO$_2$, O$_3$, and soil water deficit on spring wheat (*Triticum aestivum* L. cv. Nandu). *Agronomie* 19, 677–687. doi: 10.1051/agro:19990803

Huseynova, I. M., Aliyeva, D. R., and Aliyev, J. A. (2014). Subcellular localization and responses of superoxide dismutase isoforms in local wheat varieties subjected to continuous soil drought. *Plant Physiol. Bioch.* 81, 54–60. doi: 10. 1016/j.plaphy.2014.01.018

IPCC (2013). *Climate Change 2013: The Physical Science Basis*. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, New York, NY.

Kano-Nakata, M., Tatsumi, J., Inukai, Y., Asanuma, S., and Yamauchi, A. (2014). Effect of various intensities of drought stress on δ^{13}C variation among plant organs in rice: comparison of two cultivars. *Am. J. Plant Sci.* 5, 1686–1693. doi: 10.4236/ajps.2014.511183

Khoshro, H. H., Taleei, A., Bihamta, M. R., Shahbazi, M., and Abbasi, A. (2013). Expression analysis of the genes involved in osmotic adjustment in bread wheat (*Triticum aestivum* L.) cultivars under terminal drought stress conditions. *J. Crop Sci. Biotechnol.* 16, 173–181. doi: 10.1007/s12892-013-0040-7

Kim, S. H., Sicher, R. C., Bae, H., Gitz, D. C., Baker, J. T., Timlin, D. J., et al. (2006). Canopy photosynthesis, evapotranspiration, leaf nitrogen, and transcription profiles of maize in response to CO$_2$ enrichment. *Global Change Biol.* 12, 588–600. doi: 10.1111/j.1365-2486.2006.01110.x

Kosová, K., Vítámvás, P., and Prášil, I. T. (2014). Wheat and barley dehydrins under cold, drought, and salinity – what can LEA-II proteins tell us about plant stress response? *Front. Plant Sci.* 5:343. doi: 10.3389/fpls.2014.00343

Li, P., Ainsworth, E. A., Leakey, A. D. B., Ulanov, A., Lozovaya, V., Ort, D. R., et al. (2008). *Arabidopsis* transcript and metabolite profiles: ecotype-specific responses to open-air elevated CO$_2$. *Plant Cell Environ.* 31, 1673–1687. doi: 10. 1111/j.1365-3040.2008.01874.x

Liu, E. K., Mei, X. R., Yan, C. R., Gong, D. Z., and Zhang, Y. Q. (2016). Effects of water stress on photosynthetic characteristics, dry matter translocation and WUE in two winter wheat genotypes. *Agr. Water Manage.* 167, 75–85. doi: 10. 1016/j.agwat.2015.12.026

Long, S. P., Ainsworth, E. A., Leakey, A. D. B., Nösberger, J., and Ort, D. R. (2006). Food for thought: lower-than-expected crop yield stimulation with rising CO$_2$ concentrations. *Science* 312, 1918–1921. doi: 10.1126/science.1114722

Lopes, M. S., and Araus, J. L. (2006). Nitrogen source and water regime effects on durum wheat photosynthesis and stable carbon and nitrogen isotope composition. *Physiol. Plantarum* 126, 435–445. doi: 10.1111/j.1399-3054.2006. 00595.x

Luna, C. M., Pastori, G. M., Driscoll, S., Groten, K., Bernard, S., and Foyer, C. H. (2005). Drought controls on H$_2$O$_2$ accumulation, catalase (CAT) activity and *CAT* gene expression in wheat. *J. Exp. Bot.* 56, 417–423. doi: 10.1093/jxb/eri039

Ma, F., Xu, T. T., Ji, M. F., and Zhao, C. M. (2014). Differential drought tolerance in tree populations from contrasting elevations. *AoB Plants* 6:plu069. doi: 10.1093/aobpla/plu069

Madhu, M., and Hatfield, J. L. (2013). Dynamics of plant root growth under increased atmospheric carbon dioxide. *Agron. J.* 105, 657–669. doi: 10.2134/agronj2013.0018

Marti, J., Bort, J., Slafer, G. A., and Araus, J. L. (2007). Can wheat yield be assessed by early measurements of Normalized Difference Vegetation Index? *Ann. Appl. Biol.* 150, 253–257. doi: 10.1111/j.1744-7348.2007.00126.x

Martins, M. Q., Rodrigues, W. P., Fortunato, A. S., Leitão, A. E., Rodrigues, A. P., Pais, I. P., et al. (2016). Protective response mechanisms to heat stress in interaction with high [CO$_2$] conditions in *Coffea* spp. *Front. Plant Sci.* 7:947. doi: 10.3389/fpls.2016.00947

Maydup, M. L., Graciano, C., Guiamet, J. J., and Tambussi, E. A. (2012). Analysis of early vigour in twenty modern cultivars of bread wheat (*Triticum aestivum* L.). *Crop Pasture Sci.* 63, 987–996. doi: 10.1071/CP12169

McKersie, B. (2015). Planning for food security in a changing climate. *J. Exp. Bot.* 66, 3435–3450. doi: 10.1093/jxb/eru547

Melloul, M., Iraqi, D., Udupa, S. M., Erba, G., Alaoui, M. A. E., Ibriz, M., et al. (2013). Analysis of mRNA levels of ten genes under water stress in *Triticum turgidum* subsp. durum. *J. Plant Stud.* 3, 65–79. doi: 10.5539/jps.v3n1p65

Nagy, Z., Nemeth, E., Guoth, A., Bona, L., Wodala, B., and Pecsvaradi, A. (2013). Metabolic indicators of drought stress tolerance in wheat: glutamine synthetase isoenzymes and Rubisco. *Plant Physiol. Biochem.* 67, 48–54. doi: 10.1016/j.plaphy.2013.03.001

Nakhforoosh, A., Grausgruber, H., Kaul, H. P., and Bodner, G. (2015). Dissection of drought response of modern and underutilized wheat varieties according to Passioura's yield-water framework. *Front. Plant Sci.* 6:570. doi: 10.3389/fpls.2015.00570

Nezhadahmadi, A., Prodhan, Z. H., and Faruq, G. (2013). Drought tolerance in wheat. *Sci. World J.* 2013:610721. doi: 10.1155/2013/610721

Pang, J., Palta, J. A., Rebetzke, G. J., and Milroy, S. P. (2014). Wheat genotypes with high early vigour accumulate more nitrogen and have higher photosynthetic nitrogen use efficiency during early growth. *Funct. Plant Biol.* 41, 215–222. doi: 10.1071/FP13143

Pazzagli, P. T., Weiner, J., and Liu, F. (2016). Effects of CO$_2$ elevation and irrigation regimes on leaf gas exchange, plant water relations, and water use efficiency of two tomato cultivars. *Agr. Water Manage.* 169, 26–33. doi: 10.1016/j.agwat.2016.02.015

Peremarti, A., Mare, C., Aprile, A., Roncaglia, E., Cattivelli, L., Villegas, D., et al. (2014). Transcriptomic and proteomic analyses of a pale-green durum wheat mutant shows variations in photosystem components and metabolic deficiencies under drought stress. *BMC Genomics* 15:125. doi: 10.1186/1471-2164-15-125

Pleijel, H., Gelang, J., Sild, E., Danielsson, H., Younis, S., Karlsson, P. E., et al. (2000). Effects of elevated carbon dioxide, ozone and water availability on spring wheat growth and yield. *Physiol. Plantarum* 108, 61–70. doi: 10.1034/j.1399-3054.2000.108001061.x

Qin, N., Xu, W., Hu, L., Li, Y., Wang, H., Qi, X., et al. (2015). Drought tolerance and proteomics studies of transgenic wheat containing the maize C4 phosphoenolpyruvate carboxylase (PEPC) gene. *Protoplasma* 253, 1503–1512. doi: 10.1007/s00709-015-0906-2

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Viena: R Foundation for Statistical Computing.

Rebolledo, M. C., Luquet, D., Courtois, B., Henry, A., Soulié, J. C., Rouan, L., et al. (2013). Can early vigour occur in combination with drought tolerance and efficient water use in rice genotypes? *Funct. Plant Biol.* 40, 582–594. doi: 10.1071/FP12312

Richards, R. A. (2000). Selectable traits to increase crop photosynthesis and yield of grain crops. *J. Exp. Bot.* 51, 447–458. doi: 10.1093/jexbot/51.suppl_1.447

Robredo, A., Pérez-López, U., Miranda-Apodaca, J., Lacuesta, M., Mena-Petite, A., and Muñoz-Rueda, A. (2011). Elevated CO$_2$ reduces the drought effect on nitrogen metabolism in barley plants during drought and subsequent recovery. *Environ. Exp. Bot.* 71, 399–408. doi: 10.1016/j.envexpbot.2011.02.011

Russo, A. C., Gouveia, C. M., Trigo, R. M., Liberato, M. L. R., and DaCamara, C. (2015). The influence of circulation weather patterns at different spatial scales on drought variability in the Iberian Peninsula. *Front. Environ. Sci.* 3:1. doi: 10.3389/fenvs.2015.00001

Sanchez-Garcia, M., Royo, C., Aparicio, N., Martin-Sanchez, J. A., and Alvaro, F. (2013). Genetic improvement of bread wheat yield and associated traits in Spain during the 20th century. *J. Agric. Sci.* 151, 105–118. doi: 10.1017/S0021859612000330

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C$_T$ method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Soriano, J. M., Villegas, D., Aranzana, M. J., García del Moral, L. F., and Royo, C. (2016). Genetic structure of modern durum wheat cultivars and mediterranean landraces matches with their agronomic performance. *PLoS ONE* 11:e0160983. doi: 10.1371/journal.pone.0160983

Stitt, M., and Krapp, A. (1999). The interaction between elevated carbon dioxide and nitrogen nutrition: the physiological and molecular background. *Plant Cell Environ.* 22, 583–621. doi: 10.1046/j.1365-3040.1999.00386.x

Suzuki, Y., and Makino, A. (2012). Availability of Rubisco small subunit up-regulates the transcript levels of large subunit for stoichiometric assembly of its holoenzyme in rice. *Plant Physiol.* 160, 533–540. doi: 10.1104/pp.112.201459

Tardieu, F. (2013). Plant response to environmental conditions: assessing potential production, water demand, and negative effects of water deficit. *Front. Plant Sci.* 4:17. doi: 10.3389/fphys.2013.00017

Taub, D. R., and Wang, X. (2008). Why are nitrogen concentrations in plant tissues lower under elevated CO$_2$? A critical examination of the hypotheses. *J. Integr. Plant Biol.* 50, 1365–1374. doi: 10.1111/j.1744-7909.2008.00754.x

Tcherkez, G. (2011). Natural ^{15}N/^{14}N isotope composition in C$_3$ leaves: are enzymatic isotope effects informative for predicting the ^{15}N-abundance in key metabolites? *Funct. Plant Biol.* 38, 1–12. doi: 10.1071/FP10091

Trnka, M., Rotter, R. P., Ruiz-Ramos, M., Kersebaum, K. C., Olesen, J. E., Zalud, Z., et al. (2014). Adverse weather conditions for European wheat production will become more frequent with climate change. *Nat. Clim. Change* 4, 637–643. doi: 10.1038/nclimate2242

Upadhyaya, H., Khan, M. H., and Panda, S. K. (2007). Hydrogen peroxide induces oxidative stress in detached leaves or *Oryza sativa* L. *Gen. Appl. Plant Physiol.* 33, 83–95.

Vicente, R., Martínez-Carrasco, R., Pérez, P., and Morcuende, R. (2016). An association network reveals co-regulation of carbon and nitrogen metabolism-related parameters in durum wheat grown under different environmental conditions. *New Biotechnol.* 33, 414. doi: 10.1016/j.nbt.2015.10.079

Vicente, R., Morcuende, R., and Babiano, J. (2011). Differences in Rubisco and chlorophyll content among tissues and growth stages in two tomato (*Lycopersicon esculentum* Mill.) varieties. *Agron. Res.* 9, 501–507.

Vicente, R., Pérez, P., Martínez-Carrasco, R., Gutiérrez, E., and Morcuende, R. (2015a). Nitrate supply and plant development influence nitrogen uptake and allocation under elevated CO$_2$ in durum wheat grown hydroponically. *Acta Physiol. Plant.* 37, 114. doi: 10.1007/s11738-015-1867-y

Vicente, R., Pérez, P., Martínez-Carrasco, R., Usadel, B., Kostadinova, S., and Morcuende, R. (2015b). Quantitative RT-PCR platform to measure transcript levels of C and N metabolism-related genes in durum wheat: transcript profiles in elevated [CO$_2$] and high temperature at different nitrogen supplies. *Plant Cell Physiol.* 56, 1556–1573. doi: 10.1093/pcp/pcv079

Wall, G. W., Garcia, R. L., Kimball, B. A., Hunsaker, D. J., Pinter, P. J., Long, S. P., et al. (2006). Interactive effects of elevated carbon dioxide and drought on wheat. *Agron. J.* 98, 354–381. doi: 10.2134/agronj2004.0089

Wilson, P. B., Rebetzke, G. J., and Condon, A. G. (2015). Of growing importance: combining greater early vigour and transpiration efficiency for wheat in variable rainfed environments. *Funct. Plant Biol.* 42, 1107–1115. doi: 10.1071/fp15228

Xu, Z., Zhou, G., and Shimizu, H. (2010). Plant responses to drought and rewatering. *Plant Signal. Behav.* 5, 649–654. doi: 10.4161/psb.5.6.11398

Yousfi, S., Márquez, A. J., Betti, M., Araus, J. L., and Serret, M. D. (2016). Gene expression and physiological responses to salinity and water stress of contrasting durum wheat genotypes. *J. Integr. Plant Biol.* 58, 48–66. doi: 10.1111/jipb.12359

Yousfi, S., Serret, M. D., Márquez, A. J., Voltas, J., and Araus, J. L. (2012). Combined use of δ^{13}C, δ^{18}O and δ^{15}N tracks nitrogen metabolism and genotypic adaptation of durum wheat to salinity and water deficit. *New Phytol.* 194, 230–244. doi: 10.1111/j.1469-8137.2011.04036.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

SNP-Based QTL Mapping of 15 Complex Traits in Barley under Rain-Fed and Well-Watered Conditions by a Mixed Modeling Approach

Freddy Mora[1], Yerko A. Quitral[2], Ivan Matus[3], Joanne Russell[4], Robbie Waugh[4] and Alejandro del Pozo[2]*

[1] Instituto de Ciencias Biológicas, Área de Biología Molecular y Biotecnología, Universidad de Talca, Talca, Chile, [2] Centro de Mejoramiento Genético y Fenómica Vegetal, Facultad de Ciencias Agrarias, PIEI Adaptación de la Agricultura al Cambio Climático (A2C2), Universidad de Talca, Talca, Chile, [3] Centro Regional de Investigación Quilamapu, Instituto de Investigaciones Agropecuarias, Chillán, Chile, [4] The James Hutton Institute, Dundee, Scotland

This study identified single nucleotide polymorphism (SNP) markers associated with 15 complex traits in a breeding population of barley (*Hordeum vulgare* L.) consisting of 137 recombinant chromosome substitution lines (RCSL), evaluated under contrasting water availability conditions in the Mediterranean climatic region of central Chile. Given that markers showed a very strong segregation distortion, a quantitative trait locus/loci (QTL) mapping mixed model was used to account for the heterogeneity in genetic relatedness between genotypes. Fifty-seven QTL were detected under rain-fed conditions, which accounted for 5–22% of the phenotypic variation. In full irrigation conditions, 84 SNPs were significantly associated with the traits studied, explaining 5–35% of phenotypic variation. Most of the QTL were co-localized on chromosomes 2H and 3H. Environment-specific genomic regions were detected for 12 of the 15 traits scored. Although most QTL-trait associations were environment and trait specific, some important and stable associations were also detected. In full irrigation conditions, a relatively major genomic region was found underlying hectoliter weight (HW), on chromosome 1H, which explained between 27% (SNP 2711-234) and 35% (SNP 1923-265) of the phenotypic variation. Interestingly, the locus 1923-265 was also detected for grain yield at both environmental conditions, accounting for 9 and 18%, in the rain-fed and irrigation conditions, respectively. Analysis of QTL in this breeding population identified significant genomic regions that can be used for marker-assisted selection (MAS) of barley in areas where drought is a significant constraint.

Keywords: drought, marker segregation distortion, RCSL, physiological trait, kinship

INTRODUCTION

Abiotic stresses can significantly reduce crop yields and restrict the latitudes and soils on which commercially important species can be cultivated (Lobell and Field, 2007; Jacobsen et al., 2012). The development of drought-tolerant genotypes as well as genotypes with higher water-use efficiency is of global interest as populations continue to increase and water availability decreases (Araus et al., 2002, 2008; Cattivelli et al., 2008). In barley (*Hordeum vulgare* L.), a large number of morphological and physiological traits are linked to drought tolerance (Chen et al., 2010; del Pozo et al., 2012) which exhibit strong environmental interactions (Tondelli et al., 2006). Increasing tolerance to drought stress has become a major goal for barley breeding programs particularly in light of prolonged drought periods as a result of climate change (Wehner et al., 2015).

To determine the genetic basis of complex traits, important genetic, and genomic resources have been developed in a wide range of species (Kota et al., 2003; Mora et al., 2015), including barley (Kota et al., 2003; Close et al., 2009; Zhou et al., 2015). Single nucleotide polymorphism (SNP) markers are the most abundant sequence variations encountered in eukaryotic genomes (Griffin and Smith, 2000), and in barley offer the potential for generating very high-density genetic maps (Close et al., 2009), providing a useful tool for quantitative trait locus/loci (QTL) mapping for marker-assisted selection (MAS; Sato and Takeda, 2009; Szűcs et al., 2009).

Conventionally, quantitative trait locus/loci (QTL) mapping is carried out using markers that follow a Mendelian segregation ratio, which depends on the population under investigation (Xu, 2008). The phenomenon that alleles at a locus deviate from the Mendelian expectation has been defined as segregation distortion (Zhang et al., 2010), which has been encountered in many commercially important species, such as maize (Lu et al., 2002), rice (Xu et al., 1997), tomato (Paterson et al., 1988), and barley (Malosetti et al., 2011). Biologically, segregation distortion can be due to chromosome loss, genetic isolation mechanisms, and the presence of viability genes. Non-biological factors such as scoring errors and sampling errors also can contribute to segregation distortion (Alheit et al., 2011).

Construction of genetic maps and QTL analysis using distorted markers is risky because the basic assumption of Mendelian segregation is violated, and, consequently most of these markers are removed from the subsequent QTL analyses (Luo et al., 2005; Xu, 2008). Moreover, several authors found that distorted markers influence the estimation of genetic intervals and the markers order on a same chromosome (Lorieux et al., 1995a,b; Zhu et al., 2007). Because these markers are routinely removed, valuable information around these regions is lost (Luo et al., 2005). However, if distorted markers are handled properly, their effects on genetic map construction and QTL identification and mapping can be significantly improved (Xu, 2008; Alheit et al., 2011; Hashemi et al., 2015).

From an analytical standpoint, many QTL mapping procedures have been developed for Mendelian populations, but few are available for markers that do not segregate in a typical Mendelian fashion (Zhu et al., 2007; Xu, 2008). Recently,

Malosetti et al. (2011) found severe allele frequency distortions in many chromosomal regions in a RIL population of barley. The authors stated that violation of the basic assumptions implies that the genetic covariance between genotypes (i.e., genetic relatedness) in the population is not homogeneous, and analogous to association mapping studies, a QTL mapping mixed model that account for this heterogeneity should be used to avoid false QTL detection. In addition, mixed models have been used to investigate QTL-by-environment interaction (Malosetti et al., 2004; Boer et al., 2007) and to map QTL for several traits simultaneously (Malosetti et al., 2008).

In barley, a large number of mapping populations have been developed to map QTL. Further, advanced mapping populations, including near-isogenic lines (NILs), chromosome segment substitution lines (CSSLs), and recombinant chromosome substitution lines (RCSLs, Schmalenbach et al., 2008, 2011; Sato and Takeda, 2009; Naz et al., 2012), have also been developed to facilitate the genetic dissection of complex traits. As a consequence, many QTL controlling complex traits including agronomic, and morphological traits, yield component, disease resistance, tolerance to abiotic stress, and malting quality have been identified (Pillen et al., 2004; Talame et al., 2004; Li et al., 2005; Gyenis et al., 2007). However, studies that have identified QTL for drought-related morphological and physiological traits are still scarce in barley (Chen et al., 2010; Mir et al., 2012; Sayed et al., 2012; Kalladan et al., 2013; Li et al., 2013; Wójcik-Jagła et al., 2013; Honsdorf et al., 2014; Mansour et al., 2014; Naz et al., 2014). Thus, a breeding program has been developed in the Mediterranean area of central Chile, in areas where drought is a significant constraint to yield. This study identifies SNP markers associated with 15 complex traits (including physiological and morphological traits) in a breeding population consisting of 137 RCSLs of barley (Matus et al., 2003), evaluated under contrasting water availability. Given that markers showed a very strong segregation distortion, a QTL mapping mixed model was employed to account for the heterogeneity in genetic relatedness between genotypes.

MATERIALS AND METHODS
Plant Material and Field Evaluation

A set of 137 RCSL were evaluated in field conditions in Santa Rosa (35°78′ S, 72°17′ W), in the Mediterranean climatic region of central Chile, under rain-fed and fully irrigated conditions in 2008–2009. The average annual temperature in this region is 13°C, the minimum average is 3°C (July), and the maximum 28.6°C (January). Monthly maximum and minimum temperatures and precipitation during 2008 are in Table S1. The annual precipitation in 2008 was 992 mm but the amount during the growing season was 120 mm. The soil is a sandy loam, Humic Haploxerand, Andisol. For fully irrigated conditions, four furrow irrigation of 50 mm water was applied from heading to maturity.

The RCSL population was developed using the advanced backcross strategy of Tanksley and Nelson (1996). The accession of *Hordeum spontaneum* (Caeserea 26-24 from Israel) was the donor parent and *Hordeum vulgare* subsp. *vulgare* "Harrington" (a North American malting quality standard) was the recurrent

parent (Matus et al., 2003). The recurrent parent was used as the female and the donor as the male to obtain the F_1 generation. Finally, the lines were obtained using two backcrosses with the recurrent parent and six generations of self-pollination (BC_2F_6).

The field trial was arranged in a 14×10 alpha-lattice design with two replications ("Harrington" cultivar was replicated more times for arrangement $v = 14 \times 10 = 140$). Fertilizer and field management practices recommended for optimum barley production were used (Inostroza et al., 2009; del Pozo et al., 2012). The 15 different morphological, agronomic, and physiological traits measured during this experiment are described in **Table 1**.

Phenotypic Data Analysis

A mixed linear modeling approach was employed for phenotypic data analysis using the MIXED procedure in SAS. Field data were analyzed on the basis of the statistical model (Stich et al., 2008):

$$y_{ijno} = \mu + g_i + l_j + (gl)_{ij} + r_{nj} + b_{onj} + \varepsilon_{ijno}$$

where y_{ijno} is the observed phenotype for the ith RCSL at the jth location in the oth incomplete block of the nth replicate, μ is an intercept term (overall population mean), g_i is the genotypic effect of the ith entry, l_j is the effect of the jth location, r_{nj} is the effect of the nth replicate of the jth location, b_{onj} is the effect of the oth incomplete block of the nth replication of the jth location, and ε_{ijno} is the residual or within-plot error. To analyze the effect of environment (rain-fed and fully irrigated) on the RCSL population, environment (contrasting environmental condition), genotype, and genotype-environment interaction were considered as fixed effects while replication and block were analyzed as random effects. Error variances were assumed to be heterogeneous among locations according to Stich

et al. (2008). Adjusted entry means M_i were calculated for each RCSL as:

$$M_i = \hat{\mu} + \hat{g}_i$$

where $\hat{\mu}$ and $\hat{g}_i$ are the generalized least-squares estimates of μ and g_i, respectively. Adjusted entry means, calculated by environment and in a combined analysis, were latter used in the analysis of QTL. The CORR procedure of SAS was used to estimate Pearson correlations (r, $n = 137$) between pairs of traits.

DNA Extraction and Genotyping

Leaf tissues were harvested from each plant and crushed with liquid nitrogen. Genomic DNA was extracted from 200 to 300 mg of leaf sample using Qiagen DNeasy Plant mini kit (QIAGEN Co.). DNA concentration was determined by nanodrop (Thermo Fisher Scientific) and adjusted to 100 ng/μl. DNA samples were sent to the Southern California Genotyping Consortium, Illumina BeadLab at the University of California, Los Angeles (UCLA) for the OPA-SNP assay with the 1536-plex detection platform of barley OPA 1 (BOPA1) developed by Close et al. (2009). SNP markers are distributed across the seven barley chromosomes. The SNP loci were designated by HarvEST:Barley unigene assembly #32 numbers (http://harvest.ucr.edu/). Genotyping of the 137 RCSLs and their parents was performed by Illumina GoldenGate assay. The order of polymorphic markers from BOPA1 was performed using MEGA5 software (Tamura et al., 2011). Moreover, the chromosome segments introgressed into *H. vulgare* cv. Harrington from *Hordeum spontaneum* were estimated from the graphical haplotypes (Van Berloo et al., 2008) in each of the recombinant lines selected.

QTL Mapping by a Mixed Modeling Approach

For the segregation data of each SNP marker, deviations from the Mendelian ratios (1:1 ratio for RSCL populations) were tested using the Chi-square test. Given that SNP markers did not segregate in an expected Mendelian ratio (details in the Results Section) a mixed model (Stich et al., 2008) was employed for QTL analysis of a designed cross using a structured variance–covariance matrix, where the structure was induced by selection (Malosetti et al., 2011). This QTL mapping mixed model accounts for the heterogeneity in genetic relatedness between genotypes. The hypothesis of association of SNP markers with the 15 target traits was tested using the following mixed model implemented in the program TASSEL 3.0 (Bradbury et al., 2007):

$$M_i = \mu_1 + x_i\alpha + u_i + e_i \tag{1}$$

where M_i is the adjusted entry mean of the ith RCSL, μ_1 is an intercept term, x_i is the SNP genotype of ith RCSL, α is the additive allele substitution effect (SNP effect), u_i is the residual genetic background effect of the ith entry, and e_i is the random residual effect. u_i is assumed to follow a normal distribution with variance-covariance matrix G $= \sigma_u^2 \cdot 2 \cdot K$, where K is the coefficient of co-ancestry

TABLE 1 | List of 15 morphologic, agronomic, and physiological traits measured in a RCSL population of barley under two contrasting environmental conditions in southern Chile.

Trait	Abbreviations	Type	Unit
Peduncle length	PL	Morphologic	cm
peduncle extrusion	PE	Morphologic	cm
Spike length excluding the awns	SL	Morphologic	cm
Plant height	PH	Agronomic	cm
Tillers	TN	Agronomic	Number per m^{-2}
Dry weight in tillering	DWT	Agronomic	g·m^{-2}
Biological yield	BY	Agronomic	g·m^{-2}
Hectoliter weight	HW	Agronomic	g
Harvest index	HI	Agronomic	—
Kernel per spike	KS	Agronomic	Number
Thousand kernel weight	TKW	Agronomic	g
Grain yield	GY	Agronomic	ton·ha^{-1}
Relative water content	RWC	Physiological	%
Intercept PAR	IPAR	Physiological	—
Chlorophyll fluorescence at anthesis	Fv/Fm	Physiological	—

matrix between entries. The threshold used for declaring an association significant was $P < 0.01$. According to Malosetti et al. (2011) this mixed modeling approach for QTL detection can accommodate the extra genetic covariance by embedding kinship information in the model, leading to appropriate tests, and minimizing the rate of false QTL or gene detection. Variance components were estimated using the Restricted Maximum Likelihood (REML) method. Additionally, false discovery rate (FDR)-adjusted p-values were calculated using PROC MULTTEST in SAS software. The Bayesian information criterion (BIC) was used to compare the simple model that ignore kinship information and the model with a structured variance–covariance matrix (kinship information).

RESULTS

Phenotypic Data Analysis

The statistical analysis of fixed effects for the 15 complex traits under study are summarized in **Table 2**. According to the F-values (type III tests of fixed effects), the 137 RCSL presented significant differences at $P < 0.01$ in most traits under study. Significant differences were observed between both contrasting water regimes for most traits, including biological yield (BY) and grain yield (GY). In fact, the average GY under rain-fed was reduced by 81% in relation to fully irrigated condition, indicating that the barley RCSL lines were exposed to a severe water stress. Also, there was

a strong reduction in BY (−59.9%), dry weight at tillering (DWT; −39.9%), tiller number (TN; −48.1%) and thousand-kernel weight (TKW; −28.1%) (**Table 2**). Peduncle length (PL), peduncle extrusion (PE), plant height (PH), and the physiological trait IPAR also evidenced significant environmental effect. There were significant interactions between RCSL and environment for PE, PH, TN, hectoliter weight (HW), kernels per spike (KS), and GY.

Estimates of phenotypic correlations among traits are shown in **Table 3**. Grain yield was positively and significantly correlated with TN, BY, HW, HI, and KS, and negatively correlated with PL and PH ($P < 0.01$, **Table 3**). The physiological trait IPAR was significantly and positively correlated with biological yield ($r = 0.6$) and grain yield ($r = 0.41$). Relative water content (RWC), PE, and chlorophyll fluorescence (Fv/Fm) were not significantly correlated ($P > 0.05$) with most of the variables studied.

Kinship Matrix to Avoid False-Positives

A large number of spurious QTLs were detected when the genetic covariance matrix is ignored in the mixed model. In fact, 444, 460, and 516 false-positives were detected for all traits studied in the rain-fed and fully irrigated conditions, and combined analysis, respectively, using a structured variance–covariance matrix in a mixed model, where the structure was induced by selection. This QTL mapping mixed model accounts for the heterogeneity in genetic relatedness between genotypes. **Figure 1** shows the results of segregation distortion analysis based on the P-values, obtained by the chi-square test (P-values were plotted

TABLE 2 | Summary of statistical analysis of fixed effects for 15 complex traits measured in a RCSL population of barley under rain-fed and well-watered conditions.

Trait	Environmental condition		Environment effect	GEI	RCSL
	Rain-fed	Fully irrigated			
MORPHOLOGIC					
PL (cm)	17.8 (11.9–28.2)	22.8 (14.0–34.8)	**	NS	**
PE (mm)	2.7 (1.3–3.5)	2.4 (0.5–4.6)	**	*	*
SL (cm)	9.4 (5.6–21.6)	9.4 (6.2–12.1)	NS	NS	**
AGRONOMIC					
PH (cm)	61.9 (40–85)	83.7 (65–120)	**	**	**
TN (number·m^{-2})	357.1 (115–745)	688.3 (160–1235)	**	*	**
DWT (g·m^{-2})	121.9 (24.9–404.8)	201.0 (20.2–934.4)	**	NS	NS
BY (g·m^{-2})	434.6 (75–1325)	1085.8 (200–1950)	**	NS	*
HW	67.6 (51.3–73.2)	69.4 (63–9.1)	NS	*	NS
HI	0.3 (0.03–0.65)	0.3 (0.1–0.5)	NS	NS	**
KS	17.5 (7.9–28)	21.4 (10.8–29.3)	NS	*	**
TKW	40.1 (12.8–69.7)	55.8 (37.7–89.1)	**	NS	NS
GY (ton·ha^{-1})	1.0 (0.01–3.03)	5.3 (1.1–8.5)	**	**	**
PHYSIOLOGICAL					
RWC (%)	83.5 (52–101)	86.0 (43–127)	NS	NS	NS
IPAR	35.4 (7.5–59.9)	51.8 (14.6–79.8)	**	NS	*
Fv/Fm	0.61 (0.3–0.7)	0.65 (0.5–0.7)	NS	NS	NS

Data are presented as phenotypic means with minimum and maximum values in parentheses. GEI, genotype-environment interactions; NS, not significant.
*Significant at the 0.05 probability level.
**Significant at the 0.01 probability level.

TABLE 3 | Pearson correlation coefficients among mean variables (adjusted entry means) measured in a RCSL population of barley under rain-fed and well-watered conditions.

Trait	PL	PE	SL	PH	TN	DWT	BY	HW	HI	KS	TKW	GY	RWC	IPAR	FV/FM
PL	1														
PE	0.08	1													
SL	0.11	0.15	1												
PH	0.37**	0.19*	0.37**	1											
TN	−0.21**	−0.15	−0.31**	−0.48**	1										
DWT	−0.01	0.15	0.18*	0.12	0.16	1									
BY	−0.05	−0.10	0.23**	0.05	0.57**	0.28**	1								
HW	−0.21*	−0.08	−0.05	−0.03	0.13	0.04	0.17*	1							
HI	−0.12	−0.11	−0.08	−0.50**	0.52**	0.20*	0.40**	0.26**	1						
KS	−0.19*	−0.04	0.64**	0.15	−0.01	0.17*	0.47**	0.18*	0.24**	1					
TKW	0.22**	0.24**	0.19*	0.07	−0.17*	0.21*	0.06	0.07	0.20*	0.06	1				
GY	−0.35**	−0.07	0.09	−0.30**	0.44**	0.14	0.49**	0.48**	0.61**	0.47**	−0.08	1			
RWC	−0.06	0.09	−0.13	−0.17*	0.10	0.06	0.11	−0.08	0.14	−0.13	0.10	0.06	1		
IPAR	−0.11	−0.01	0.22*	0.14	0.26**	0.25**	0.60**	0.09	0.11	0.34**	−0.16	0.41**	0.05	1	
Fv/Fm	0.13	0.05	0.14	0.09	−0.10	−0.17*	−0.09	0.01	−0.09	0.07	−0.12	0.03	−0.11	−0.02	1

Significant at the 0.05 probability level.
**Significant at the 0.01 probability level.*

FIGURE 1 | Chi-square test for segregation distortion of SNP markers along the seven linkage groups in the RCSL population of barley.

on a −Log10 scale). According to the Bayesian information criterion (BIC) there was significant evidence against the simple models that ignore kinship information in most of the traits studied (Table S2); between the two competing models, the best model is the one that has the smallest BIC-value. In the combined analysis, BIC was smaller for the mixed model that includes the kinship information (K model), for most of the traits. Therefore, the use of a QTL detection model that accounts for the heterogeneous genetic relatedness between RCSLs lines, caused by the uneven sharing of genetic background, is clearly necessary.

SNP-Based QTL Mapping of Complex Traits

The number of QTLs detected for all the traits under study, including the chromosome number and the percentage of the total phenotypic variation explained by the QTLs, are summarized in **Table 4** (details of all significant SNP–trait associations are given in Table S3). The Manhattan plots of the genome-wide QTL study showed a lower number of significant QTLs under rain-fed (**Figure 2**) in comparison with fully irrigated conditions (**Figure 3**). In fact, 57 QTLs were detected in rain-fed, which accounted for 5–22% of the phenotypic variation (KS and PH, respectively). In the fully irrigated condition, a total of 84 SNPs were significantly associated with the traits studied, which explained from 5 to 35% of phenotypic variation (TN and HW, respectively). In the combined analysis, 92 QTLs were detected to be highly significant, which accounted for 5–21% of the phenotypic variation.

Most of the QTLs were co-localized on chromosomes 2H and 3H; i.e., about 53% (30/57) and 62% (52/84) in the rain-fed and irrigation conditions, respectively. The chromosomes with lowest number of QTLs found were 4H and 6H, with two QTLs each, which accounted for 5.1–8.3% of the phenotypic variation. Only two QTLs were detected for grain yield in rain-fed condition, SNPs 2711-234 (1H) and 1923-265 (1H), which explained 7 and 8.6% of phenotypic variation, respectively, but importantly, they were stables across the two contrasting water regimes. Under full irrigation, six QTLs were detected for grain yield, SNPs 2711-234 (1H), 1923-265 (1H), ConsensusGBS0598-3 (3H), 9018-522 (3H), 4105-1417 (3H), 7045-950 (3H), and explained from 5.2 to 18%. Different to grain yield, the putative QTLs underlying biological yield were detected on chromosomes 5H and 7H, in the rain-fed condition, and on chromosomes 2H, 3H, and 5H in the fully irrigated condition.

TABLE 4 | Summary of QTLs detected for 15 complex traits in a RCSL population of barley evaluated under contrasting environment conditions.

Trait	Rain-fed			Fully irrigated plants			Combined		
	NQ	ChN	PV (%)	NQ	ChN	PV (%)	NQ	ChN	PV (%)
MORPHOLOGIC									
PL	9	3H, 7H	5.3–13	11	1H, 3H, 7H	5.5–15	12	3H, 4H, 7H	5.1–18
PE	2	2H, 4H	5.3–5.5	4	6H, 7H	5.6–6.7	4	4H, 6H	5.3–6.8
SL	1	7H	6.6	2	3H, 5H	5.1–7.2	11	2H, 3H, 7H	5.1–7.5
AGRONOMIC									
PH	7	2H, 3H, 4H	5.8–22	7	2H, 4H, 6H	5.7–8.3	13	2H, 3H, 4H	5.1–13
TN	5	1H, 3H, 6H	5.3–6.4	4	4H, 5H	5.0–5.9	0		
DWT	3	3H	5.3–7.7	1	6H	6.0	1	6H	5.5
BY	5	5H, 7H	5.2–6.3	10	2H, 3H, 5H	5.4–7.1	14	2H, 3H, 5H	5.0–8.2
HW	9	1H, 2H, 4H	5.6–8.3	4	1H	6.5–35	3	1H	8.9–21
HI	1	2H	5.9	7	3H, 5H	5.2–20	4	3H	8.3–11
KS	3	3H, 5H	5.0–6.1	12	2H, 3H, 7H	6.0–11	10	2H, 3H	5.4–8.0
TKW	5	2H, 5H, 7H	5.2–6.3	0			3	2H, 5H, 7H	5.1–6.7
GY	2	1H	6.9–8.6	6	1H, 3H	5.2–18	5	1H, 3H	6.3–20
PHYSIOLOGICAL									
RWC	0			5	1H, 3H, 7H	5.1–6.1	1	1H	8.0
IPAR	3	2H, 3H	5.3–6.2	7	2H	8.0–17	6	2H	6.1–11
Fv/Fm	2	4H	5.4–5.7	4	1H, 3H	5.1–6.5	5	1H, 4H	5.8–8.8

PV (%), percentage of the phenotypic variation explained by SNP markers; NQ, number of significant QTLs; ChN, chromosome number.

All QTLs detected for BY, Fv/Fm, DWT, HI, IPAR, KS, PE, PH, SL, RWC, TKW, and TN were environment-specific. Although most QTL-trait associations were environment specific, some stable associations were also detected for HW, SNPs 2711-234 (1H) and 1923-265 (1H), and PL, SNPs 9282-205 (3H), 3965-353 (3H), 2335-1614 (3H), and 5212-1409 (7H).

Fifteen and ten SNPs, respectively, were associated with more than one trait in the fully irrigated and rain-fed conditions. Importantly, the SNPs 4105-1417, 7045-950, 9018-522, ConsensusGBS0598-3 (all on chromosome 3H) were concomitantly associated with grain yield, harvest index, and kernel per spike, in the fully irrigated condition, which accounted for 5.2–20% of the phenotypic variation. The mentioned SNPs ConsensusGBS0598-3, 7045-950, and 4105-1417 were also associated to biological yield, and the SNP 9018-522 with spike length. In both environmental conditions, grain yield and HW shared QTLs linked to the SNPs 1923-265 and 2711-234 (on chromosome 1H); both QTLs correspond to stable associations across environments.

QTLs with >15% of the phenotypic variation explained by SNP markers are shown in **Table 5**; these are relatively moderate (to major) QTLs detected in this RSCL population of barley. The results confirmed the presence of more QTLs with >15% in the fully irrigated (9) than the rain-fed condition (2), and the majority of these QTLs were environment-specific. In the fully irrigated condition, a relatively major QTL was detected underlying HW, i.e., SNP marker 1923-265 on chromosome 1H at 140 cM, which accounted for 35.3% of the total phenotypic variation. In the same genomic region, the SNP marker 2711-234 (on 1H at 139 cM) explained 27% of the phenotypic variation for HW. These QTLs were stable across the

contrasting environmental conditions. Interestingly, the major locus 1923-265 (1H) associated with HW, was also detected for grain yield in both environmental conditions, which explained between 8.6 and 18.1% of the phenotypic variation, in the rain-fed and fully irrigated conditions, respectively. Finally, one genomic region on chromosome 4H, was moderately associated with plant height (i.e., SNPs ABC08009-1-2-304 and 954-1377) and was environment-specific in the rain-fed condition.

DISCUSSION

In this study, a RCSL population consisting of 137 lines was evaluated for 15 complex traits, including plant height, grain yield, and yield-related traits in two contrasting environment conditions. Severe allele frequency distortions was evidenced along the seven linkage groups in the RCSL population. By incorporating kinship into the model, as proposed by Malosetti et al. (2011) for populations that have undergone some selection resulting in a departure from Mendelian segregation ratios, we identified QTLs associated with key traits in two contrasting field conditions.

Environment-specific genomic regions were detected for the majority of the traits (12/15). These findings are consistent with the study conducted by Wang et al. (2014), in which most of the QTL for different traits varied between environments in a doubled haploid population of barley.

Most of the economically important traits in barley are inherited quantitatively. Plant height (PH), for instance, is under polygenic control, and represents one of the most important agronomic traits for barley (Wang et al., 2014; Zhou et al., 2015). In this study, the highest number of QTL was detected for plant
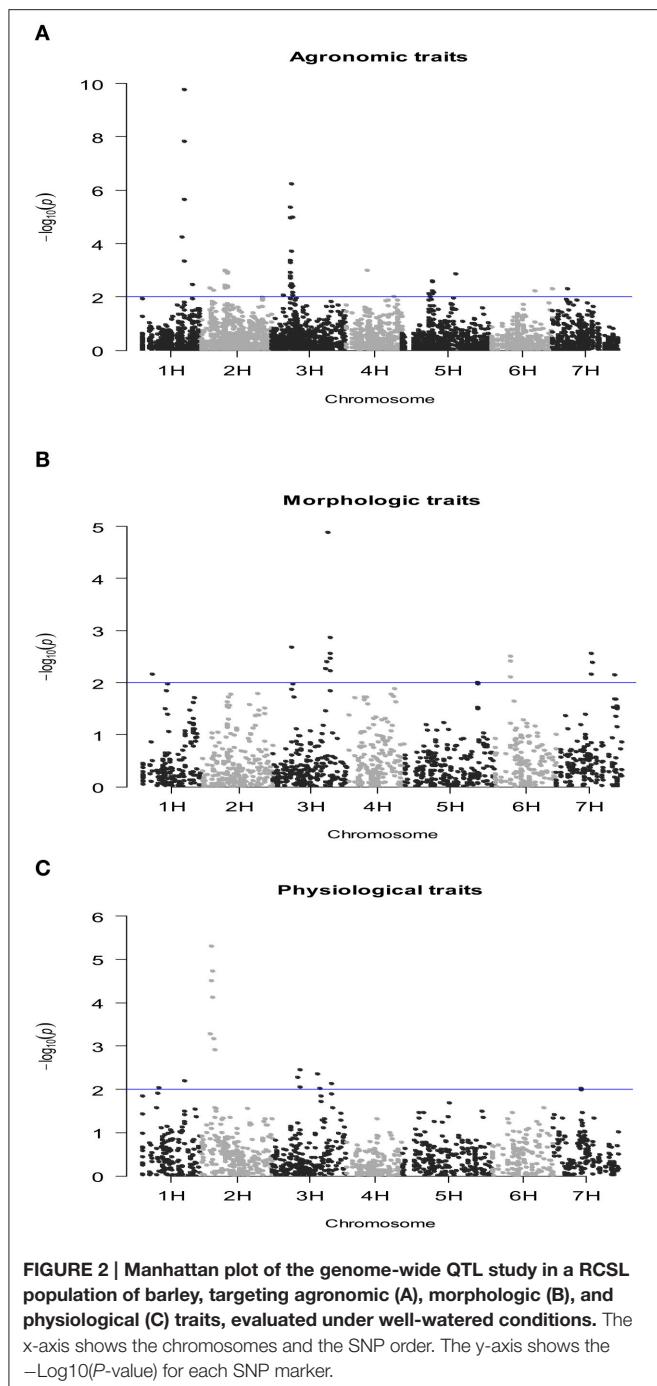
FIGURE 2 | Manhattan plot of the genome-wide QTL study in a RCSL population of barley, targeting agronomic (A), morphologic (B), and physiological (C) traits, evaluated under well-watered conditions. The x-axis shows the chromosomes and the SNP order. The y-axis shows the −Log10(*P*-value) for each SNP marker.

FIGURE 3 | Manhattan plot of the genome-wide QTL study in a RCSL population of barley, targeting agronomic (A), morphologic (B), and physiological (C) traits, evaluated under rain-fed conditions. The x-axis shows the chromosomes and the SNP order. The y-axis shows the −Log10(*P*-value) for each SNP marker.

height (20 QTL in both conditions) which were observed on chromosomes 2H, 3H, 4H, and 6H. Similarly, Honsdorf et al. (2014) found the highest number of associations for this trait, which were located on all chromosomes except 5H. Inostroza et al. (2009) found SSR-trait associations for PH on chromosomes 1H, 2H, 4H, 5H, 6H, and 7H, evidencing a genome-wide distribution. Interestingly, in the present study, a significant genomic region (explained 22% of phenotypic variance) on chromosome 4H, which comprises the SNPs ABC08009-1-2-304

and 954-1377, controls PH in barley under rain-fed conditions. The majority of the QTLs were detected on chromosomes 2H (4/7) under fully irrigated and 4H (5/7) under rain-fed condition. This result is partially consistent with Pasam et al. (2012), who found 32 associations with plant height with the majority located on chromosomes 2H and 3H. Malosetti et al. (2011) found significant SNP associations with PH on chromosomes 2H, 3H, 5H, and 7H, with two important QTLs on chromosomes 3H (127.1 cM) and 5H (69.3 cM), which are known to carry semi-dwarfing genes in barley (Malosetti et al., 2011; Wang et al., 2014). The use of semi-dwarf genes has greatly improved barley

Trait	SNP	ChN	Pos	PV(%)	Situation
FULLY IRRIGATED PLANTS					
PL	3965-353	3H	176	15.1	Both environments
IPAR	8787-1459	2H	39	16.8	Environment-specific
HI	ConsensusGBS0598-3	3H	60	15.6	Environment-specific
HI	9018-522	3H	61	16.9	Environment-specific
HI	4105-1417	3H	64	20.3	Environment-specific
HI	7045-950	3H	67	15.4	Environment-specific
GY	1923-265	1H	140	18.1	Both environments
HW	2711-234	1H	139	26.9	Both environments
HW	1923-265	1H	140	35.3	Both environments
RAIN-FED					
PH	ABC08009-1-2-304	4H	180	22.2	Environment-specific
PH	954-1377	4H	182	22.2	Environment-specific
COMBINED					
PL	3965-353	3H	176	18.3	Both environments
GY	1923-265	1H	140	20.4	Both environments
HW	2711-234	1H	139	17.1	Both environments
HW	1923-265	1H	140	21.4	Both environments

PV(%), percentage of the phenotypic variation explained by SNP markers; ChN, chromosome number; Pos, SNP position in cM.

yields with controlled plant height being used to reduce yield loss arising from lodging and to increase the harvest index (Wang et al., 2014). Zhou et al. (2015) found a major QTL for plant height mapped at 105.5 cM on chromosome 3H, which had a LOD score of 13.01 and explained 44.5% of phenotypic variation. In this study, one QTL was detected at 197 cM on chromosome 3H (SNP 9610-1195), in the rain-fed condition, which explained 7.3% of phenotypic variation.

In the fully irrigated condition, a relatively major QTL was found for HW on chromosome 1H at 140 cM (SNP 1923-265), which explained 35.3% of phenotypic variation. This result is consistent with the study of Rode et al. (2012) who found three QTLs for HW on chromosome 1H, including the SNP 1923-265. In rain-fed condition, this SNP explained 8.6% of phenotypic variation for grain yield, and under fully irrigated condition, it explained 5.7 and 18.1% for RWC and GY, respectively. In contrast, Rode et al. (2012) did not find any SNP controlling HW associated with another related-trait. As expected, the SNP marker 2711-234 (in the same genomic region of SNP 1923-265 on chromosome 1H, at 139 cM) that explained 27% of the phenotypic variation for HW, was associated with grain yield in both environmental conditions, explaining 10 and 7% of phenotypic variation. The correlation coefficient between both traits was positively correlated ($r = 0.48$; $P < 0.01$; **Table 3**).

According to Naz et al. (2014), tiller number per plant is a major determinant of yield in crops like barley. In this study, the nine QTL detected for TN were located on all chromosomes except 2H and 7H, and all were environment-specific. There have

also been conflicting reports on the QTL detection for TN and their chromosomal location in barley, although chromosomes 3H and 4H appear to be consistent (Elberse et al., 2004; Wang et al., 2010; Honsdorf et al., 2014). Naz et al. (2014) identified five QTL for TN on chromosomes 1H, 2H, 4H, and 5H; of which one QTL (located on 5H between 203.85 and 231.75 cM) accounted for 70.5% increase in TN. On the other hand, Honsdorf et al. (2014), for instance, found two QTL for this trait on chromosomes 3H and 4H.

Thousand kernel weight (TKW) is one of the major yield components having direct effect on the final yield (Pasam et al., 2012). In this study, five significant QTLs associated with TKW were found on chromosomes 2H (three QTLs), 5H (one QTL), and 7H (one QTL), under rain-fed condition, which explained between 5.2 and 6.3% of the phenotypic variation (**Table 3**). In contrast, Pasam et al. (2012) found 21 QTL associated with thousand grain weight, which were present on all chromosomes. Comadran et al. (2011) detected three QTL associated with TKW on chromosome 2H, and, similarly with our study, none of these associations accounted for more than 10% of the phenotypic variation; the largest effect was over 5% of the trait mean. Kalladan et al. (2013) found consistent QTL for TKW across the environments (stable QTL), which were mapped to all seven linkage groups except chromosomes 4H and 5H.

Two genomic region on chromosomes 1H and 3H were associated with grain yield under favorable conditions. These correspond to the SNP markers 2711-234 and 1923-265, on chromosome 1H, and ConsensusGBS0598-3 (60 cM), 9018-522 (61 cM), 4105-1417 (64 cM), and 7045-950 (67 cM) on chromosome 3H. Similarly, Rode et al. (2012) found two QTLs associated with grain yield on chromosome 3H, but at 42.1 cM (SNP 15141-288) and 169.3 cM (SNP ConsensusGBS0632-3). The two QTLs detected for grain yield in rain-fed conditions, SNPs 2711-234 (1H) and 1923-265 (1H), were stables across the two contrasting water regimes. This result is in accordance with the findings of Kalladan et al. (2013) who found altogether five stable QTL for yield; of them three were mapped to chromosome 1H: Moreover, the QTL explaining most of the phenotypic variations for yield were found on chromosome 1H and 2H. In contrast, seven QTLs were found in the Rode's study on chromosome 5H. Comadran et al. (2011) found three main QTL for grain yield located on chromosomes 2H and 7H. Inostroza et al. (2009) although using simple sequence repeats (SSRs) found that the yield QTLs are distributed throughout the genome, on chromosomes 1H, 2H, 3H, 5H, 6H, and 7H; similar to the findings of Mansour et al. (2014).

Importantly, Comadran et al. (2011) mentioned that co-localization of several QTL related to yield components traits suggest that major developmental loci may be linked to most of the associations in barley. In our study, some markers on chromosome 3H were concomitantly associated with biological and grain yield, harvest index, spike length, and kernel per spike, in the irrigated environment. Most of the correlation coefficients among these traits were positive and statistically different from zero ($P < 0.01$) varying from $r = 0.24$ to 0.64. In both environmental conditions, grain yield and HW (correlation coefficient $r = 0.48$, $P < 0.01$) shared

two QTLs located on chromosome 1H (SNPs 1923-265 and 2711-234), both QTLs correspond to stable associations across environments.

In this study, eight putative QTLs underlying harvest index were identified on chromosomes 2H, 3H, and 5H. Most of them were localized on chromosome 3H (6/8) under irrigation conditions. Only one significant QTL (SNP 5880-2547) controlling HI was found on 2H under rain-fed conditions, which explained 6% of phenotypic variation. This QTL controlling HI was also concomitantly associated with thousand kernel weight, in rain-fed (explained 6%); the correlation coefficient between these traits was statistically different from zero $r = 0.2$, $P < 0.05$. Comadran et al. (2011) detected three QTL for HI on chromosomes 1H, 2H, and 3H, with the most significant being located on chromosome 2H, in the same region as QTL for heading date and yield.

Fan et al. (2015) found that the physiological trait RWC had a very close correlation ($r = 0.73$, $P < 0.01$) with drought tolerance in barley; in fact, one QTL for RWC was identified on chromosome 2H and it explained 44.3% of phenotypic variation. In the current study, no QTL controlling RWC was identified in rain-fed, and only a small effect was identified on 1H, 3H, and 7H accounting for around 5% of the phenotypic variation. This was expected considering that there was no significant difference in RWC between rain-fed and irrigated environments. Li et al. (2013) carried out a meta-analysis of QTL associated with tolerance to abiotic stresses in barley, identifying MetaQTL for RCW under abiotic stress on H1, H2, H5, H5, and H7. Wójcik-Jagła et al. (2013) in a comparative QTL analysis of early short-time drought tolerance in Polish fodder and malting spring barleys, found 18 QTLs for nine physiological traits on all chromosomes except 1H in malting barley and 15 QTLs for five physiological traits on chromosomes 2H, 4H, 5H, and 6H in fodder barley.

Fluorescence parameters for dark adapted flag leaves (Fo, Fm, Fv, Fv/Fm) of 194 recombinant inbred lines (RILs), developed from the cross between the cultivar "Arta" and *H spontaneum* 41-1, measured under well-watered and drought stress conditions, showed significant differences among RILs but no differences between water regimes (Guo et al., 2008). Using SSRs and AFLPs markers they were able to identified nine and five QTLs, under well-watered and drought stress conditions, respectively; a QTL for Fv/Fm [i.e., (Fm – Fo)/Fm], which explained 15% of the phenotypic variance, was identified on chromosome 2H at 116 cM in the linkage map under drought stress. In our study two QTLs for quantum yield of PSII on chromosome 4H were identify under rain-fed condition and five in the combined analysis (**Table 4**). In the study of Wójcik-Jagła et al. (2013) one major QTL related to photochemical quenching of chlorophyll fluorescence was located on chromosome 4H in fodder barley. In an advanced backcross quantitative trait locus (AB-QTL) analysis performed by Sayed et al. (2012) to elucidate genetic mechanisms controlling proline content (PC) and leaf wilting (WS) in barley under drought stress conditions, QTL for WS were localized on

chromosome 1H, 2H, 3H, and 4H. Among these, QWS.S42.1H and QWS.S42.4H were associated to decrease in WS due to the introgression of exotic alleles. QTL for PC were localized on chromosome 3H, 4H, 5H, and 6H. QTL effects on 3H, 4H, and 6H were responsible to heighten PC due to the preeminence of elite alleles over the exotic alleles which ranged from 26 to 43%.

The results are in agreement with the findings of Malosetti et al. (2011) who screened a population of 161 inbred lines of barley with 1536 SNPs, which were used for gene and QTL detection. The model incorporating kinship, co-ancestry information, was consistently superior to the one without kinship (according to the Akaike information criterion), similarly with this study. Importantly, Malosetti et al. (2011) showed that ignoring this type of information results in an unrealistically high number of marker–trait associations, without providing clear conclusions about QTL locations. In this work, a large number of spurious QTLs were detected when the genetic covariance matrix is ignored in the mixed model (444 and 460 false-positives for all traits studied in the rain-fed and full irrigation conditions, respectively) confirming that ignoring this type of genetic relatedness will increase the rate of false-positives. As with the study carried out by Malosetti et al. (2011) we highlight the importance of the inclusion of kinship information when detecting QTL in populations that have undergone some process of selection. This research provides useful information for MAS programs in areas where drought is a significant constraint. As drought stress tolerance has become an important goal, this analysis of QTL identified significant genomic regions that can be used for breeding purposes.

AUTHOR CONTRIBUTIONS

FM and YQ made the QTLs analysis and FM wrote the first manuscript; IM was responsible of the field trial and agronomic trait determination; JR and RW were responsible of the research grants CGIAR-GCP Challenge and performed the SNPs analysis; AD performed the physiological evaluations and is the leader of the drought tolerance studies of barley in Chile. All the authors contributed to the final manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00909

REFERENCES

Alheit, K. V., Reif, J. C., Maurer, H. P., Hahn, V., Weissmann, E. A., Miedaner, T., et al. (2011). Detection of segregation distortion loci in Triticale (x Triticosecale Wittmack) based on a high-density DArT marker consensus genetic linkage map. *BMC Genomics* 12:380. doi: 10.1186/1471-2164-12-380

Araus, J. L., Slafer, G. A., Reynolds, M., and Royo, C. (2002). Plant breeding and drought in C3 cereals: what should we breed for? *Ann. Bot.* 89, 925–940. doi: 10.1093/aob/mcf049

Araus, J. L., Slafer, G. A., Royo, C., and Serret, M. D. (2008). Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant. Sci.* 27, 377–412. doi: 10.1080/07352680802467736

Boer, M. P., Wright, D., Feng, L., Podlich, D. W., Luo, L., Cooper, M., et al. (2007). A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. *Genetics* 177, 1801–1813. doi: 10.1534/genetics.107.071068

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Cattivelli, L., Rizza, F., Badeck, F. W., Mazzucotelli, E., Mastrangelo, A. M., Francia, E., et al. (2008). Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. *Field. Crop. Res.* 105, 1–14. doi: 10.1016/j.fcr.2007.07.004

Chen, G., Krugman, T., Fahima, T., Chen, K., Hu, Y., Roder, M., et al. (2010). Chromosomal regions controlling seedling drought resistance in Israeli wild barley, *Hordeum spontaneum* C. *Koch. Genet. Resour. Crop. Evol.* 57, 85–99. doi: 10.1007/s10722-009-9453-z

Close, T. J., Bhat, P. R., Lonardi, S., Wu, Y., Rostoks, N., Ramsay, L., et al. (2009). Development and implementation of high-throughput SNP genotyping in barley. *BMC Genomics* 10:582. doi: 10.1186/1471-2164-10-582

Comadran, J., Russell, J. R., Booth, A., Pswarayi, A., Ceccarelli, S., Grando, S., et al. (2011). Mixed model association scans of multi-environmental trial data reveal major loci controlling yield and yield related traits in *Hordeum vulgare* in Mediterranean environments. *Theor. Appl. Genet.* 122, 1363–1373. doi: 10.1007/s00122-011-1537-4

del Pozo, A., Castillo, D., Inostroza, L., Matus, I., Méndez, A. M., and Morcuende, R. (2012). Physiological and yield responses of recombinant chromosome substitution lines of barley to terminal drought in a Mediterranean-type environment. *Ann. Appl. Biol.* 160, 157–167. doi: 10.1111/j.1744-7348.2011.00528.x

Elberse, I. A. M., Vanhala, T. K., Turin, J. H. B., Stam, P., van Damme, J. M. M., and van Tienderen, P. H. (2004). Quantitative trait loci affecting growth-related traits in wild barley (*Hordeum spontaneum*) grown under different levels of nutrient supply. *Heredity* 93, 22–33. doi: 10.1038/sj.hdy.6800467

Fan, Y., Shabala, S., Ma, Y., Xu, R., and Zhou, M. (2015). Using QTL mapping to investigate the relationships between abiotic stress tolerance (drought and salinity) and agronomic and physiological traits. *BMC Genomics* 16:43. doi: 10.1186/s12864-015-1243-8

Griffin, T. J., and Smith, L. M. (2000). Single-nucleotide polymorphism analysis by MALDI–TOF mass spectrometry. *Trends Biotechnol.* 18, 77–84. doi: 10.1016/S0167-7799(99)01401-8

Guo, P., Baum, M., Varshney, R. K., Graner, A., Grando, S., and Ceccarelli, S. (2008). QTLs for chlorophyll and chlorophyll fluorescence parameters in barley under post-flowering drought. *Euphytica* 163, 203–214. doi: 10.1007/s10681-007-9629-6

Gyenis, L., Yun, S. J., Smith, K. P., Steffenson, B. J., Bossolini, E., Sanguineti, M. C., et al. (2007). Genetic architecture of quantitative trait loci associated with morphological and agronomic trait differences in a wild by cultivated barley cross. *Genome* 50, 714–723. doi: 10.1139/G07-054

Hashemi, F. S. G., Rafii, M. Y., Ismail, M. R., Mohamed, M. T. M., Rahim, H. A., Latif, M. A., et al. (2015). Comparative mapping and discovery of segregation distortion and linkage disequilibrium across the known fragrance chromosomal regions in a rice F2 population. *Euphytica* 204, 557–569. doi: 10.1007/s10681-014-1322-y

Honsdorf, N., March, T. J., Berger, B., Tester, M., and Pillen, K. (2014). High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. *PLoS ONE* 9:e97047. doi: 10.1371/journal.pone.0097047

Inostroza, L., Del Pozo, A., Matus, I., Castillo, D., Hayes, P., Machado, S., et al. (2009). Association mapping of plant height, yield, and yield stability in recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp. *spontaneum* as a source of donor alleles in a *Hordeum vulgare* subsp. *vulgare* background. *Mol. Breeding* 23, 365–376. doi: 10.1007/s11032-008-9239-6

Jacobsen, S. E., Jensen, C. R., and Liu, F., (2012). Improving crop production in the arid Mediterranean climate. *Field Crop Res.* 28, 34–47. doi: 10.1016/j.fcr.2011.12.001

Kalladan, R., Worch, S., Rolletschek, H., Harshavardhan, V. T., Kuntze, L., Seiler, C., et al. (2013). Identification of quantitative trait loci contributing to yield and seed quality parameters under terminal drought in barley advanced backcross lines. *Mol. Breeding* 32, 71–90. doi: 10.1007/s11032-013-9853-9

Kota, R., Rudd, S., Facius, A., Kolesov, G., Thiel, T., Zhang, H., et al. (2003). Snipping polymorphisms from large EST collections in barley (*Hordeum vulgare* L.). *Mol. Genet. Genomics.* 270, 24–33. doi: 10.1007/s00438-003-0891-6

Li, J., Huang, X. Q., Heinrichs, F., Ganal, M. W., and Röder, M. S. (2005). Analysis of QTLs for yield, yield components, and malting quality in a BC3-DH population of spring barley. *Theor. Appl. Genet.* 110, 356–363. doi: 10.1007/s00122-004-1847-x

Li, W. T., Liu, C. J., Liu, Y. X., Pu, Z. E., Dai, S. F., Wang, J. R., et al. (2013). Meta-analysis of QTL associated with tolerance to abiotic stresses in barley. *Euphytica* 189, 31–49. doi: 10.1007/s10681-012-0683-3

Lobell, D. B., and Field, C. B. (2007). Global scale climate–crop yield relationships and the impacts of recent warming. *Environ. Res. Lett.* 2:014002. doi: 10.1088/1748-9326/2/1/014002

Lorieux, M., Perrier, G. X., Gonzalez de Leon, D., and Lanaud, C. (1995a). Maximum likelihood models for mapping genetic markers showing segregation distortion. 1. Backcross population. *Theor. Appl. Genet.* 90, 73–80. doi: 10.1007/BF00220998

Lorieux, M., Perrier, X., Goffinet, B., Lanaud, C., and Gonzalez de Leon, D. (1995b). Maximum likelihood models for mapping genetic markers showing segregation distortion. 2. F2 population. *Theor. Appl. Genet.* 90, 81–89. doi: 10.1007/BF00220999

Lu, H., Romero-Severson, J., and Bernardo, R. (2002). Chromosomal regions associated with segregation distortion in maize. *Theor. Appl. Genet.* 105, 622–628. doi: 10.1007/s00122-002-0970-9

Luo, L., Zhang, Y. M., and Xu, S. (2005). A quantitative genetics model for viability selection. *Heredity* 94, 347–355. doi: 10.1038/sj.hdy.6800615

Malosetti, M., Ribaut, J., Vargas, M., Crossa, J., and Van Euuwijk, F. (2008). A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). *Euphytica* 161, 241–257. doi: 10.1007/s10681-007-9594-0

Malosetti, M., Van Eeuwijk, F. A., Boer, M. P., Casas, A. M., Elía, M., Moralejo, M., et al. (2011). Gene and QTL detection in a three-way barley cross under selection by a mixed model with kinship information using SNPs. *Theor. Appl. Genet.* 122, 1605–1616. doi: 10.1007/s00122-011-1558-z

Malosetti, M., Voltas, J., Romagosa, I., Ullrich, S. E., and Van Eeuwijk, F. A. (2004). Mixed models including environmental covariables for studying QTL by environment interaction. *Euphytica* 137, 139–145. doi: 10.1023/B:EUPH.0000040511.46388.ef

Mansour, E., Casas, A. M., Gracia, M. P., Molina-Cano, J. L., Moralejo, M., Cattivelli, L., et al. (2014). Quantitative trait loci for agronomic traits in an elite barley population for Mediterranean conditions. *Mol. Breeding* 33, 249–265. doi: 10.1007/s11032-013-9946-5

Matus, I., Corey, A., Filichkin, T., Hayes, P. M., Vales, M. I., Kling, J., et al. (2003). Development and characterization of recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp spontaneum as a source of donor alleles in a *Hordeum vulgare* subsp. *vulgare* background. *Genome* 46, 1010–1023. doi: 10.1139/g03-080

Mir, R. R., Zaman-Allah, M., Sreenivasulu, N., Trethowan, R., and Varshney, R. K. (2012). Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. *Theor. Appl. Genet.* 125, 625–645. doi: 10.1007/s00122-012-1904-9

Mora, F., Castillo, D., Lado, B., Matus, I., Poland, J., Belzile, F., et al. (2015). Genome-wide association mapping of agronomic traits and carbon

isotope discrimination in a worldwide germplasm collection of spring wheat using SNP markers. *Mol. Breeding* 35, 1–12. doi: 10.1007/s11032-015-0264-y

Naz, A. A., Arifuzzaman, M., Muzammil, S., Pillen, K., and Léon, J. (2014). Wild barley introgression lines revealed novel QTL alleles for root and related shoot traits in the cultivated barley (*Hordeum vulgare* L.). *BMC Genetics* 15:107. doi: 10.1186/s12863-014-0107-6

Naz, A. A., Ehl, A., Pillen, K., and León, J. (2012). Validation for root-related quantitative trait locus effects of wild origin in the cultivated background of barley (*Hordeum vulgare* L.). *Plant Breeding* 131, 392–398. doi: 10.1111/j.1439-0523.2012.01972.x

Pasam, R. K., Sharma, R., Malosetti, M., Van Eeuwijk, F. A., Haseneyer, G., Kilian, B., et al. (2012). Genome-wide association studies for agronomical traits in a world wide spring barley collection. *BMC Plant Biol.* 12:16. doi: 10.1186/1471-2229-12-16

Paterson, A., Lander, E., Hewitt, J. D., Peterson, S., Lincoln, S. E., and Tanksley, S. D. (1988). Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335, 721–726. doi: 10.1038/335721a0

Pillen, K., Zacharias, A., and Léon, J. (2004). Comparative AB-QTL analysis in barley using a single exotic donor of *Hordeum vulgare* ssp. *spontaneum*. *Theor. Appl. Genet.* 108, 1591–1601. doi: 10.1007/s00122-004-1586-z

Rode, J., Ahlemeyer, J., Friedt, W., and Ordon, F. (2012). Identification of marker-trait associations in the German winter barley breeding gene pool (*Hordeum vulgare* L.). *Mol. Breeding* 30, 831–843. doi: 10.1007/s11032-011-9667-6

Sato, K., and Takeda, K. (2009). An application of high-throughput SNP genotyping for barley genome mapping and characterization of recombinant chromosome substitution lines. *Theor. Appl. Genet.* 119, 613–619. doi: 10.1007/s00122-009-1071-9

Sayed, M. A., Schumann, H., Pillen, K., Naz, A. A., and Léon, J. (2012). AB-QTL analysis reveals new alleles associated to proline accumulation and leaf wilting under drought stress conditions in barley (*Hordeum vulgare* L.). *BMC Genet.* 13:61. doi: 10.1186/1471-2156-13-61

Schmalenbach, I., Körber, N., and Pillen, K. (2008). Selecting a set of wild barley introgression lines and verification of QTL effects for resistance to powdery mildew and leaf rust. *Theor. Appl. Genet.* 117, 1093–1106. doi: 10.1007/C

Schmalenbach, I., March, T. J., Bringezu, T., Waugh, R., and Pillen, K. (2011). High-resolution genotyping of wild barley introgression lines and fine-mapping of the threshability locus thresh-1 using the Illumina GoldenGate assay. *G3 GenesGenomes Genet.* 1, 187–196. doi: 10.1534/g3.111.000182

Stich, B., Möhring, J., Piepho, H. P., Heckenberger, M., Buckler, E. S., and Melchinger, A. E. (2008). Comparison of mixed-model approaches for association mapping. *Genetics* 178, 1745–1754. doi: 10.1534/genetics.107.079707

Szűcs, P., Blake, V. C., Bhat, P. R., Chao, S., Close, T. J., Cuesta-Marcos, A., et al. (2009). An integrated resource for barley linkage map and malting quality QTL alignment. *Plant. Gen.* 2, 134–140. doi: 10.3835/plantgenome2008.01.0005

Talame, V., Sanguineti, M. C., Chiapparino, E., Bahri, H., Ben Salem, M., Forster, B. P., et al. (2004). Identification of *Hordeum spontaneum* QTL alleles improving field performance of barley grown under rainfed conditions. *Ann. Appl. Bot.* 144, 309–319. doi: 10.1111/j.1744-7348.2004.tb00346.x

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Tanksley, S. D., and Nelson, J. C. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92, 191–203. doi: 10.1007/BF00223376

Tondelli, A., Francia, E., Barabaschi, D., Aprile, A., Skinner, J. S., Stockinger, E. J., et al. (2006). Mapping regulatory genes as candidates for cold and drought stress tolerance in barley. *Theor. Appl. Genet.* 112, 445–454. doi: 10.1007/s00122-005-0144-7

Van Berloo, R., Van Heusden, S., Bovy, A., Meijer-Dekens, F., Lindhout, P., and Van Eeuwijk, F. (2008). Genetic research in a public–private research consortium: prospects for indirect use of Elite breeding germplasm in academic research. *Euphytica* 161, 293–300. doi: 10.1007/s10681-007-9519-y

Wang, G., Schmalenbach, I., von Korff, M., León, J., Kilian, B., Rode, J., et al. (2010). Association of barley photoperiod and vernalization genes with QTLs for flowering time and agronomic traits in a BC(2)DH population and a set of wild barley introgression lines. *Theor. Appl. Genet.* 120, 1559–1574. doi: 10.1007/s00122-010-1276-y

Wang, J., Yang, J., Jia, Q., Zhu, J., Shang, Y., Hua, W., et al. (2014). A new QTL for plant height in barley (*Hordeum vulgare* L.) showing no negative effects on grain yield. *PLoS ONE* 9:e90144. doi: 10.1371/journal.pone.0090144

Wehner, G. G., Balko, C. C., Enders, M. M., Humbeck, K. K., and Ordon, F. F. (2015). Identification of genomic regions involved in tolerance to drought stress and drought stress induced leaf senescence in juvenile barley. *BMC Plant Biology* 15:125. doi: 10.1186/s12870-015-0524-3

Wójcik-Jagła, M., Rapacz, M., Tyrka, M., Kościelniak, J., Crissy, K., and Żmuda, K. (2013). Comparative QTL analysis of early short-time drought tolerance in Polish fodder and malting spring barleys. *Theor. Appl. Genet.* 126, 3021–3034. doi: 10.1007/s00122-013-2190-x

Xu, S. (2008). Quantitative trait locus mapping can benefit from segregation distortion. *Genetics* 180, 2201–2208. doi: 10.1534/genetics.108.090688

Xu, Y., Zhu, L., Xiao, J., Huang, N., and McCouch, S. R. (1997). Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled-haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* 253, 535–545. doi: 10.1007/s004380050355

Zhang, L., Wang, S., Li, H., Deng, Q., Zheng, A., Li, S., et al. (2010). Effects of missing marker and segregation distortion on QTL mapping in F2 populations. *Theor. Appl. Genet.* 121, 1071–1082. doi: 10.1007/s00122-010-1372-z

Zhou, G., Zhang, Q., Zhang, X. Q., Tan, C., and Li, C. (2015). Construction of high-density genetic map in barley through restriction-site associated DNA sequencing. *PLoS ONE* 10:e0133161. doi: 10.1371/journal.pone.0133161

Zhu, C., Wang, C., and Zhang, Y. M. (2007). Modeling segregation distortion for viability selection I. Reconstruction of linkage maps with distorted markers. *Theor. Appl. Genet.* 114, 295–305. doi: 10.1007/s00122-006-0432-x

Glutamine synthetase in Durum Wheat: Genotypic Variation and Relationship with Grain Protein Content

Domenica Nigro[1], Stefania Fortunato[1,2], Stefania L. Giove[1], Annalisa Paradiso[3], Yong Q. Gu[4], Antonio Blanco[1], Maria C. de Pinto[3] and Agata Gadaleta[2]*

[1] Department of Soil, Plant and Food Sciences, University of Bari Aldo Moro, Bari, Italy, [2] Department of Agricultural and Environmental Sciences, Research Unity of Genetic and Plant Biotechnology, University of Bari Aldo Moro, Bari, Italy, [3] Department of Biology, University of Bari Aldo Moro, Bari, Italy, [4] Crop Improvement and Genetics Research, Western Regional Research Center, United States Department of Agriculture – Agricultural Research Service, Albany, CA, USA

Grain protein content (GPC), is one of the most important trait in wheat and its characterized by a very complex genetic control. The identification of wheat varieties with high GPC (HGPC), as well as the characterization of central enzymes involved in these processes, are important for more sustainable agricultural practices. In this study, we focused on *Glutamine synthetase* (*GS*) as a candidate to study GPC in wheat. We analyzed *GS* expression and its enzymatic activity in different tissues and phenological stages in 10 durum wheat genotypes with different GPC. Although each genotype performed quite differently from the others, both because their genetic variability and their adaptability to specific environmental conditions, the highest GS activity and expression were found in genotypes with HGPC and *vice versa* the lowest ones in genotypes with low GPC (LGPC). Moreover, in genotypes contrasting in GPC bred at different nitrogen regimes (0, 60, 140 N Unit/ha) GS behaved differently in diverse organs. Nitrogen supplement increased *GS* expression and activity in roots of all genotypes, highlighting the key role of this enzyme in nitrogen assimilation and ammonium detoxification in roots. Otherwise, nitrogen treatments decreased GS expression and activity in the leaves of HGPC genotypes and did not affect GS in the leaves of LGPC genotypes. Finally, no changes in *GS* and soluble protein content occurred at the filling stage in the caryopses of all analyzed genotypes.

Keywords: wheat, grain protein content, GS (Glutamine synthetase), qRT-PCR, enzyme activity, western blot

INTRODUCTION

Global agriculture urgently requires a modification of standard breeding practices and management policies. A recent report by the United Nation (The Millennium Development Goals Report, 2014) highlighted that the world's population reached 7.2 billion in 2014 and is expected to increase by more than 2 billion by 2050. This means that in the very near future even higher production will be needed to maintain food supplies. Indeed, breeders and scientists have focused their efforts on the identification of agricultural practices and the development of new genetic technologies.

Agricultural productivity had increased in recent decades through the diffusion of modern crop production practices, such as the spread of high-yielding crop varieties and a heavier use of mineral fertilizers. Nitrogen is the most important nutrient and, secondly only to water, a limiting factor for plant growth and development (Kraiser et al., 2011). In the last 40 years, the amount of nitrogen fertilizers supplied to crops has risen dramatically from 12 to 104 Tg/year (Mulvaney et al., 2009). This excess in synthetic N supply significantly affected yield increase. However, as reported in the statistics from the Food and Agricultural Organization of the United Nations, the yield of crops, especially wheat, soybean and maize, have slowed to a growth rate of about 1% annually, and in some specific cases, as in developed countries, the growth rate is quite close to zero (Fischer et al., 2009). Much of this nitrogen is wasted, as well – of the total amount of N supplied, only 30–50% is actually taken up by the plant (depending on the species and cultivar) and used in different biochemical pathways. Most is lost to the environment in several ways, such as surface run-off, leaching of nitrates, ammonia (NH_3) volatilization or bacterial competition (Garnett et al., 2009). This represents a considerable expense both in terms of cost and environmental impact. Control directives and best management practices have been implemented several years ago to minimize environmental damage from nitrogen run-off (The Nitrates Directive, EC91/676/EEC, The EU Water Framework Directive, 2000/60/EC). Several studies and international projects have since highlighted the importance of defining the optimum timing and rate of nitrogen application during plant growth to maximize yield.

One of the most valuable agronomic and physiological indicators of how plants respond and use available N is nitrogen use efficiency (NUE), at first defined as the yield of grain per unit of available nitrogen in the soil (Moll et al., 1982, 1987). Currently, NUE could be defined as the ratio among plant grain yield and plant-available N in the soil, including soil-native N and N applied as fertilizer, and is composed of N-uptake efficiency and physiological N-use efficiency (De Macale and Velk, 2004). There is a need to diversify NUE significance, as there are several interpretations of this agronomic trait, depending on species and parameters of interest to be evaluated (Pathak et al., 2011; Hawkesford et al., 2013). Barraclough et al. (2014) studied how to quantify genetic variation in the uptake, portioning and remobilization of nitrogen in individual plant organs at extreme rates on N supply and can influence grain protein content (GPC). They found out that biggest contributor to variation in plant and crop performance was N-rate, followed by growth stage and finally genotype.

Glutamine synthetase (GS), an enzyme with an essential role in the assimilation of inorganic N, has been proposed as a candidate for improving NUE in wheat (Habash et al., 2007; Gadaleta et al., 2011, 2014; Thomsen et al., 2014). GS is present in most species, with three to five isoforms localized in the cytosol (GS1) and a single isoform (GS2) in plastids (Swarbreck et al., 2011). On the bases of phylogenetic studies and mapping data in wheat, 10 GS cDNA sequences were classified into four sub-families denominate *GS1* (a, b, and c), *GS2* (a, b, and c), *GSr* (1 and 2), and *GSe* (1 and 2; Bernard et al., 2008; Thomsen et al., 2014).

Bernard et al. (2008) reported that QTLs for flag leaf and total GS activity were positively co-localized with QTLs for grain and stem nitrogen amount, but smaller correlations were established with loci for grain yield components; they identified QTLs for GS activity co-localized to a *GS2* gene mapped on chromosome 2A and to the *GSr* gene on 4A. Genetic studies in rice (Obara et al., 2004) and maize (Gallais and Hirel, 2004) demonstrated co-localizations of QTLs for GS protein or activity with QTLs relating to grain parameters at the mapped GS genes.

To date few studies are available on the role of genotypic variation of GS for GPC. In this work, we present data on total GS activity and expression in 10 wheat genotypes in relation to their final GPC. Moreover, the response to nitrogen supplies in terms of total *GS* expression and activity of four different wheat genotypes, differing in GPC, has been investigated.

MATERIALS AND METHODS

Plant Material and Field Experiment Design

Ten different durum wheat genotypes (the breeding lines PI191145 and PC32, and genotypes Svevo, Cannizzo, Gianni, Ciccio, Appio, Lucanica, Canyon, and Vesuvio) were chosen from a collection of tetraploid wheat genotypes described by Laidò et al. (2013) and Marcotuli et al. (2015). Wheat genotypes were grown for 6 years (2009–2013) without any external nitrogen supply at Valenzano (Bari, Italy); geographical coordinates: 41° 2′ 0″ North, 16° 53′ 0″ East. A randomized complete block design with three replications and plots consisting in 2.0 m × 1.5 m, with a seed density of 350 germinated seeds/m^2. According to the standard agronomic practices in the study's area, fertilizer applications were made at pre-sowing (90 kg/ha P_2O_5). During the growing season standard cultivation practices were adopted without water supply. The plants were harvested after physiological maturity, on July 10, of each year.

The different selected genotypes were chosen according to previous evaluation of yield and quality component trait (unpublished data). In order to evaluated the involvement of candidate enzymes and genes only in the accumulation of GPC, genotypes with similar value of grain yield per spike (GYS) and thousand kernel weight (TKW) were chosen, in order to avoid the negative correlation between GPC and GYS or dilution factor due to TKW (Supplemetary Table S1).

In 2014, four genotypes (PC32, Cannizzo, Ciccio, and Vesuvio) were grown in Valenzano (BA) at three different nitrogen regimes: 0, 60, and 140 N Unit/ha in randomized blocks with replicates (indicated as N0, N60, and N140). Each genotype was sown in one linear meter row and 20 cm apart. Nitrogen was supplied, in the form of ammonia nitrate, in three equal rates, 10 days before collecting samples at stages of first leaf, flowering, and grain filling. Roots were collected from plants at the seedling stage immediately washed, removed excess of water, frozen in liquid nitrogen, and stored at −80°C. Leaf tissues of each sample were collected in each phase 10 days after nitrogen implementation, immediately frozen in liquid nitrogen, and stored at −80°C until used in further assays.

Protein Content Quantification

Total GPC was assessed on 3 g of whole meal flour using a dual beam near infrared reflectance spectrophotometer (Zeutec Spectra Alyzer Premium, Zeutec Büchi, Rendsburg, Germany). Soluble proteins were assayed according to Bradford (1976), using bovine albumin as a standard.

GS Activity Determination

Plant tissues were frozen in liquid N and ground in a mortar with 1:10 (w/v) extraction buffer (100 mM triethanolamine, 1 mM EDTA, 10 mM $MgSO_4$, 5 mM glutamate, 10% v/v ethylene glycol, 10 μM leupeptin and 6 mM DTT- pH7.6). Crude extracts were centrifuged at $21000 \times g$ for 30 min at 4°C and the supernatant used for GS activity determination. GS activity was measured according to Bernard et al. (2008).

GS Immunoblotting

The soluble proteins in each extract were separated through SDS-PAGE (Laemmli, 1970). Equal concentrations of denatured proteins (5 μg) were loaded in each track of a 12% polyacrylamide gel. The proteins were electrophoretically transferred to an ImmunoBlot PVDF membrane (Bio-Rad, München, Germany) with a Trans-Blot Semi-Dry (Bio-Rad, München, Germany) by using a transfer buffer containing 25 mM Tris, 190 mM Glycine, 20% methanol. The electrophoretic transfer was conducted at 15 V for 60 min. After the transfer, the PVDF membrane was soaked in blocking solution (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0,05% Tween 20, and 1% BSA) for 30 min; incubated overnight with primary antibody, and incubated with secondary antibody for 60 min. GS proteins were detected with GS1/GS2 glutamine synthetase global primary polyclonal antibody (Agrisera Vännäs, Sweden); which recognizes both cytoplasmic and chloroplastic forms of the GS enzyme. The secondary antibody was Anti-Rabbit IgG (H+L) HRP conjugate (Promega, Madison, WI, USA). The antibody-protein complex was detected with enhanced chemiluminescence (ECL)-detection reagents (Amersham Buchler Ltd); ImmunoBlot PVDF membrane was incubated for 2 min in ECL, then exposed in an X-ray cassette with Amersham Hyperfilm ECL (Amersham Buchler Ltd) for 2 min. The hyperfilm was soaked in a developing and fixing solution (Kodak Inc.).

Quantitative RT-PCR

Total RNA was extracted with the RNeasy Plant Mini Kit (QIAGEN®), and checked on 1.5% denaturing agarose gels. The total amount of RNA and its purity was determined using a Nano-Drop ND1000 spectrophotometer (Thermo Scientific, Walthman, MA, USA). All RNA samples were adjusted to the same concentration (1 μg) for subsequent treatment with recombinant DNase I (Roche Applied Science, Mannheim, Germany) to remove genomic DNA, and then reverse-transcribed into double stranded cDNA with the Transcriptor First Strand cDNA Synthesis Kit (Roche Applied Science, Mannheim, Germany).

Data were normalized using three reference genes: Cell Division Control AAA-Superfamily of ATPases (CDC),

ADP-Ribosylation Factor (ADP-RF), and RNase L Inhibitor-like protein (RLI; Paolacci et al., 2009; Giménez et al., 2011). These genes were previously used as references in other wheat gene expression studies (Nigro et al., 2013); all three have a stability value around 0.035 when evaluated with NormFinder software (Andersen et al., 2004). In order to pick a primer combination which could detect total GS expression, sequences of known GS genes were aligned in order to find conserved regions. Specifically, cDNAs sequences of both plastidic GS2 (DQ124212, DQ124213 and DQ124214) and cytosolic isoforms GS1 (DQ124209, DQ124210 and DQ124211), GSe (AY491970 and AY491971), and GSr (AY491968 and AY491969), reported by Bernard et al. (2008), were aligned and compared (**Supplementary Figure S1**). The primer combination was chosen in the region with higher homology among them, in particular a fragment of 149 bp (F 5′–3′: CCCTGGCCCCCAGGGTCCATACTACTG; R 5′–3′: GTCATGCCTGGTCAGTGGGAGT).

Quantitative Real-Time PCR analyses to determine GS genes expression levels were carried out using EVA GREEN® in the CFX96™ Real-Time PCR System (Bio-rad). The PCR cycle was 95°C for 3 min, followed by 40 cycles of 95°C for 10 s, 60°C for 30 s. Amplification efficiency (98–100%) for the primer set was determined by amplification of cDNA with a series of six scalar dilutions (1:5) per reaction. Each 10 μl PCR reaction contained 1 μl of a 1:5 dilution of cDNA, 5 μl of EvaGreen Mix 10X (Bio-Rad), and 500 nM of each primer. All experiments were performed in Hard-Shell 96-well skirted PCR plates (HSP9601) with Microseal® 'B' Adhesive Seals (MSB-1001) from Bio-Rad. Fluorescence signals were recorded each cycle. The specificity of each amplicon was confirmed by the presence of a single band of the expected size during agarose gel electrophoresis (2% w/v), single peak melting curves of the PCR products, and sequencing of the amplified fragment. qRT-PCR data for both GS and endogenous controls genes are derived from the mean values of three independent amplification reactions carried out on five different plants harvested in the same phenotypic stage (biological replicates). All calculations and analyses were performed using CFX Manager 2.1 software (Bio-Rad Laboratories) using the ΔCt method, which uses the relative quantity (RQ) calculated with a ratio of the RQ of the target gene to the relative expression of the reference gene (including the three reference targets in each sample). Standard deviations were used to normalize values for the highest or lowest individual expression levels (CFX Manager 2.1 software user manual, Bio-Rad Laboratories).

Statistical Analysis

Values are expressed as mean ± SEM. The medium values reported for GPC, GS activity, and GS expression in the high (H) and low (L) GPC groups were obtained mediating the values of all genotypes belonging to each group (Lucanica, PI191145, PC32, Cannizzo and Svevo for the HGPC; Ciccio, Vesuvio, Appio, Gianni, and Canyon for LGPC). One-way analysis of variance was conducted to calculate differences within and among the groups and for each treatment.

FIGURE 1 | Grain protein content (GPC) of 10 *Triticum durum* genotypes. (A) GPC in 10 wheat genotypes cultivated for 5 years (2009–2013). Values are the mean ± SE of the medium value obtained in each year; different letters indicate significant differences (one-way ANOVA test; $P < 0.05$). **(B)** Average values of GPC for the HGPC and LGPC subgroups; Data are the means ± SE of the GPC of the cultivars belonging to HGPC and LGPC subgroups; different letters indicate values significantly different from one another (one-way ANOVA; $P < 0.05$).

The Dumm's test was used for comparisons among 10 genotypes with no treatment.

The Tukey's test was used for comparisons among four treated genotypes. Correlations were calculated using the Spearman test. Differences were considered significant at P-values <0.05 (two-tailed). Analyses were performed using Sigma Plot software 12.0 (Systat Software, Inc., San Jose, CA, USA).

RESULTS

Glutamine Synthetase Activity and Expression in 10 Durum Wheat Genotypes

The GPC, expressed as percentage of protein per dry weight, was analyzed for five consecutive years from 10 wheat genotypes. The genotypes were classified into either high GPC (HGPC: Lucanica, PI191145, PC32, Cannizzo, and Svevo), or low GPC (LGPC: Ciccio, Vesuvio, Appio, Gianni, and Canyon) groups. The average GPC values of the two groups were significantly different (**Figure 1**).

Both enzyme activity and gene expression of GS, a candidate gene for NUE and GPC, were analyzed in roots and leaves at different phenological stages in the 10 selected wheat genotypes grown in the field during the 2014 season.

Glutamine synthetase activity in roots significantly differ among genotypes. However, the highest GS activities were found in two HGPC genotypes (Cannizzo and Svevo) and the lowest ones in the LGPC genotypes Vesuvio and Canyon (**Figure 2A**). As a consequence the overall mean of GS specific activity of HGPC genotypes was significantly higher than the average value

of LGPC genotypes (**Figure 2B**). Expression data of GS in the roots of each cultivar were consistent with enzyme activity (**Figure 2C**) and again an higher mean value of GS expression was found in the HGPC group when compared with the LGPC one (**Figure 2D**).

Similar trends were observed for GS activity and expression in the leaves at the first leaf stage. Indeed, although differences were found among GS activity and expression of each cultivar, the highest values were found in the HGPC group (PI191145, PC32, and Cannizzo) and the lowest ones in the LGPC group (Vesuvio, Appio, and Canyon; **Figures 3A,C**). Also in this case the medium values of enzyme activity and expression for HGPC genotypes, were significantly higher than that observed in LGPC genotypes (**Figures 3B,D**).

At the flowering stage, differences in GS activity and expression in the leaves of the different genotypes were less marked, even if, also in this case, two HGPC genotypes (PI191145 and Svevo) showed the highest values and two LGPC genotypes (Vesuvio and Appio) the lowest ones (**Figures 4A,C**); the overall mean of GS activity and expression of HGPC genotypes resulted higher than that observed in LGPC genotypes (**Figures 4B,D**).

In the caryopses at the filling stage, the activity and the expression of GS did not change significantly among the genotypes of the two groups, with the exception of PI191145 that showed the highest values, and Vesuvio that had the lowest ones (**Figures 5A,C**). In this case the average values of *GS* activity and expression in the HGPC and LGPC groups did not differ significantly (**Figures 5B,D**).

Regression analysis conducted between GPC, enzymatic activity, and gene expression revealed significant correlation and were reported in **Table 1**.
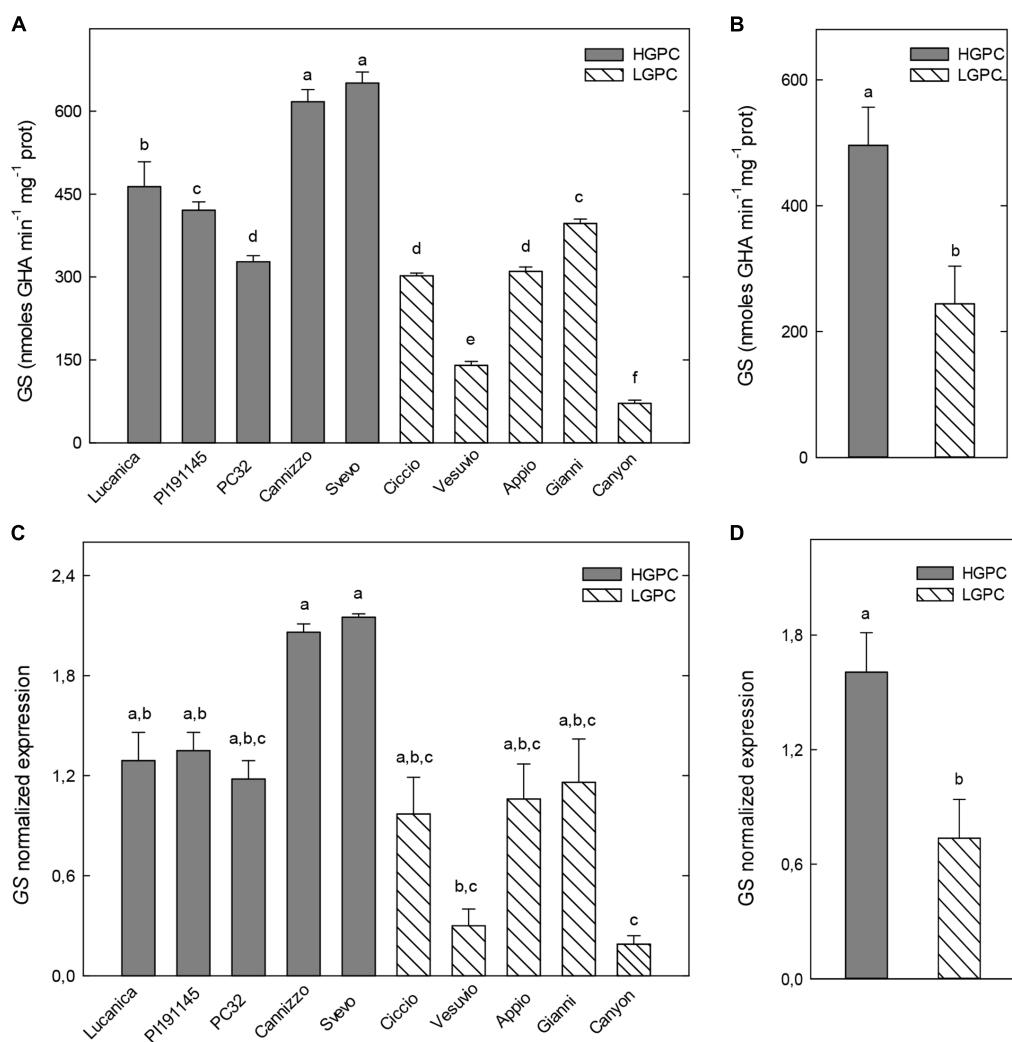
FIGURE 2 | Glutamine synthetase (GS) activity and expression in roots of 10 durum wheat genotypes differing in GPC. (A) GS activity in 10 wheat genotypes. The mean ($\pm$SE; $n = 5$) is presented with different letters representing significant differences (one-way ANOVA test; $P < 0.05$). **(B)** GS activity in HGPC and LGPC subgroups; Data are the means $\pm$ SE of the GS activity of the cultivars belonging to HGPC and LGPC subgroups; different letters indicate significant differences (one-way ANOVA test; $P < 0.05$). **(C)** *GS* normalized fold expression in 10 wheat genotypes The mean ($\pm$SE; $n = 5$) is presented with different letters representing significant differences ($P < 0.05$). **(D)** *GS* expression in HGPC and LGPC subgroups; Data are the means $\pm$ SE of the GS activity of the cultivars belonging to HGPC and LGPC subgroups; different letters indicate significant differences (one-way ANOVA test; $P < 0.05$).

Effect of Nitrogen Treatments on GS Activity and Expression in Different Wheat Genotypes

Two wheat genotypes from each group (PC32 and Cannizzo from the HGPC and Ciccio and Vesuvio from the LGPC) were grown in 2014 at Valenzano (BA) under three different rates of nitrogen application, N0, N60, and N140 units/ha (see Mat and Meth). GS activity and expression were followed in roots, leaves and caryopses of the four selected genotypes grown at different N fertilization.

Root GS activity increased in all genotypes after the N treatment (**Supplementary Figure S2**). However, in the HGPC genotypes PC32 and Cannizzo, the maximum increase in GS

was already evident after application of 60 N units/ha and no further increase occurred after application of 140 N units/ha. The LGPC genotypes, Ciccio and Vesuvio, behaved differently. Ciccio increased root GS activity proportionally to the N application, whereas only the application of 140 N units/ha increased GS activity in Vesuvio roots (**Supplementary Figure S2A**). GS expression in roots had almost the same behavior of GS activity: a maximum increase in gene expression was observed after the application of 60 N units/ha in PC32 and Cannizzo; Ciccio showed an increase of root GS expression proportional to nitrogen supply and Vesuvio had no significant differences between gene expression at N0 and N60, but a significant increase occurred when 140 N units/ha was supplied (**Supplementary Figure S2B**).
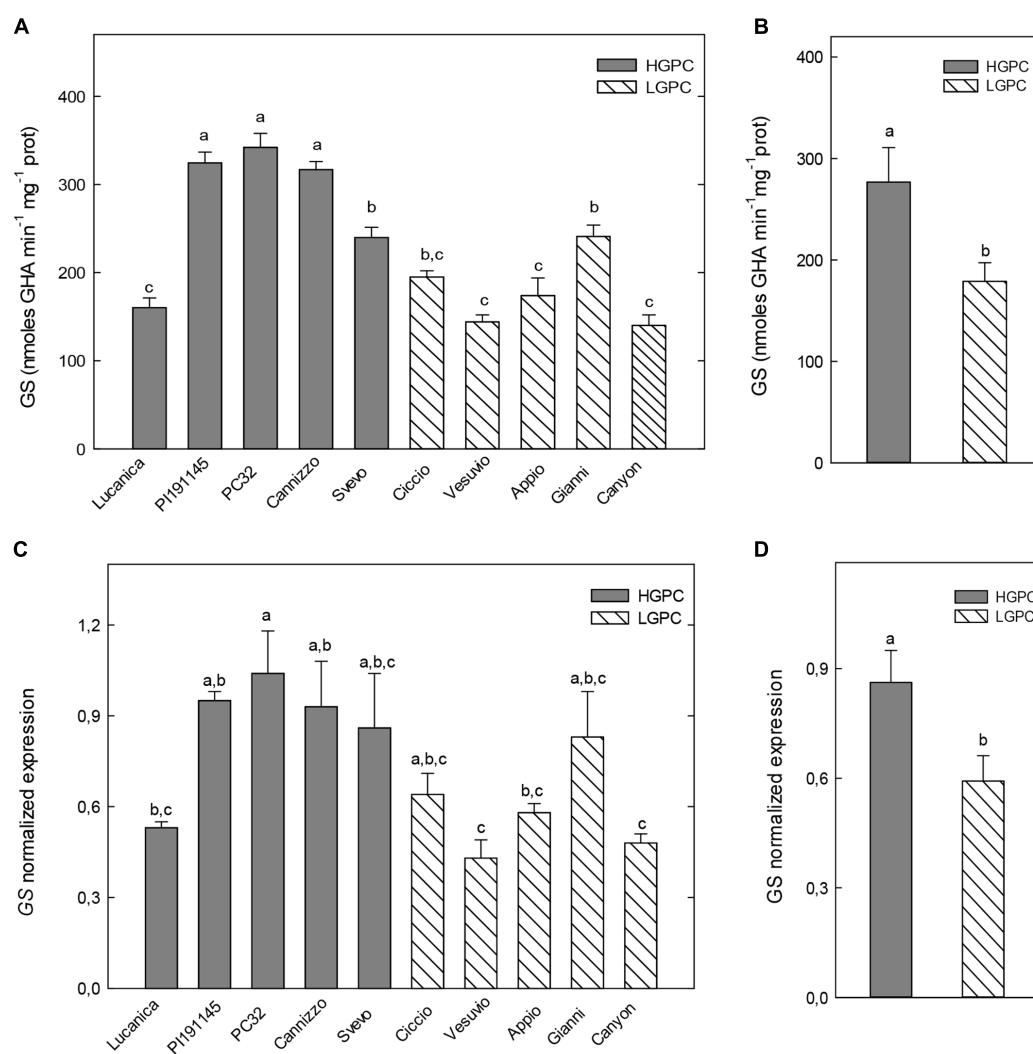
FIGURE 3 | Glutamine synthetase (GS) activity and expression in leaves at the first leaf stage of 10 durum wheat genotypes differing in GPC. (A) GS activity in 10 wheat genotypes. The mean (±SE; $n = 5$) is presented with different letters representing significant differences (one-way ANOVA test; $P < 0.05$). **(B)** GS activity in HGPC and LGPC subgroups; Data are the means ± SE of the GS activity of the cultivars belonging to HGPC and LGPC subgroups; different letters indicate significant differences (one-way ANOVA test; $P < 0.05$). **(C)** *GS* normalized fold expression in 10 wheat genotypes. The mean (±SE; $n = 5$) is presented with different letters representing significant differences ($P < 0.05$). **(D)** *GS* expression in HGPC and LGPC subgroups; Data are the means ± SE of the GS activity of the cultivars belonging to HGPC and LGPC subgroups; different letters indicate significant differences (one-way ANOVA test; $P < 0.05$).

The western blot analysis show the presence of a 40 kDa band in the four genotypes, indicating that only the cytosolic GS isoenzyme was present in the roots. The band intensity in the three different N treatments was consistent with GS transcript level and activity (**Supplementary Figure S2C**).

In leaf tissues at the first leaf stage, GS activity and expression significantly decreased with nitrogen application (N60 and N140 units/ha) in the two HGPC genotypes (PC32 and Cannizzo). On the other hand, nitrogen did not significantly change GS activity and expression in the Ciccio and Vesuvio genotypes (**Supplementary Figures S3A,B**). The Western blot analysis highlighted the presence of two bands of 44 and 40 kDa, indicating that both plastidic and cytosolic isoenzymes, respectively, were active in the leaves. Moreover, GS activity in the

leaves at this stage seemed to be principally due to the plastidic GS, that was more abundant compared to the cytosolic one. Consistently with GS activity and expression, the intensity of the bands of GS proteins after the application of 140 N units/ha decreased in PC32 and Cannizzo and did not show differences in Ciccio and Vesuvio (**Supplementary Figure S3C**).

Glutamine synthetase activity in the leaves at the flowering stage was similar to that observed in the first leaf stage. In PC32 and Cannizzo, GS activity was still highest when no nitrogen was supplied, and decreased significantly with nitrogen applications. On the other hand, Ciccio and Vesuvio genotypes did not show significant differences in GS activity in all N regimes (**Supplementary Figure S4A**). RT-PCR analysis and western blot showed a decrease in the transcript and protein levels
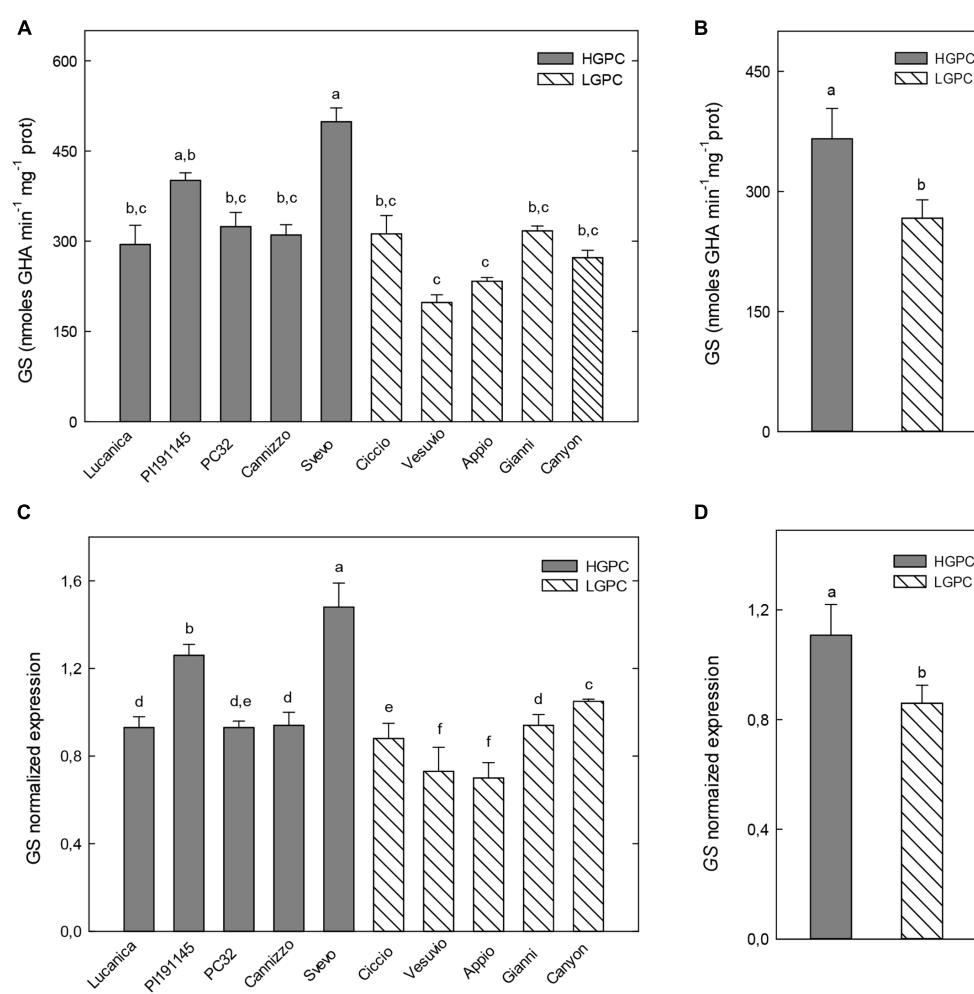
FIGURE 4 | Glutamine synthetase activity and expression in leaves at the flowering stage of 10 durum wheat genotypes differing in GPC. (A) GS activity in 10 wheat genotypes. The mean (±SE; $n = 5$) is presented with different letters representing significant differences (one-way ANOVA test; $P < 0.05$). **(B)** GS activity in HGPC and LGPC subgroups; Data are the means ± SE of the GS activity of the cultivars belonging to HGPC and LGPC subgroups; different letters indicate significant differences (one-way ANOVA test; $P < 0.05$). **(C)** GS normalized fold expression in 10 wheat genotypes. The mean (±SE; $n = 5$) is presented with different letters representing significant differences ($P < 0.05$). **(D)** GS expression in HGPC and LGPC subgroups; Data are the means ± SE of the GS activity of the cultivars belonging to HGPC and LGPC subgroups; different letters indicate significant differences (one-way ANOVA test; $P < 0.05$).

only in the PC32 genotype after application of 140 N units/ha (**Supplementary Figures S4B,C**).

Glutamine synthetase activity in the caryopses at the filling stage significantly increased only in the Vesuvio cultivar, which has the lowest GPC (**Supplementary Figure S5A**). However, when soluble GPC was measured in the three analyzed genotypes under different N treatment, no statistically significant differences were observed in the four genotypes after N supplies (**Supplementary Figure S5B**), indicating that final GPC was not affected by nitrogen application.

DISCUSSION

Nitrogen uptake and utilization is a very complex process in plants, and deciphering all its components is a challenge for scientists and breeders (Hawkesford et al., 2013). The quantitative traits of NUE and GPC are influenced both by the actions of multiple genes and environmental influence (Blanco et al., 2012). In the present work, the enzyme activity and expression of GS, a candidate gene for N-utilization efficiency, were studied in wheat in order to define its role in NUE and GPC. Genetic studies on NUE in maize and rice have shown that GS activity of a cytosolic GS isoform 1 co-localized with QTLs for N remobilization and grain size (Gallais and Hirel, 2004; Obara et al., 2004). In addition, rice mutants lacking the cytosolic GS gene OsGS1;1 were severely limited in growth and grain filling (Tabuchi et al., 2005). In *Triticum aestivum* a QTL for leaf GS activity, mapped to the *TaGSr* locus, co-localized with a QTL for grain N concentration. In this case, increased GS activity was associated with higher grain N. Phenotypic and genotypic correlations between flag leaf weight, soluble protein content and GS activity suggest shared
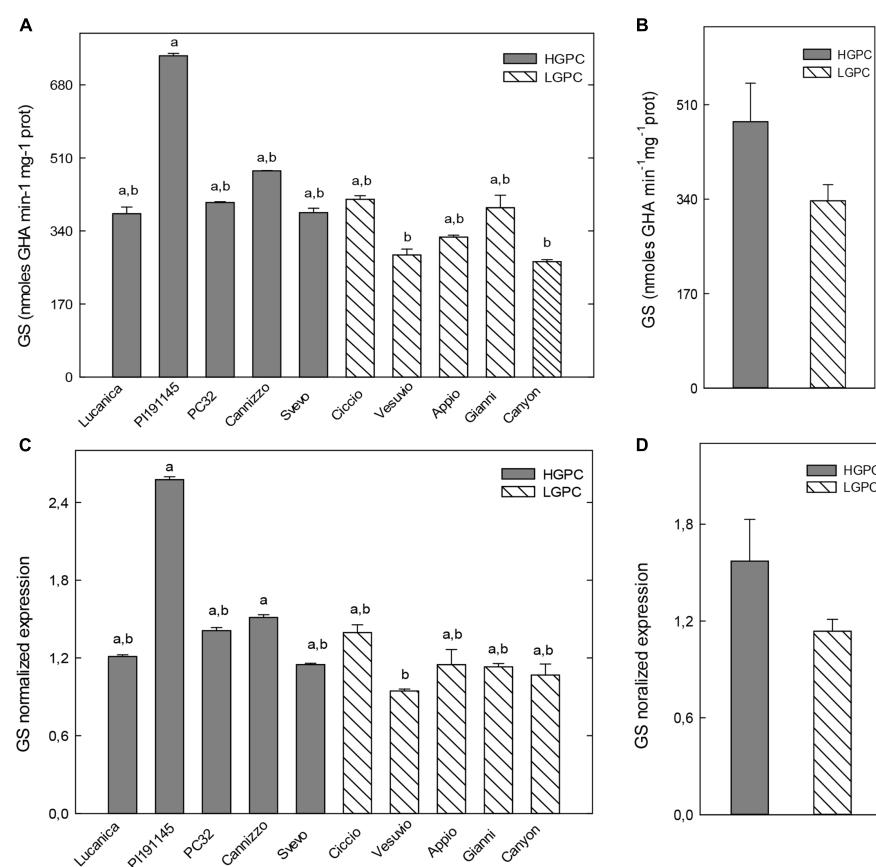
FIGURE 5 | Glutamine synthetase activity and expression in caryopses at the filling stage of 10 durum wheat genotypes differing in GPC. (A) GS activity in 10 wheat genotypes. The mean (±SE; $n = 5$) is presented with different letters representing significant differences (one-way ANOVA test; $P < 0.05$). **(B)** GS activity in HGPC and LGPC subgroups; Data are the means ± SE of the GS activity of the cultivars belonging to HGPC and LGPC subgroups; different letters indicate significant differences (one-way ANOVA test; $P < 0.05$). **(C)** *GS* normalized fold expression in 10 wheat genotypes The mean (±SE; $n = 5$) is presented with different letters representing significant differences ($P < 0.05$). **(D)** *GS* expression in HGPC and LGPC subgroups. Data are the means ± SE of the GS activity of the cultivars belonging to HGPC and LGPC subgroups; different letters indicate significant differences (one-way ANOVA test; $P < 0.05$).

TABLE 1 | Coefficients of correlation (R^2) and probability (*P*-value) between GPC, enzymatic activity, and gene expression.

Trait	R^2	P
GPC-REA	0.35	0.05
GPC-CEA	0.42	0.05
GPC-CGE	0.36	0.05
REA-RGE	0.97	0.001
LIEA-LIGE	0.94	0.001
LIIEA-LIIGE	0.87	0.001
CEA-CGE	0.94	0.001

GPC, grain protein content; REA, root enzymatic activity; RGE, root gene expression; LIEA, leaf I stage enzymatic activity; LIGE, leaf I stage gene expression; LIIEA, leaf II stage enzymatic activity; LIIGE, leaf II stage gene expression; CEA, caryopsis enzymatic activity; CGE, caryopsis gene expression. Significance at $P \leq 0.05$, $P \leq 0.01$, and $P \leq 0.001$ levels, respectively.

control of leaf size and metabolic capacity during grain filling in wheat (Habash et al., 2007). Other two wheat *GS* genes, the plastidic *GS2* and the cytosolic *GS1.3*, have been associated with QTLs for GPC (Gadaleta et al., 2011, 2014). Moreover, in winter

wheat GPC is positively correlated with amino acid and soluble protein content, and with GS activity (Fontaine et al., 2009). Our results show that a clear genotypic variation in GS activity and expression occurs in roots and leaves of the 10 durum wheat genotypes analyzed. However, despite the genotypic variation, the highest GS activities and expression have been found in genotypes of the HGPC group and *vice versa* the lowest ones in the genotypes of the LGPC group. As a consequence, GS activity and expression are on average higher in the HGPC group than in the LGPC one. Another study on five wheat cultivars exhibiting different NUE showed a good correlation between GS activity and the amount of N re-mobilized from the top section of the plant, or even from the flag leaf alone, to the grain (Kichey et al., 2006).

The situation is different in the caryopses at the filling stage, where no significant differences in GS activity and expression between the LGPC and HGPC genotypes were observed. This suggests that GS could be related to the maintenance of critical N flows and sensing during crucial developmental stages, as proposed by Thomsen et al. (2014).

To assess the effect of GS on NUE and GPC, four wheat genotypes were grown under different nitrogen regimes in field conditions. The obtained results are reasonably different for roots and leaves. In roots of all selected wheat genotypes, only cytosolic GS was present. Moreover, after N supply, an increase in GS expression and activity occurred both in the HGPC and LGPC genotypes. These data are consistent with results obtained in *Arabidopsis*; in roots, cytosolic GS is essential for ammonium detoxification and nitrogen assimilation under ample nitrate supply (Lothier et al., 2011). In rice, most of the ammonium taken up by the roots can be assimilated within the organ, as shown by the rapid up-regulation of *OsGS1;2* in the cell layers of the root surface following the supply of ammonium ions (Tabuchi et al., 2007).

The results are quite different in leaves at both the phenological stages considered. In accordance with what reported by Bernard et al. (2008), both plastidic and cytosolic enzymes were detected by Western blot, as two proteins of 44 and 40 kDa, respectively. After supplying nitrogen, total GS activity and expression in leaves did not change in the LGPC genotypes (Ciccio and Vesuvio) and significantly decreased in the HGPC genotypes (Cannizzo and PC32). These results are in accordance with data reported by Tian et al. (2015) showing that the expression of GS genes was higher in the N-efficient wheat genotype than in the N-inefficient one regardless of N treatment. Soluble GPC was not statistically significant in our genotypes after N treatments, implying that the genetic difference between cultivars caused differences in GPC. This is also consistent with the results by Gaju et al. (2011), who in analyzing fourteen UK and French wheat cultivars and two French advanced breeding lines showed that genetic variability in NUE related mainly to differences in N-utilization efficiency, rather than N-uptake efficiency.

Previous studies have reported that when NUE is calculated as a function of grain yield per estimated N input, this decreases with the increasing N input (Grant et al., 1991; Muurinen et al., 2007; Sylvester-Bradley and Kindred, 2009; Anbessa and Juskiw, 2012). The total N uptake of each cultivar in these studies was quite similar, implying that the differences observed in terms of grain yield in response to different N regimes and NUE was best assessed as differences in the efficiency of utilization. This suggests that the rate of nitrogen fertilizer application might be adjusted according to the individual cultivar to improve NUE, while maintaining potential grain yield.

Nitrogen use efficiency is a complex trait that cannot be explained by the action of a single gene. In a recent study on 24 Australian spring wheat genotypes, Mahjourimajd et al. (2016) analyzed how nitrogen supplies can affect NUE and yield in different environmental conditions. They demonstrated that there was significant genetic variation for NUE-related traits among wheat genotypes, allowing them to define a ranking of genotypes for NUE stability. Focusing and explaining the genetic mechanisms underlying traits associated with NUE are essential to contribute to wheat breeding efforts in order to develop high NUE genotypes. In this context, our data contribute to highlight that NUE is a genotype-dependent parameter, and that GS plays a very important role in terms of N utilization. So far, these studies confirm that the efficient management of N through the use of appropriate germplasm is essential for sustainability of agricultural production and that the use of genotypes optimized for traits relating to N-use efficiency rather than yield alone is of primary importance (Hawkesford, 2014). In this view, a more "precision farming" approach could be helpful to guarantee high grain yield while wasting little fertilizer, leading to both economic and environmental benefits.

AUTHOR CONTRIBUTIONS

DN, AG, and MP: Conceived and designed the experiments. DN, SF, SG, and AP: Performed the experiments. AG, AB, and MP: Contributed reagents/materials/analysis tools. DN, AG, MP, YG, and AB: Wrote the paper. DN, AG, MP, YG, and AB: Analyzed the data.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00971

FIGURE S1 | Alignment of glutamine synthetase (GS) genes region chosen for RT-PCR primer design.

FIGURE S2 | Glutamine synthetase (GS) in roots of four selected wheat genotypes bred at three different nitrogen rates. **(A)** GS specific activity; **(B)** GS normalized fold expression; **(A)** GS specific activity and **(B)** GS expression; Data are the means ± SE of five experiments; different letters indicate significant differences after N treatment in each cultivar (one-way ANOVA test; $P < 0.05$). **(C)** Representative image of western blotting analysis with the GS1/GS2 antibody; each well was loaded with 5 μg soluble proteins.

FIGURE S3 | Glutamine synthetase in leaves at the first leaf stage, of four selected wheat genotypes bred at three different nitrogen rates. **(A)** GS specific activity and **(B)** GS expression; data are the means ± SE of five experiments; different letters indicate significant differences after N treatment in each cultivar (one-way ANOVA test; $P < 0.05$). **(C)** Representative image of western blotting analysis with the GS1/GS2 antibody; each well was loaded with 5 μg soluble proteins.

FIGURE S4 | Glutamine synthetase in leaves at the flowering leaf stage, of four selected wheat genotypes bred at three different nitrogen rates. **(A)** GS specific activity and **(B)** GS expression; data are the means ± SE of five experiments; different letters indicate significant differences after N treatment in each cultivar (one-way ANOVA test; $P < 0.05$). **(C)** Representative image of western blotting analysis with the GS1/GS2 antibody; each well was loaded with 5 μg soluble proteins.

FIGURE S5 | Glutamine synthetase and soluble protein content in caryopses at the filling stage of four selected wheat genotypes bred at three different nitrogen rates. **(A)** GS activity and **(B)** soluble protein content. Data are the means ± SE of five experiments; different letters indicate significant differences after N treatment in each cultivar (one-way ANOVA test; $P < 0.05$).

REFERENCES

Anbessa, Y., and Juskiw, P. (2012). Review: strategies to increase nitrogen use efficiency of spring barley. *Can. J. Plant Sci.* 92, 617–625. doi: 10.4141/cjps2011-207

Andersen, C. L., Jensen, J. L., and Orntoft, T. F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64, 5245–5250. doi: 10.1158/0008-5472.CAN-04-0496

Barraclough, P. B., Lopez-Bellido, R., and Hawkesford, M. J. (2014). Genotypic variation in the uptake, partitioning and remobilisation of nitrogen during grain-filling in wheat. *Field Crop Res.* 156, 242–248. doi: 10.1016/j.fcr.2013.10.004

Bernard, S. M., Molle, R. A. L. B., Dionisio, G., Kichey, T., Jahn, T. P., Dubois, F., et al. (2008). Gene expression, cellular localisation and function of glutamine synthetase isozymes in wheat (*Triticum aestivum* L.). *Plant Mol. Biol.* 67, 89–105. doi: 10.1007/s11103-008-9303-y

Blanco, A., Mangini, G., Giancaspro, A., Giove, S., Colasuonno, P., Simeone, R., et al. (2012). Relationships between grain protein content and grain yield components through QTL analyses in a RIL population derived from two elite durum wheat cultivars. *Mol. Breed.* 30, 79–92. doi: 10.1007/s11032-011-9600-z

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3

De Macale, M. A. R., and Velk, P. L. G. (2004). The role of *Azolla* cover in improving the nitrogen use efficiency of lowland rice. *Plant Soil.* 263, 311–321. doi: 10.1023/B:PLSO.0000047742.67467.50

Fischer, R. A., Byerlee, D., and Edmeades, G. O. (2009). *Can technology deliver on the yield challenge to 2050? FAO Expert Meeting on How to Feed the World in 2050*, Rome: FAO, 24–26.

Fontaine, J. X., Ravel, C., Pageau, K., Heumez, E., Dubois, F., Hirel, B., et al. (2009). A quantitative genetic study for elucidating the contribution of glutamine synthetase, glutamate dehydrogenase and other nitrogen-related physiological traits to the agronomic performance of common wheat. *Theor. Appl. Genet.* 119, 645–662. doi: 10.1007/s00122-009-1076-4

Gadaleta, A., Nigro, D., Giancaspro, A., and Blanco, A. (2011). The glutamine synthetase (GS2) genes in relation to grain protein content of durum wheat. *Funct. Integr. Genomics* 11, 665–670. doi: 10.1007/s10142-011-0235-2

Gadaleta, A., Nigro, D., Marcotuli, I., Giancaspro, A., Giove, S. L., and Blanco, A. (2014). Isolation and characterisation of cytosolic glutamine synthetase (GSe) genes and association with grain protein content in durum wheat. *Crop Pasture Sci.* 65, 38–45.

Gaju, O., Allard, V., Martre, P., Snape, J. W., Heumez, E., Le Gouis, J., et al. (2011). Identification of traits to improve the nitrogen-use efficiency of wheat genotypes. *Field Crop Res.* 123, 139–152. doi: 10.1016/j.fcr.2011.05.010

Gallais, A., and Hirel, B. (2004). An approach to the genetics of nitrogen use efficiency in maize. *J. Exp. Bot.* 55, 295–306. doi: 10.1093/jxb/erh006

Garnett, T., Conn, V., and Kaiser, B. N. (2009). Root based approaches to improving nitrogen use efficiency in plants. *Plant Cell Environ.* 32, 1272–1283. doi: 10.1111/j.1365-3040.2009.02011.x

Giménez, M. J., Piston, F., and Atienza, S. G. (2011). Identification of suitable reference genes for normalization of qPCR data in comparative transcriptomics analyses in the Triticeae. *Planta* 233, 163–173. doi: 10.1007/s00425-010-1290-y

Grant, C. A., Gauer, L. E., Gehl, D. T., and Bailey, L. D. (1991). Protein-production and nitrogen-utilization by barley cultivars in response to nitrogen-fertilizer under varying moisture conditions. *Can. J. Plant Sci.* 71, 997–1009. doi: 10.4141/cjps91-142

Habash, D. Z., Bernard, S., Schondelmaier, J., Weyen, J., and Quarrie, S. A. (2007). The genetics of nitrogen use in hexaploid wheat: N utilisation, development and yield. *Theor. Appl. Genet.* 114, 403–419. doi: 10.1007/s00122-006-0429-5

Hawkesford, M. J. (2014). Reducing the reliance on nitrogen fertilizer for wheat production. *J. Cereal Sci.* 59, 276–283. doi: 10.1016/j.jcs.2013.12.001

Hawkesford, M. J., Araus, J. L., Park, R., Calderini, D., Miralles, D., Shen, T., et al. (2013). Prospects of doubling global wheat yields. *Food Energy Secur.* 2, 34–48. doi: 10.1002/fes3.15

Kichey, T., Heumez, E., Pocholle, D., Pageau, K., Vanacker, H., Dubois, F., et al. (2006). Combined agronomic and physiological aspects of nitrogen management in wheat highlight a central role for glutamine synthetase. *New Phytol.* 169, 265–278. doi: 10.1111/j.1469-8137.2005.01606.x

Kraiser, T., Gras, D. E., Gutierrez, A. G., Gonzalez, B., and Gutierrez, R. A. (2011). A holistic view of nitrogen acquisition in plants. *J. Exp. Bot.* 62, 1455–1466. doi: 10.1093/jxb/erq425

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0

Laidò, G., Mangini, G., Taranto, F., Gadaleta, A., Blanco, A., Cattivelli, L., et al. (2013). Genetic diversity and population structure of tetraploid wheats (*Triticum turgidum* L.) estimated by SSR, DArT and pedigree data. *PLoS ONE* 8:e67280. doi: 10.1371/journal.pone.0067280

Lothier, J., Gaufichon, L., Sormani, R., Lemaitre, T., Azzopardi, M., Morin, H., et al. (2011). The cytosolic glutamine synthetase GLN1;2 plays a role in the control of plant growth and ammonium homeostasis in *Arabidopsis* rosettes when nitrate supply is not limiting. *J. Exp. Bot.* 62, 1375–1390. doi: 10.1093/jxb/erq299

Mahjourimajd, S., Kuchel, H., Langridge, P., and Okamoto, M. (2016). Evaluation of Australian wheat genotypes for response to variable nitrogen application. *Plant Soil* 399, 247–255.

Marcotuli, I., Houston, K., Waugh, R., Fincher, G. B., Burton, R. A., Blanco, A., et al. (2015). Genome wide association mapping for arabinoxylan content in a collection of tetraploid wheats. *PLoS ONE* 10:e0132787. doi: 10.1371/journal.pone.0132787

Moll, R. H., Kamprath, E. J., and Jackson, W. A. (1982). Analysis and interpretation of factors which contribute to efficiency of nitrogen utilisation. *Agron. J.* 74, 562–564. doi: 10.2134/agronj1982.00021962007400030037x

Moll, R. H., Kamprath, E. J., and Jackson, W. A. (1987). Development of nitrogen efficient prolific hybrids of maize. *Crop Sci.* 27, 181–186. doi: 10.2135/cropsci1987.0011183X002700020007x

Mulvaney, R. L., Khan, S. A., and Ellsworth, T. R. (2009). Synthetic nitrogen fertilizers deplete soil nitrogen: a global dilemma for sustainable cereal production. *J. Environ. Qual.* 38, 2295–2314. doi: 10.2134/jeq2008.0527

Muurinen, S., Kleemola, J., and Peltonen-Sainio, P. (2007). Accumulation and translocation of nitrogen in spring cereal cultivars differing in nitrogen use efficiency. *Agron. J.* 99, 441–449. doi: 10.2134/agronj2006.0107

Nigro, D., Gu, Y. Q., Huo, N. X., Marcotuli, I., Blanco, A., Gadaleta, A., et al. (2013). Structural analysis of the wheat genes encoding NADH-dependent glutamine-2-oxoglutarate amidotransferases and correlation with grain protein content. *PLoS ONE* 8:e73751. doi: 10.1371/journal.pone.0073751

Obara, M., Sato, T., Sasaki, S., Kashiba, K., Nagano, A., Nakamura, I., et al. (2004). Identification and characterization of a QTL on chromosome 2 for cytosolic glutamine synthetase content and panicle number in rice. *Theor. Appl. Genet.* 110, 1–11. doi: 10.1007/s00122-004-1828-0

Paolacci, A. R., Tanzarella, O. A., Porceddu, E., and Ciaffi, M. (2009). Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Mol. Biol.* 10:11. doi: 10.1186/1471-2199-10-11

Pathak, R. R., Lochab, S., and Raghuram, N. (2011). Plant Systems/ Improving plant nitrogen -use efficiency. *Compr. Biotechnol.* 4, 209–218.

Swarbreck, S. M., Defoin-Platel, M., Hindle, M., Saqi, M., and Habash, D. Z. (2011). New perspectives on glutamine synthetase in grasses. *J. Exp. Bot.* 62, 1511–1522. doi: 10.1093/jxb/erq356

Sylvester-Bradley, R., and Kindred, D. R. (2009). Analysing nitrogen responses of cereals to prioritize routes to the improvement of nitrogen use efficiency. *J. Exp. Bot.* 60, 1939–1951. doi: 10.1093/jxb/erp116

Tabuchi, M., Abiko, T., and Yamaya, T. (2007). Assimilation of ammonium ions and reutilization of nitrogen in rice (*Oryza sativa* L.). *J. Exp. Bot.* 58, 2319–2327. doi: 10.1093/jxb/erm016

Tabuchi, M., Sugiyama, K., Ishiyama, K., Inoue, E., Sato, T., Takahashi, H., et al. (2005). Severe reduction in growth rate and grain filling of rice mutants lacking OsGS1;1, a cytosolic glutamine synthetase1;1. *Plant J.* 42, 641–651. doi: 10.1111/j.1365-313X.2005.02406.x

The Millennium Development Goals Report (2014). *The Millennium Development Goals Report*. New York: United Nations. Available at: http://www.un.org/millenniumgoals/2014%20MDG%20report/MDG%202014%20English%20web.pdf

Thomsen, H. C., Eriksson, D., Moller, I. S., and Schjoerring, J. K. (2014). Cytosolic glutamine synthetase: a target for improvement of crop nitrogen use efficiency? *Trends Plant Sci.* 19, 656–663. doi: 10.1016/j.tplants.2014.06.002

Tian, H., Fu, J., Drijber, R. A., and Gao, Y. J. (2015). Expression patterns of five genes involved in nitrogen metabolism in two winter wheat (*Triticum aestivum* L.) genotypes with high and low nitrogen utilization efficiencies. *J. Cereal Sci.* 61, 48–54. doi: 10.1016/j.jcs.2014.09.007

The Nitrogen Contribution of Different Plant Parts to Wheat Grains: Exploring Genotype, Water, and Nitrogen Effects

Rut Sanchez-Bragado, M. Dolors Serret and José L. Araus*

Plant Physiology Department, University of Barcelona, Barcelona, Spain

The flag leaf has been traditionally considered as the main contributor to grain nitrogen. However, during the reproductive stage, other organs besides the flag leaf may supply nitrogen to developing grains. Therefore, the contribution of the ear and other organs to the nitrogen supplied to the growing grains remains unclear. It is important to develop phenotypic tools to assess the relative contribution of different plant parts to the N accumulated in the grains of wheat which may helps to develop genotypes that use N more efficiently. We studied the effect of growing conditions (different levels of water and nitrogen in the field) on the nitrogen contribution of the spike and different vegetative organs of the plant to the grains. The natural abundance of $\delta^{15}N$ and total N content in the flag blade, peduncle, whole spike, glumes and awns were compared to the $\delta^{15}N$ and total N in mature grains to trace the origin of nitrogen redistribution to the grains. The $\delta^{15}N$ and total N content of the different plant parts correlated positively with the $\delta^{15}N$ and total N content of mature grains suggesting that all organs may contribute a portion of their N content to the grains. The potential contribution of the flag blade to grain N increased (by 46%) as the growing conditions improved, whereas the potential contribution of the glumes plus awns and the peduncle increased (46 and 31%, respectively) as water and nitrogen stress increased. In general, potential contribution of the ear providing N to growing grains was similar (42%) than that of the vegetative parts of the plants (30–40%), regardless of the growing conditions. Thus, the potential ear N content could be a positive trait for plant phenotyping, especially under water and nitrogen limiting conditions. In that sense, genotypic variability existed at least between old (tall) and modern (semidwarf) cultivars, with the ear from modern genotypes exhibiting less relative contribution to the total grain N. The combined use of $\delta^{15}N$ and N content may be used as an affordable tool to assess the relative contribution of different plant parts to the grain N in wheat.

Keywords: nitrogen content, nitrogen isotope composition, ear, grains, wheat

INTRODUCTION

In terrestrial ecosystems, nitrogen is often the most limiting element in plant growth (Vitousek, 1994) and alongside drought stress, a lack of nitrogen can limit crop quality and productivity (Passioura, 2002). In particular, water availability can affect crop growth, soil nitrogen dynamics and the utilization of plant nitrogen from soil and fertilizers (Raimanová and Haberle, 2010).

In addition, grain filling and by extension grain yield is dependent on carbon and nitrogen metabolism (Zhang et al., 2010). The nitrogen requirements of developing grains in wheat are mainly supplied by degradation of proteins derived from different plant organs (Dalling et al., 1976; Simpson et al., 1983; Bancal, 2009). Cereals may accumulate most of their nitrogen in vegetative organs before ear emergence and then redistribute it during grain development (Dalling et al., 1976; Hortensteiner, 2002). Nevertheless, even though an additional source of N may be taken up by roots directly from the soil and assimilated between anthesis and physiological maturity (Dupont and Altenbach, 2003), the greatest proportion of N in the grain is redistributed from the vegetative parts (Perez et al., 1989; Jukanti et al., 2008). In fact, N uptake after anthesis is considered in most cases to be minimal (Perez et al., 1989). Moreover, grain filling is the period when the nitrogen content of different plant parts is substantially reduced (Lopes et al., 2006) as a consequence of protein hydrolysis, which remobilizes amino acids for export to developing grains (Feller and Fischer, 1994).

Traditionally, the flag leaf has been considered as the main contributor to grain nitrogen due to its large protein content (Millard and Grelet, 2010). Indeed, during the last decade, most of the studies dealing with N accumulation in grains have only focused on leaf lamina nitrogen (Hortensteiner, 2002; Bahrani and Joo, 2010). However, during the reproductive stage, as sink strength (grain N) increases after anthesis, other organs besides the flag leaf may become organs of nitrogen supply to developing grains (Waters et al., 1980). Thus, apart from the flag leaf blade, the contribution of the ear as well as the lower parts of the plant may be relevant. Ear photosynthesis is considered as an important source of assimilates for grain filling in wheat and other cereals (Araus et al., 1993; Bort et al., 1994; Tambussi et al., 2007; Sanchez-Bragado et al., 2014a,b) especially under drought conditions (Tambussi et al., 2005). To date, several studies have analyzed the photosynthetic contribution of the ear to grain filling (Tambussi et al., 2007; Maydup et al., 2010, 2012, 2014). Nevertheless, the ear's contribution in terms of nitrogen supply to growing grains still remains unclear, despite studies that have emphasized its importance (Simpson et al., 1983; Lopes et al., 2006). Thus, contribution of other parts of the spike such as the glumes and awns to grain nitrogen might be relevant because these tissues have the longest period of metabolic activity during grain filling (Simpson et al., 1983; Jukanti et al., 2008; Bahrani and Joo, 2010). In fact, measurements of enzyme activities related to nitrogen metabolism (i.e., glutamine synthetase and glutamate dehydrogenase) in floral parts of wheat (awns, glumes) and the flag leaf reveal the existence of ammonia turnover activity in all these organs (Maheswari et al., 1992), suggesting that glumes and awns in addition to the flag leaf can play an important role in nitrogen metabolism during grain filling. For instance, the relative daily contribution of different organs to N accumulation in wheat grains during mid grain filling has been reported as 40% from leaves (leaf lamina and sheath), 23% from glumes, 23% from stems, and 16% from roots (Simpson et al., 1983). However, in advanced stages of grain development, the role of the glumes in terms of contribution to grain nitrogen has been observed to be greater (38%) than in the flag leaf (19%) (Lopes et al., 2006).

Thus, protein content in the flag leaf blade seems to be constant until anthesis but reduces during grain filling, whereas glumes can accumulate proteins until 5 days after anthesis (Waters et al., 1980). Thus, glumes can act as a temporal sink for nitrogen in the absence of alternative sinks prior to rapid grain filling (Dalling et al., 1976). Then, during rapid grain filling as the sink strength of the grain increases, glumes are converted into a nitrogen source, presumably remobilizing their accumulated nitrogen to the grains (Waters et al., 1980; Lopes et al., 2006). In contrast, the nitrogen contribution of the stem to grain nitrogen has been observed to be minor due to its low protein content available for mobilization (and low N content loss) compared to the glumes and the flag leaf (Waters et al., 1980). However it may be relevant to study the potential contribution of this plant part since all the N that leaves remobilize to the grains has to pass through that plant part, especially the upper part of the stem (peduncle). This is particularly important for the leaves below the flag leaf since start to senesce earlier (compared with the flag leaf) during grain filling. Thus, it is important to have a better understanding of the role of the spike as a source of N to growing grains, in order to develop phenotypic tools to assess the relative contribution of different plant parts to the N accumulated in the grains of wheat. Long term, the objective is to develop genotypes that use N more efficiently.

The natural variation of the stable nitrogen isotopes (^{15}N/^{14}N) has been considered as a tool to study nitrogen plant dynamics and as a tracer of the nitrogen sources used by the plant (Evans, 2001; Rossato, 2002; Malagoli et al., 2005). Although it is known that plant nitrogen isotope abundance (δ^{15}N) is linked to nitrogen metabolism (Robinson, 2000; Ellis, 2002; Pritchard and Guy, 2004), underlying biochemical mechanisms that affect nitrogen isotope composition (δ^{15}N) are not yet completely understood (Cernusak et al., 2009). One of the reasons for this might be the fractionation that the nitrogen isotope undergo during enzymatic assimilation of ammonium or nitrate into other forms (Yousfi et al., 2012). Besides, other processes such as volatilization, translocation, or nitrogen recycling in the plant can discriminate positively or negatively against ^{15}N (Robinson et al., 1998; Evans, 2001). Nevertheless, in spite of the discrimination processes affecting δ^{15}N (Evans, 2001), vegetative organs in wheat have been reported to exhibit different δ^{15}N (Lopes et al., 2006; Yousfi et al., 2009, 2013). Furthermore, the natural abundance of the stable N isotopes has been observed to be affected by water availability (Handley et al., 1994; Robinson, 2000; Lopes et al., 2004, 2006; Araus et al., 2013) and the nitrogen source (Cliquet et al., 1990). Thus, providing that nitrogen isotope fractionation from the vegetative organs to the growing grains is negligible (or at least constant; Dawson et al., 2002; Serret et al., 2008), an alternative method to study N uptake and remobilization by the plant could be to compare the δ^{15}N in its natural abundance in the different plant organs during grain filling.

The aim of this work is to study the contribution of different plant parts to the nitrogen in grains under different growing conditions (using different water and nitrogen levels) in a set of old (i.e., tall) and modern (i.e., semidwarf) durum wheat genotypes. To this end, the natural abundance of δ^{15}N and the total N content in the flag leaf blade, peduncle, roots, the whole

spike and different tissues (glumes and awns) of the plant were compared to values of $\delta^{15}N$ and total N content in mature grains in order to trace the origin of the nitrogen redistributed into the grains. The final objective is to assess the performance of the total N content together with the $\delta^{15}N$ in its natural abundance as an affordable tool to assess the relative contribution of different plant parts to the N accumulated in the grains of wheat.

MATERIALS AND METHODS

Germplasm Used and Experimental Conditions

Ten durum wheat [*Triticum turgidum* L. ssp. *durum* (Desf.) Husn.] genotypes were studied: five old Spanish cultivars (*Blanqueta, Griego de Baleares, Negro, Jerez 37,* and *Forment de Artes*) and five modern (i.e., semidwarf) Spanish cultivars delivered after 1990 (*Anton, Bolo, Don Pedro, Regallo,* and *Sula*). Old cultivars were chosen based on their similarity to the phenology of modern cultivars. Field experiments were conducted during two growing seasons, one in 2011 and the other in 2012 (Sanchez-Bragado et al., 2014a), at the experimental station of the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) in Aranjuez (40°03′N, 3°31′E, 500 m asl). In the experimental field, soil is Entisol Fluvent Xerofluvent, with the upper 0.4 m having an organic matter content of 4.9 g/kg, total nitrogen content of 0.37 g/kg, carbonate content of 233 g/kg, pH of 8.1 and electric conductivity of 0.164 dS/m (Araus et al., 2013). During 2012 the five old Spanish genotypes grown under support irrigation conditions were also discarded due to lodging. Two water treatments (support irrigation, SI, and rain-fed, RF) combined with two nitrogen regimes (fertilized, HN, and non-fertilized, LN) were assayed. The trials were planted on 30 December 2010 and 18 November 2011 (from now on designated by their harvest year 2011 and 2012, respectively) in plots with six rows 0.20 m apart, covering a total area of 7.1 m^2 (5 m length and 1.42 m width) per plot. Total accumulated precipitation during the 2011 and 2012 seasons was 275.4 and 126.1 mm, respectively. For both years sprinkler irrigation was applied to irrigated plots around initiation of booting (beginning of April) and grain filling (around May 15th and 30th) with ∼60 mm of water on each date. Prior to sowing, all trials received 60 kg ha^{-1} of phosphorous as superphosphate (18%) and 60 kg ha^{-1} potassium as potassium chloride (60%). Further, the HN plants were dressed with nitrogen applied at the beginning of tillering (January 27th in 2011 and December 29th in 2012) and jointing (March 20th in 2011 and February 20th in 2012) using a dose of 45 and 105 kg ha^{-1} of urea (46%), respectively. The LN plants were not N fertilized, relying exclusively on the N availability in the soil before sowing. Water and nitrogen treatments were arranged according to a split-split plot design with three replicates. Experiment plots were kept free of weeds, insect pests, and diseases by recommended chemical measures (Sanchez-Bragado et al., 2014a).

Phenology was recorded throughout the growth cycle (Zadoks et al., 1974). Sampling was performed around 7 days after anthesis (7th May) in 2011 and 10 days after anthesis (18th

April) in 2012. In 2011 stomatal conductance (g$_s$) was measured with a leaf porometer (Decagon; Pullman, USA) in one leaf per plot at the mid point of grain filling. Similarly, chlorophyll content was measured with a SPAD-502 Minolta chlorophyll meter (Spectrum Technologies, Plainfield, IL, USA). In 2011, roots were collected from the upper layer (0–10 cm) with a split tube (Eijkelkamp Soil & Water, The Netherlands), rinsed with distilled water and then placed inside a paper envelope. Based on similar values of GY obtained from previous studies two old Spanish cultivars (*Blanqueta* and *Negro*) and two modern cultivars (*Anton* and *Bolo*) with three replicates for each growing condition (4 treatments) were selected for root extraction (48 plots). Thereafter, five representative flag leaves, peduncle and ears were collected per plot, and oven dried together with the collected roots at 70°C for 48 h. Once dried, in 2011 the entire spike and flag leaf blade were weighed and ground, whereas in 2012 the glumes, awns, flag leaf blade, and peduncle were separated after drying, weighed, and finely ground for total nitrogen content and nitrogen isotope signature analyses as described below. In addition, based on similar grain nitrogen contents obtained in previous studies, developing grains of a modern cultivar (*Regallo*) and an old Spanish cultivar (*Jerez 37*) from the same set of samples (with four growing conditions and three replicates per genotype) were also separated, weighed and finely ground for total nitrogen content and nitrogen isotope analyses as explained below. At maturity, the central four rows of each plot were harvested and grain yield (GY) recorded. Harvesting was performed manually and by machine in 2011 and 2012, respectively. Subsequently mature grains were processed for N content and isotope analysis. Total nitrogen GY was then calculated as the product of GY by N content on a dry matter basis of mature grains.

Nitrogen Concentration and Stable Isotope Composition

The total N content and stable nitrogen isotope signature in the dry matter of the entire spike, glumes, awns, flag leaf, peduncle, roots, developing, and mature grains were analyzed. Approximately 1 mg of each sample was weighed into tin capsules and measured with an elemental analyser (Flash 1112 EA; ThermoFinnigan, Bremen, Germany) coupled with an Isotope Ratio Mass Spectrometer (Delta C IRMS, ThermoFinnigan, Bremen, Germany) operating in continuous flow mode in order to determine the total N content and the stable nitrogen ($^{15}N/^{14}N$) isotope ratios. The ($^{15}N/^{14}N$) ratios of plant material were expressed in δ notation (Coplen, 2008): $\delta^{15}N = (^{15}N/^{14}N)_{sample}/(^{15}N/^{14}N)_{standard} - 1$, where "sample" refers to plant material and "standard" N$_2$ in air.

Water-Soluble Fraction

The protein-free water-soluble fractions (WSFs) of the flag leaf and spike (entire spike, awns, and glumes) were extracted from the same dry samples tested for nitrogen isotopes, as described previously (Cabrera-Bosquet et al., 2011; Yousfi et al., 2013). Summarizing, 50 mg of either fine leaf or ear powder were suspended with 1 ml of MilliQ water in an Eppendorf tube (Eppendorf Scientific, Hamburg, Germany) for 20 min at

about 5°C. After centrifugation (12,000 g for 5 min at 5°C), the pellet was discarded and the supernatant containing the WSF was heated at 100°C for 3 min, where the heat-denatured proteins precipitated. Subsequently, samples were centrifuged again (12,000 g for 5 min at 5°C) to separate previously denatured proteins from the soluble fraction. Aliquots of 40 μl of supernatant containing protein-free WSF were transferred into tin capsules for nitrogen analysis. The capsules containing the aliquots were oven dried at 60°C.

Total Organ N and Potential, Relative Organ N Contribution to Grain N

Total organ nitrogen content of the flag leaf blade, peduncle, whole spike, glumes, and awns was calculated as the product of nitrogen content on a dry matter basis in the different organs multiplied by their respective dry weight. For the whole spike N calculation, the total grain N of developing grains was subtracted from the calculation (taking into account the treatment and genotype). The potential organ nitrogen contributions of the flag leaf blade, peduncle, spike, glumes, and awns to the nitrogen accumulated in the grains were calculated as the product of N content of each organ multiplied by its respective dry weight and standardized (i.e., divided) by the total N content of mature grains per spike. The relative contribution of the different organs was calculated as the potential organ N contribution for the specific organ divided by the sum of the potential organ N of all organs studied. In addition the ratio between mean values of dry weight (g) under rainfed (RF) vs. supplemental irrigation (SI) was calculated for the spike, flag leaf, peduncle, glumes, and awns.

Statistical Analysis

Treatment, organ, and genotype effects were assessed by means of Analysis of Variance (ANOVA). Water regime, nitrogen supply, organ and their interactions were included as fixed factors. Means were compared by Tukey's HSD test. A bivariate correlation procedure was constructed to analyse the relationships between the measured traits. Statistical analyses were performed using the SPSS 18.0 statistical package (SPSS Inc., Chicago, IL, USA). Figures were created using Microsoft Excel 2010 (Microsoft Corporation).

RESULTS

Grain Yield and Organ Nitrogen Content

Average grain yield (GY) and nitrogen grain yield (nitrogen GY) across treatments and genotypes were lower in 2011 (1.7 Mg·ha^{-1}) than in 2012 (3.1 Mg·ha^{-1}; **Tables 1, 2**, respectively). However, the interaction of nitrogen and water showed an effect on GY in the two-way ANOVA analysis (Table S1). Considering the different treatments, the highest average GY and nitrogen GY was observed under supplemental irrigation, regardless of the N fertilization (fertilized, SI+HN or non-fertilized, SI-LN) conditions for both 2011 (**Table 1**) and 2012 (**Table 2**). Modern cultivars showed higher GY and nitrogen GY than the old cultivars during both growing seasons (Table S1).

In 2011 the average leaf chlorophyll content (SPAD) values across water and nitrogen conditions at the time of organ collection ranged between 52.9 and 52.6 for supplemental irrigation vs. rain-fed and between 55.8 and 49.8 for the fertilized and non-fertilized treatments, respectively (being only significant in fertilized conditions, data not shown). Regardless of the growing conditions, the total grain N content per spike (Grain N·spike^{-1}) was much higher than the total N content of any of the different organs analyzed in 2011 and 2012 (**Tables 1, 2**). Moreover in 2011, irrespective of the organ analyzed and the water regime, the N content on a dry matter basis and the total organ N in fertilized conditions were higher than non-fertilized conditions (**Table 1**). Conversely, in 2012 and regardless of water regime, even though the N content on a dry matter basis was lower in non-fertilized (LN) than in fertilized (HN) conditions, the total organ N of all organs was higher in LN conditions with the exception of the peduncle (**Table 2**). At the same time, the flag leaf blade, peduncle, spike, roots, and awns together with mature grains exhibited higher total organ N content under support irrigation (SI) than rain-fed (RF) conditions (irrespective of nitrogen level) with the exception of the glumes (**Tables 1, 2**). In fact, organs that developed under SI conditions also showed larger dry weight than under RF conditions (Table S2). The organ with the highest total N content on a dry matter basis in 2011 was the spike (**Tables 1**), whereas in 2012 it was the flag leaf blade. Conversely, the glumes in 2012 showed the lowest N content on a dry matter basis (**Table 2**). With regards to genotypic differences, modern cultivars showed higher N content on a dry matter basis in the flag leaf, spike, peduncle, and awns than old cultivars, whereas total organ N was only higher in mature grains, the peduncle, and awns for modern cultivars compared to old cultivars (Table S3). Similarly, only N content showed a genotypic effect in all studied organs, whereas water supply, fertilization and organ showed an effect on almost all studied parameters (**Tables 1, 2**). Besides, the interaction of water supply, fertilization and organ was only significant for N content and for δ^{15}N in WSF (see Table S1).

In 2011, the average nitrogen isotope composition in the dry matter (δ^{15}N DM) within the four growing conditions was enriched in the grains (4.6%) and depleted in the roots (2.2%), whereas the spike and the flag leaf blade were between these values (3.8 and 3.5% for the flag leaf blade and the spike, respectively; **Table 1**). Similarly, in 2012 the δ^{15}N DM of all organs studied (flag leaf blade, peduncle, glumes, and awns) was depleted compared to mature grains (4.5%) (**Table 2**). However, during both seasons the δ^{15}N was enriched in all organs under non-fertilized conditions (SI-LN and RF-LN) compared to fertilized conditions (SI+HN and RF+HN). In particular, the δ^{15}N in all organs under LN conditions was enriched when associated with SI plots (SI-LN) compared with RF plots (RF-LN) for both growing seasons (**Tables 1, 2**). Similarly, under fertilized conditions, the δ^{15}N in all organs was enriched when associated with SI plots (SI+HN) compared to RF plots (RF+HN). With regard to the WSF, δ^{15}N was generally lower (depleted) than in the DM but the trends across water and fertilization regimes and genotypes were similar (**Tables 1, 2**). However, modern cultivars showed enriched δ^{15}N only in mature grains compared to old cultivars (Table S3).

TABLE 1 | Genotype (G), water supply (W), nitrogen (N), and organ (O) effects (ANOVA) and mean values and of nitrogen isotope composition (δ^{15}N) in the water-soluble fraction (WSF) and dry matter (DM) of the flag leaf blade and entire spike as well as in mature grains and roots, nitrogen content (N content), total nitrogen content per organ (Total organ N), grain yield (GY) and total nitrogen content of mature grains per spike (Grain N·spike^{-1}).

	Mean Values				ANOVA			
	SI+HN	SI-LN	WS+HN	WS-LN	G	W	N	O
ROOTS								
δ^{15}N DM (‰)	0.94^b	4.78^c	−0.84^a	4.06^c	ns	ns	***	–
FLAG LEAF								
N content (%DM)	3.98^c	2.85^a	3.67^b	2.85^a	ns	**	***	–
Total organ N (mg)	4.22^a	3.29^a	4.34^a	3.36^a	ns	ns	ns	–
δ^{15}N DM (‰)	1.29^a	6.68^b	1.02^a	6.06^b	ns	ns	***	–
δ^{15}N WSF (‰)	0.53^a	6.20^b	−0.09^a	5.67^b	ns	ns	***	–
SPIKE								
N content (%DM)	1.79^b	1.48^a	1.73^b	1.49^a	***	ns	***	–
Total organ N (mg)	12.38^{ab}	10.49^a	12.83^b	11.09^{ab}	ns	*	*	–
δ^{15}N DM (‰)	1.61^b	6.15^c	0.86^a	5.51^c	ns	**	***	–
δ^{15}N WSF (‰)	3.21^b	7.68^d	2.44^a	7.05^c	ns	**	***	–
GRAIN								
N content (%DM)	2.54^c	1.81^a	2.93^d	2.18^b	*	***	***	–
Grain N·spike^{-1}	24.41^a	21.38^a	21.66^a	20.94^a	**	ns	ns	–
δ^{15}N DM (‰)	3.36^b	6.58^d	2.26^a	6.18^c	ns	***	***	–
GY (Mg·ha^{-1})	2.31^c	1.87^b	1.26^a	1.41^a	ns	***	**	–
Nitrogen GY (Mg·ha^{-1})	0.06^c	0.03^{ab}	0.04^b	0.03^a	ns	***	***	–
N content (%DM)	–	–	–	–	–	–	–	***
Total organ N (mg)	–	–	–	–	–	–	–	***
δ^{15}N DM (‰)	–	–	–	–	–	–	–	***
δ^{15}N WSF (‰)	–	–	–	–	–	–	–	***

Nine durum wheat genotypes (genotype Forment de Artes was discarded due to late phenology) and three replicates per genotype (5 modern, i.e., semidwarf) cultivars for SI and RF conditions and four old (i.e., tall) cultivars only under RF conditions, 84 plots) were considered under rainfed N fertilized (RF+HN) and non-fertilized conditions (RF−LN) and supplemental irrigation N fertilized (SI+HN) and non-fertilized conditions (SI−LN). Sampling was performed 7 days after anthesis and the experiment was performed under field conditions during the 2011 crop season at the INIA's Experimental Station, Aranjuez, Spain. Mean values across plant tissues with different letters are significantly different according to the Tukey's honestly significant difference test (P < 0.05).

Level of significance: ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001.

Potential and Relative Organ N Contribution to Grain N

The potential contribution of the entire ear as a source of N to the grain nitrogen (grain N) was on average (across all growing conditions) almost double (42%) to that of the flag leaf blade (17%) in 2011 (**Figure 1**, upper panel), whereas in 2012 the potential N contribution of the flag leaf blade (33%) to grain N was higher than in the main individual tissues of the spike such as the glumes (13%) or awns (13%) as well as higher than in the peduncle (28%) (**Figure 1**, lower panel). Further, the potential N contribution (i.e., compared with the other plant parts studied) of the flag leaf as a source of N to grain N increased to a 46% with better growing conditions (SI+HN) in 2012 (**Figure 1**, lower panel), whereas the opposite trend occurred for the entire spike (50% in 2011), peduncle (31% in 2012) and glumes and awns (23% in 2012). Indeed, the sum of the potential N contribution of awns and glumes to the total N accumulated in the grains (46%) under RF was comparable to that of the flag leaf blade (46%) under SI conditions (**Figure 1**, lower panel). Moreover, the ratio between the weights of all analyzed organs under RF divided by the weight under SI conditions (Table S2) was lower in the flag leaf (0.42) than other spike organs such as glumes (0.88) and awns

(0.77). Furthermore, regardless of the growing conditions, old cultivars showed higher potential N content in the whole spike, its specific parts (glumes and awns) and the organs below the spike (peduncle and flag leaf) than in the modern cultivars (Figure S1). However, no genotypic effect was observed on the potential N content of the different organs (Table S4).

Correlations across growing conditions between nitrogen grain yield and the relative N content of glumes ($R^2 = 0.69$, $P < 0.001$), awns ($R^2 = 0.46$, $P < 0.001$), and peduncle ($R^2 = 0.42$, $P < 0.001$) were negative (**Figure 2**), whereas the same category of relationships in the flag leaf were positive ($R^2 = 0.78$, $P < 0.001$). Thus, whereas the relative N content of the flag leaf with regard to the nitrogen GY increased with better growing conditions (and thus higher nitrogen grain yield), the relative N contribution in the glumes, awns, and peduncle increased under more stressed conditions (low nitrogen grain yield; **Figure 2**).

The correlations of total N content per organ against the total grain N content per spike across growing conditions supported a greater role for the spike as a whole than the flag leaf blade, at least under the low yielding conditions of 2011. Thus the total N in the spike was better related to total grain N content per spike than the total N in the flag leaf (**Figure 3**, left panel). However, for the

TABLE 2 | Genotype (G), water supply (W), nitrogen (N), and organ (O) effects (ANOVA) and mean values of nitrogen isotope composition (δ^{15}N) in the water-soluble fraction (WSF) and dry matter (DM) of the flag leaf blade, peduncle, glumes, awns as well as in mature grains, nitrogen content (N content DM), total nitrogen content per organ (Total organ N), grain yield (GY), and total nitrogen of mature grains per spike (Grain N·spike^{-1}).

	Mean Values				ANOVA			
	SI+HN	SI−LN	WS+HN	WS−LN	G	W	N	O
FLAG LEAF								
N content (%DM)	3.78a	3.40c	2.66b	2.08a	**	***	***	–
Total organ N (mg)	5.80c	6.53b	1.40a	1.46a	ns	***	ns	–
δ^{15}N DM (‰)	3.70b	6.34c	−1.10a	4.52b	ns	***	***	–
δ^{15}N WSF (‰)	3.55b	2.66b	−0.16a	1.74b	ns	**	ns	–
PEDUNCLE								
N content (%DM)	1.77c	1.59b	1.50b	1.33a	***	***	**	–
Total organ N (mg)	3.47c	3.13bc	2.57ab	2.23a	ns	***	ns	–
δ^{15}N DM (‰)	4.07b	6.98c	−0.12a	4.29b	ns	***	***	–
δ^{15}N WSF (‰)	3.57b	2.48b	−0.79a	1.46b	ns	***	ns	–
GLUMES								
N content (%DM)	1.63d	1.39b	1.53c	1.23a	***	***	***	–
Total organ N (mg)	1.13a	1.20a	1.28a	1.41a	ns	ns	ns	–
δ^{15}N DM (‰)	3.22b	5.83c	−0.03a	4.12b	*	***	***	–
δ^{15}N WSF (‰)	4.14b	3.25b	0.94a	2.17b	ns	**	ns	–
AWNS								
N content (%DM)	2.17c	1.88b	1.41a	1.24a	***	***	***	–
Total organ N (mg)	3.20b	2.93b	1.77a	1.90a	ns	***	ns	–
δ^{15}N DM (‰)	3.09b	5.91c	0.35a	4.37b	**	***	***	–
δ^{15}N WSF (‰)	5.1c	3.47b	0.71a	1.83ab	ns	***	ns	–
GRAINS								
N content (%DM)	2.48b	2.07a	2.39ab	2.21a	ns	ns	**	–
Grain N·spike^{-1}	20.75b	21.94b	11.49a	13.92a	ns	***	***	–
δ^{15}N DM (‰)	4.27b	6.6d	1.9a	5.53c	**	***	***	–
GY (Mg·ha^{-1})	4.37b	4.63b	1.68a	1.66a	ns	**	***	–
Nitrogen GY (Mg·ha^{-1})	0.12c	0.09b	0.04a	0.04a	***	***	*	–
N content (%DM)	–	–	–	–	–	–	–	***
Total organ N (mg)	–	–	–	–	–	–	–	***
δ^{15}N DM (‰)	–	–	–	–	–	–	–	***
δ^{15}N WSF (‰)	–	–	–	–	–	–	–	***

Ten durum wheat genotypes and three replicates per genotype (5 modern cultivars for SI and RF conditions and five old cultivars under RF conditions alone) were considered under rainfed N fertilized (RF+HN) and non-fertilized conditions (RF-LN) and supplemental irrigation N fertilized (SI+HN) and non-fertilized conditions (SI-LN). Sampling was performed 10 days after anthesis and the experiment was performed under field conditions during the 2012 crop season at the INIA's Experimental Station, Aranjuez, Spain. Mean values across plant tissues with different letters are significantly different according to the Tukey's honestly significant difference test (P < 0.05). Level of significance: ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001.

2012 season, correlations of total organ N content, and total grain N content per spike were positive and significant for all the plant parts studied, but those of the flag leaf blade and also the awns were slightly higher ($P < 0.001$) than those of the peduncle and the glumes (**Figure 3**, right panel) in 2012. These results suggest an increase in the role of the flag leaf blade as a source of N to the grains as growing conditions improved (**Figure 3**, right panel).

The relationships between HI and the potential N contribution of the different organs across all genotypes and growing conditions were studied (**Table 4**). Except for the flag leaf in 2012 a negative correlation between the potential N contribution of a given organ and HI was observed, with old cultivars exhibiting, in general, higher ear contributions and lower HIs than the modern cultivars.

Fractionation Nitrogen Isotope Composition

Strong linear correlations (including all growing conditions) were observed between the δ^{15}N in the mature grains and the δ^{15}N in the dry matter of all studied plant organs and growing seasons ($P < 0.001$; **Table 3**, **Figure 4**). Nevertheless, the strongest correlation against the δ^{15}N of the grains was achieved by the flag leaf blade ($r = 0.96$ $P < 0.001$; **Table 3**, **Figure 4**, high-left panel) followed by the entire spike ($r = 0.92$ $P < 0.001$) and the glumes ($r = 0.91$ $P < 0.001$). In addition, in the relationship between the δ^{15}N in the grains and the corresponding values within each organ represented in **Figure 4**, the higher the δ^{15}N in the grains, the more similar the δ^{15}N values of the different organs and the δ^{15}N in the grains (**Figure 4**).

FIGURE 1 | Potential N contribution of the flag leaf blade, peduncle, spike, glumes, and awns to the nitrogen accumulated in the grains. Values were calculated as the product of nitrogen content (N content) of the different organs multiplied by their respective dry weight and standardized by the total nitrogen content of mature grains per spike (Grain N·spike^{-1}). For the total spike N calculation, total grain N of developing grains was subtracted from the calculation (see Section Materials and Methods). Ten durum wheat genotypes (genotype Forment de Artes was discarded due to late phenology in 2011) and three replicates per genotype (totalling 84 plots in 2011 and 90 plots in 2012) were considered under rainfed N fertilized (RF+HN) and non-fertilized conditions (RF−LN) and supplemental irrigation N fertilized (SI+HN) and non-fertilized conditions (SI−LN). Sampling was performed 7 and 10 days after anthesis (2011 and 2012, respectively) and the experiment was performed under field conditions at the INIA's Experimental Station, Aranjuez, Spain during the 2011 (upper panel) and 2012 (lower panel) growing seasons. Mean values across organs and different growing conditions with different letter are significantly different according to the Tukey's honestly significant difference test ($P < 0.05$).

Conversely, the more depleted the $\delta^{15}N$ in the grains the more different (further below) the $\delta^{15}N$ values in the different organs and the $\delta^{15}N$ in the grains in 2011 (Figure S2) and 2012 (**Figure 4**). Specifically, the $\delta^{15}N$ in the flag leaf blade showed the most different values (further below) compared with the $\delta^{15}N$ of the grains, especially under RF conditions. Conversely the $\delta^{15}N$ in the glumes and awns showed constant (i.e., regardless of the $\delta^{15}N$ of the grains) equidistant values away from the $\delta^{15}N$ in the grains (**Figure 4**, left-upper and lower panel). Nevertheless, nitrogen fractionation from source organs to the sink (assumed to be the difference between $\delta^{15}N_{organs}$ minus

$\delta^{15}N_{grains}$) increased with improvements in growing conditions as observed by the increases in the difference between the $\delta^{15}N_{organs}$ minus $\delta^{15}N_{grains}$ associated with an increase in the GY (**Figure 4**, see figure inset). In particular, the flag leaf showed the highest nitrogen fractionation associated with the increase in GY, whereas the awns showed the lowest nitrogen fractionation associated with GY (**Figure 4**, right lower panel, see figure inset). In addition, in 2011 the $\delta^{15}N$ in the flag leaf was positively and strongly related to stomatal conductance ($r = 0.75$; $P < 0.001$) under RF conditions (Figure S3). Moreover, $\delta^{15}N$ in the different organs was positively related to GY and nitrogen grain yield in 2011 and 2012 (**Figure 5**). The organ where $\delta^{15}N$ was best related to GY was the spike ($P < 0.001$) in 2011 and the flag leaf blade ($P < 0.001$) in 2012 (**Figure 5**, right and left panels, respectively).

DISCUSSION

Grain yield (GY) during 2011 was lower than in 2012, and was within the range of GY previously reported under very severe drought stress conditions in the Mediterranean basin (Araus et al., 1998; Oweis et al., 2000). In addition, the interaction of nitrogen and water showed an effect on GY (Table S1). During both years support irrigation had a positive effect on GY, whereas, nitrogen fertilization did not have any positive effect on GY except for SI in 2011, and otherwise tended to have the opposite effect (**Tables 1, 2**). N fertilization could have caused haying off, thus decreasing GY, which might have been triggered by a terminal stress during reproductive stage (Araus et al., 2013).

Effect of Water and N Fertilization in Nitrogen Isotope Composition

RF conditions caused a decrease in organ $\delta^{15}N$ (flag leaf, peduncle, glumes, awns, and grains) compared with the irrigated trial. Likewise, a decrease in $\delta^{15}N$ under stress conditions has been previously reported in shoots (Yousfi et al., 2012) and mature grains of durum wheat (Araus et al., 2013) and bread wheat (Robinson, 2000). The increase in $\delta^{15}N$ in response to SI may be the consequence of labile nitrate derived from chemical fertilizers (urea in our case) with a depleted (near to zero) $\delta^{15}N$ (Bateman and Kelly, 2007) leaching out of the root zone (Hawkesford, 2014) to lower subsoil layers (Raimanová and Haberle, 2010). Alternatively, the increase in plant biomass due to irrigation may lead to exhaustion of the nitrogen fertilizer, therefore causing the crop to rely on natural sources of soil N, which are characterized by higher $\delta^{15}N$ values (Serret et al., 2008). Therefore, apart from chemical fertilizers, nitrogen pools derived from the mineralization of soil organic matter and with an $\delta^{15}N$ enriched signature (Raimanová and Haberle, 2010) may become available to the plant (Evans and Belnap, 1999).

Nitrogen Isotope Composition and Grain Yield

Linear correlations of organ $\delta^{15}N$ with GY and nitrogen GY were strong and positive for both growing seasons (**Figure 5**), as has also been observed in the past with maize (Coque et al., 2006) and durum wheat (Yousfi et al., 2009). Such
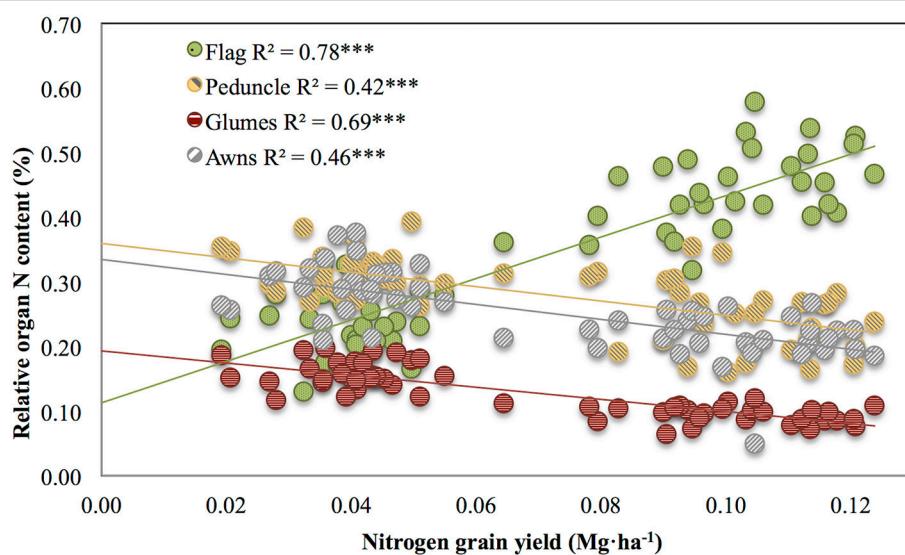
FIGURE 2 | Pearson correlations between the nitrogen grain yield and the relative organ N content of the flag leaf blade, peduncle, glumes, and awns across a wide range of environmental conditions. The relative contribution of the different organs was calculated as the value of the potential organ N contribution for the specific organ divided by the sum of the potential organ N contribution of all the organs studied (flag leaf blade, peduncle, glumes, and awns). Sampling was performed 10 days after anthesis. Ten durum wheat genotypes and three replicates per genotype (totalling 90 plots) were considered, including the four growing conditions (RF+HN, RF−LN, SI+HN, and SI−LN) tested at the INIA's Experimental Station, Aranjuez, Spain in 2012. Levels of significance: ***$P < 0.001$.
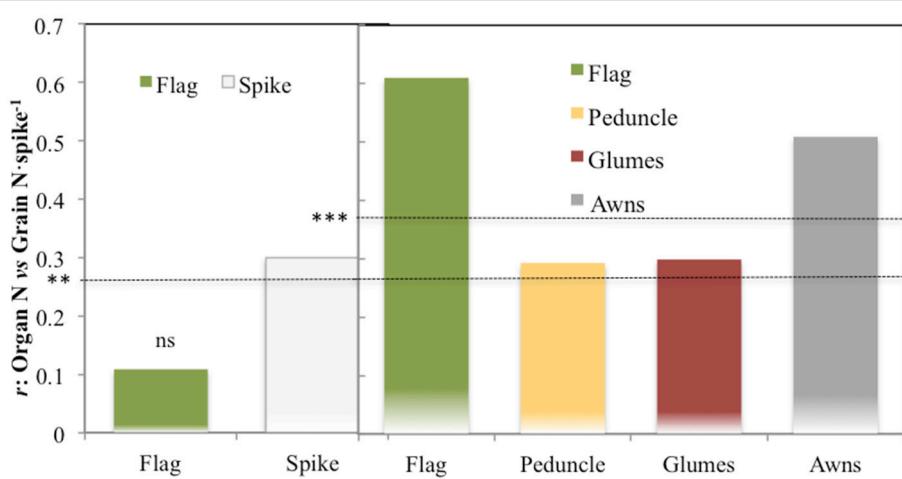
FIGURE 3 | Correlation coefficients of the relationships of the total N content per organ against the total N accumulated in the grains of the spike. The different plant organs assayed were: the flag leaf blade and spike (2011) and the flag leaf blade, peduncle, glumes, and awns (2012). For the calculation, five modern cultivars for SI and RF conditions and five old cultivars under RF conditions alone were considered, with 84 plots in 2011 (as genotype Forment de Artes was discarded due to late phenology in 2011) and 90 in 2012 including the four growing conditions (RF+HN, RF−LN, SI+HN and SI−LN) tested in the INIA's Experimental Station in Aranjuez, Spain in 2011 and 2012. Growing conditions as detailed in the legend of **Figure 1**. Levels of significance: ns, not significant; **$P < 0.01$; ***$P < 0.001$.

positive relationships between $\delta^{15}N$ and GY could be related to some extent with stomatal conductance (Farquhar et al., 1980; Araus et al., 2013). Accordingly, stomatal conductance under RF conditions was positively related to $\delta^{15}N$ in the flag leaf blade in 2011 ($r = 0.75$, $P < 0.001$, Figure S3). Thus, low stomatal conductance may reduce losses of ammonia and nitrous oxides which would reduce the enrichment of $\delta^{15}N$ (Farquhar et al., 1980; Smart and Bloom, 2001). In fact, linear regression

between $\delta^{15}N$ in the mature grains and $\delta^{15}N$ in the dry matter of the different organs (**Figure 4**) supports this finding. Thus, the enricher was the $\delta^{15}N$ in the grains; the more similar were the $\delta^{15}N$ values of the different organs to the $\delta^{15}N$ in the grains (slope 1:1 of the relationships between $\delta^{15}N$ of grains and organs, respectively). Thus, taking into account that for a given nitrogen fertilization value higher values of $\delta^{15}N$ represented better growing conditions (observed by the positive correlation

TABLE 3 | Linear regression of the relationship between the nitrogen isotope composition (δ^{15}N) in the mature grains (δ^{15}N grain DM) against the δ^{15}N in the dry matter (DM) and water-soluble fraction (WSF) of the flag leaf blade, spike and roots (2011) and flag leaf blade, peduncle, glumes and awns (2012).

	vs. δ^{15}N grain DM									
	SI+HN		SI−LN		WS+HN		WS−LN		Global	
2011										
δ^{15}N DM	−0.136	ns	0.545	ns	−0.086	ns	−0.559	ns	**0.801**	***
Flag Leaf										
δ^{15}N DM	**0.810**	***	**0.636**	*	**0.566**	**	0.242	ns	**0.957**	***
δ^{15}N WSF	**0.807**	**	**0.686**	**	0.323	ns	0.083	ns	**0.948**	***
Spike										
δ^{15}N DM	0.457	ns	0.247	ns	**0.431**	*	0.204	ns	**0.920**	***
δ^{15}N WSF	**0.510**	*	**0.704**	*	**0.566**	**	−0.019	ns	**0.935**	***
2012										
Flag Leaf										
δ^{15}N DM	0.043	ns	**0.761**	**	0.533	*	**0.609**	*	**0.932**	***
δ^{15}N WSF	0.198	ns	**0.803**	***	0.002	ns	**−0.562**	**	**0.397**	***
Peduncle										
δ^{15}N DM	0.364	ns	**0.857**	***	**0.610**	***	0.212	ns	**0.869**	***
δ^{15}N WSF	0.156	ns	0.441	ns	−0.034	ns	**−0.562**	**	**0.453**	***
Glumes										
δ^{15}N DM	0.377	ns	**0.860**	***	**0.541**	**	**0.508**	**	**0.910**	***
δ^{15}N WSF	0.112	ns	0.157	ns	−0.076	ns	**−0.431**	*	**0.313**	**
Awns										
δ^{15}N DM	0.165	ns	**0.747**	**	0.263	ns	**0.585**	**	**0.836**	***
δ^{15}N WSF	-0.059	ns	0.434	ns	−0.181	ns	**−0.518**	**	**0.293**	**

Ten durum wheat genotypes and three replicates per genotype (five modern cultivars for SI and RF conditions and five old cultivars only under RF conditions, 84 plots in 2011, and 90 in 2012 as genotype Forment de Artes was discarded due to late phenology in 2011) were considered under rainfed N fertilized (RF+HN) and non-fertilized conditions (RF-LN) and supplemental irrigation N fertilized (SI+HN) and non-fertilized conditions (SI-LN). Sampling was performed 7 and 10 days after anthesis in 2011 and 2012, respectively and the experiment was performed under field conditions at the INIA's Experimental Station, Aranjuez, Spain, during two consecutive seasons (2011 and 2012). The significant correlations are marked in bold. Levels of significance: ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001.

between GY and δ^{15}N), enrichment of δ^{15}N due to either N volatilization or leakage and/or exhaustion of chemical fertilizer (with a lower δ^{15}N) probably occurred; as a consequence values of organ δ^{15}N became closer to the δ^{15}N in the grains (Raimanová and Haberle, 2010). Conversely, different values of δ^{15}N in the grains in comparison to the δ^{15}N of the different organs suggest that reduction in stomatal conductance in response to RF conditions prevented enrichment in the δ^{15}N of the grains. In addition, the flag leaf blade showed the highest nitrogen fractionation associated with the increase in GY compared with the awns, glumes, and peduncle (**Figure 4**, see figure inset). This finding suggests that losses of ^{14}N due to volatilization were higher in the flag leaf due in part to higher stomata density (and thus conductance and transpiration) in the flag leaf (which is an amphystomatic organ) than in the awns (Tambussi et al., 2005; Li et al., 2006). In addition, the δ^{15}N in the roots also showed a similar trend, suggesting that N losses through volatilization or exudation might also be occurring in the roots (Johansson et al., 2009; Figure S1). However, the δ^{15}N in field experiments should be interpreted with caution as miscellaneous biotic and abiotic factors can affect the natural abundance and discrimination of δ^{15}N in the soil-plant system (Hogberg, 1997; Robinson et al., 1998; Evans, 2001; Robinson, 2001; Cernusak et al., 2009; Yousfi et al., 2012). In fact, in our study fractionation of δ^{15}N was

present (**Figure 4**) because the δ^{15}N of the grains was increased in comparison to the values in the individual photosynthetic organs (**Figure 4**). Linear correlations of GY and nitrogen GY against the organ δ^{15}N in the WSF were positive for both growing seasons but weaker (data not shown) compared with the correlation against organ δ^{15}N in the dry matter. Such weaker correlations in the organ δ^{15}N WSF might be related to the fact that the WSF is protein-free because enzymatic N is removed from the WSF and it only contains free amines (Cabrera-Bosquet et al., 2011).

Relative and Potential Organ Contribution to Grain Nitrogen

The relative (i.e., compared with the other plant parts studied) and potential (i.e., relative value with regard to the total N accumulated in the grains of a spike) contributions of the flag leaf as a source of N for grain nitrogen (grain N) increased as growing conditions improved at least during the second year, whereas the opposite occurred for the peduncle, glumes, awns, and entire spike (**Figure 1**). The flag leaf blade has been reported as the main nitrogen exporter to the grains in bread wheat (Simpson et al., 1983). In spite of the high contribution traditionally assigned to the flag leaf blade as a source of N to the growing grains (Evans, 1983; Araus and Tapia, 1987),
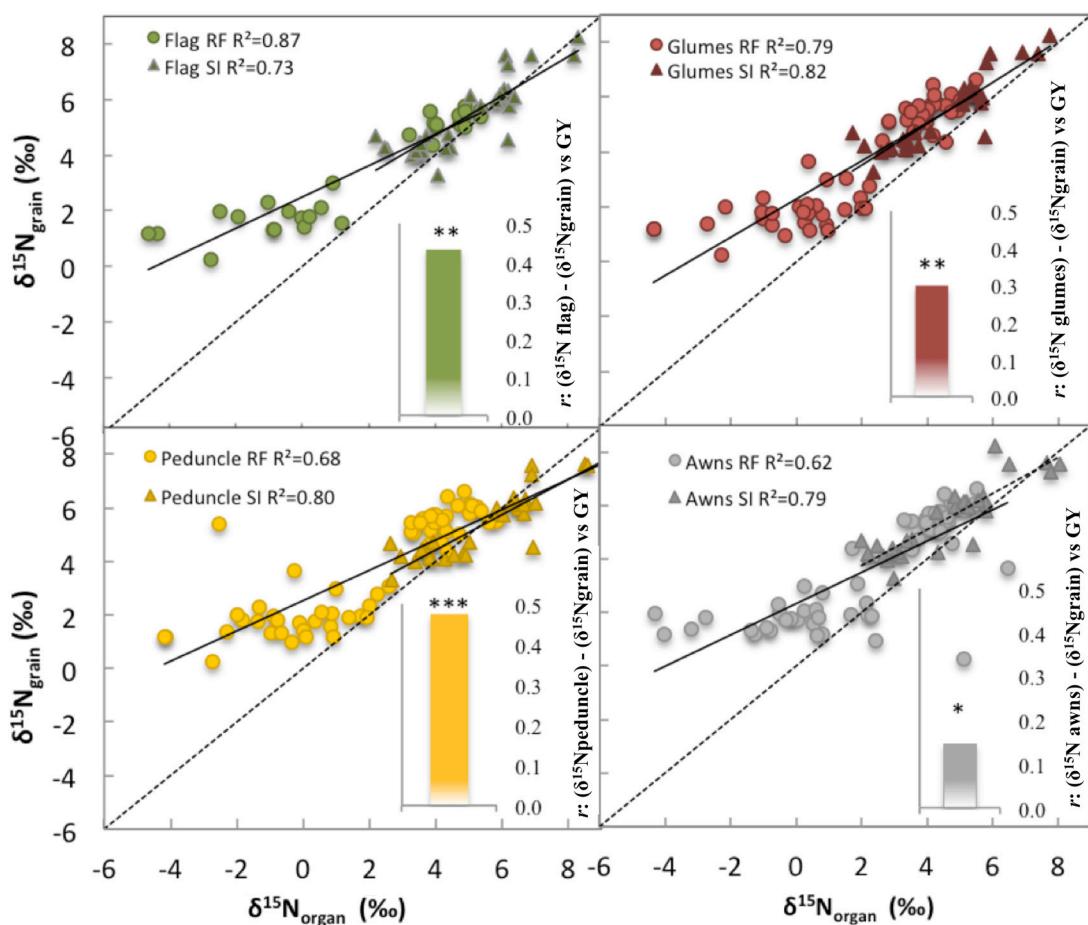
FIGURE 4 | Linear regression of the relationship between the nitrogen isotope composition (δ^{15}N) in the mature grains (δ^{15}Ngrain) against the δ^{15}N in the dry matter of the flag leaf blade, peduncle, glumes, and awns. Figure inset: nitrogen isotope fractionation during translocation to the grains from source organs (flag leaf blade, peduncle, glumes, and awns) to grains in wheat (see figure inset), are displayed as the relationship between the difference in δ^{15}Norgans minus δ^{15}Ngrain (fractionation) vs. grain yield (GY). Ten durum wheat genotypes and three replicates per genotype (90 plots in 2012) were considered including the four growing conditions (RF+HN, RF−LN, SI+HN, and SI−LN) tested at the INIA's Experimental Station in Aranjuez, Spain in 2011 (upper panel) and 2012 (lower panel). Sampling was performed 7 and 10 days after anthesis in 2011 and 2012, respectively. Levels of significance: ns, not significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

in our study, the potential contribution of the flag leaf blade shortly after anthesis was <50% (regardless of the growing conditions; **Figure 1**). Leaf N remobilization to grain N in rice, wheat or maize has been observed to vary from 50 to 90% (Masclaux et al., 2001). Contrastingly, Simpson et al. (1983) reported that less than half of the nitrogen retranslocated from the leaves arrives directly in the grain, whereas the rest is mostly translocated to the roots. However, most of the nitrogen translocated to the roots is further retranslocated to the shoots via xylem sap (Simpson et al., 1983) where it supplies transpirative organs such as the glumes, leaves, and stem (Simpson et al., 1983). However, N directly exported from the roots to non-transpirative organs such as the grains may be a minor player because grains may receive only 1% of the nitrogen exported from the roots (Simpson et al., 1983). Additionally, the relatively low potential N contribution of the flag leaf blade to grain

N may also indicate that aside from other parts of this leaf (such as the sheath), other organs of the plant contribute to the N accumulated in the grains. Leaves below the flag leaf were not considered in this study as their potential contribution to grain N is reported to be minor Del Pozo et al. (2007). Therefore we only considered the upper part of the plant, including the peduncle and the spike (**Figure 1**, upper panel and lower panel). This view is supported by the strong linear correlations (including all growing conditions) between the δ^{15}N in the mature grains and the δ^{15}N of the dry matter of these plant organs ($P < 0.001$; **Table 3**). In spite of this, the strongest correlation against δ^{15}N in the mature grains was achieved by the flag leaf blade followed by the glumes and the entire spike (**Table 3**). Additionally, the correlations of the total N content per organ against the total grain N content per spike also support a slightly greater role for the flag leaf than the
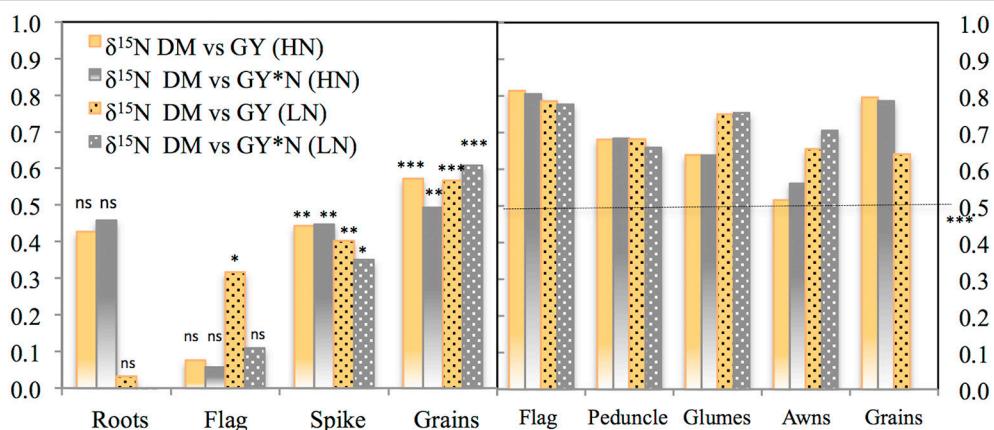
FIGURE 5 | Pearson correlations of the nitrogen grain yield (GY*Ncontent grain) and grain yield (GY) against the nitrogen isotope composition ($\delta^{15}N$) in the dry matter (DM) of the roots, flag leaf blade, peduncle, spike, glumes, and awns. Ten durum wheat genotypes and three replicates per genotype (84 plot in 2011, left panel; and 90 plots in 2012, right panel) were considered including the four growing conditions (RF+HN, RF−LN, SI+HN, and SI−LN) tested at the INIA's Experimental Station, Aranjuez, Spain in 2011 (left panel) and 2012 (right panel). Sampling was performed 7 and 10 days after anthesis in 2011 and 2012, respectively. Levels of significance: ns, not significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

other plant parts (**Figure 3**, right panel). Besides, the experiment conducted under more severe stress in 2011 (growth cycle with low GY) supports the important role of the spike to grain N, since N content in the entire spike was better related to Grain N·spike^{-1} in comparison to the flag leaf (**Figure 3**, left panel). Furthermore, the role of the flag leaf sheath supplying N to the growing grains should be not neglected as a potential source of N (Araus and Tapia, 1987).

Therefore, apart from the potential contribution of the flag leaf (and eventually other vegetative organs), the role of the ear as a supplier of N should be taken into account, as total N allocation in the ear has been observed to be higher than in the flag leaf blade in durum wheat in field chambers (Vicente et al., 2015a). Indeed, accumulation of nitrogen in the grains is closely dependent on N mobilization originating from the glumes (23% of N contribution to the grain) in the absence of an exogenous supply of nitrogen (Simpson et al., 1983). Similarly, in our study the potential and relative N contribution of the awns and glumes as well as the peduncle with respect to the flag leaf blade increased under water stress and non-fertilized conditions (**Figures 1**, **2**, respectively). In fact, the sum of the potential N contribution of awns and glumes under RF conditions was comparable to that of the flag leaf blade under SI conditions (**Figure 1**). In a recent study performed in wheat grown hydroponically, the total N accumulated in the flag leaf blade was comparatively lower than the amount accumulated in the remaining upper parts of the plant (Vicente et al., 2015b). The lower potential N content of the flag leaf blade under RF conditions could be related to organ size but was not concomitant with leaf senescence (as chlorophyll content was similar under RF and SI conditions at the moment of sampling, data not shown). Thus, the ratio between the weight of all analyzed organs under RF divided by the SI conditions (Table S2) was lower in the flag leaf (0.42) than other spike organs such as glumes (0.88) and awns (0.77).

These results suggest that the leaf lamina was smaller under RF, whereas the spike size (entire spike, awns, and glumes) was not reduced under RF conditions (Table S2). On the other hand, growing conditions may also affect the efficiency of N transfer from source organs to the grains (Masclaux-Daubresse et al., 2010). For example, part of the N accumulated in the flag leaf and other leaves may be exported back to the roots, particularly under stress (drought, low fertility) to promote root development mainly in vegetative stages (Palta and Gregory, 1997) and to a lesser extent during grain filling (Jensen, 1994; Swinnen et al., 1994). Conversely, during rapid grain filling glumes may only retranslocate N to the grains, thus increasing the role of N under water stress conditions (Waters et al., 1980). In addition, the potential N contribution of the ear, peduncle and glumes was higher in old than modern cultivars (Figure S2). Such differences might be related, at least in part, to the larger harvest index (HI) in modern cultivars than old cultivars (Reynolds et al., 2009; Foulkes et al., 2011; Sanchez-Bragado et al., 2014a) due to the introduction of *dwarfism alleles* (Maydup et al., 2012). The negative correlations between the potential organ nitrogen contribution and HI (**Table 4**) suggest that old cultivars may provide a greater contribution of N to the grains in relative terms, especially from the ear due to a sink driven phenomenon. In a semidwarf (i.e., modern) genotype the absolute amount of N accumulated in the peduncle may be lower than in a tall (i.e., old) cultivar, although not significant differences were observed (Table S1). Conversely, larger spike in the modern cultivars decrease the surface (where the photosynthetic tissues accumulating N are placed) to volume (where grain N is accumulated) ratio in comparison to spikes in the old cultivars. That is, the spike surface (assumed to be the photosynthetic tissues of the spike such as the glumes and awns) decreases relative to the spike volume (assumed to be the grains). Thus, in modern cultivars a lower relation between spike surface and volume may have

TABLE 4 | Linear regression of the relationship between the HI and Potential N contribution of the flag leaf blade, peduncle, spike, glumes, and awns to the nitrogen accumulated in the grains.

		vs. HI	
	Potential organ N contribution	r	Sig.
2011	Flag	−0.353	0.001
	Spike	−0.315	0.004
2012	Flag	0.227	0.083
	Peduncle	−0.382	0.000
	Glumes	−0.539	0.000
	Awns	−0.357	0.001

Values were calculated as the product of nitrogen content (N content) of the different organs multiplied by their respective dry weight and standardized by the total nitrogen content of mature grains per spike (Grain N·spike^{-1}). For the total spike N calculation, total grain N of developing grains was subtracted from the calculation (see Section Materials and Methods). Ten durum wheat genotypes and three replicates per genotype (five modern cultivars for SI and RF conditions and five old cultivars under RF conditions alone, 84 plots in 2011 and 90 in 2012 as the genotype Forment de Artes was discarded due to late phenology in 2011) were considered combining all growing conditions under rainfed N fertilized (RF+HN) and non-fertilized conditions (RF-LN) and supplemental irrigation N fertilized (SI+HN) and non-fertilized conditions (SI-LN). Sampling was performed 7 and 10 days after anthesis (2011 and 2012, respectively) and the experiment was performed under field conditions at the INIA's Experimental Station, Aranjuez, Spain during the 2011 and 2012.

decreased the potential of the ear to provide nitrogen to the grains compared to old cultivars. Consequently, these results suggest that among the set of genotypes studied, potential spike N content is not only dependent on growing conditions but also on genotype (sink strength and plant height).

Summarizing, the significant correlations between the total N content of the different plant organs studied (flag leaf blade, peduncle, glumes, and awns) against the grain N per spike, suggest that all these organs can potentially export a proportion of their N to the grains. This view was supported by the strong linear correlations (including all growing conditions) between the δ^{15}N in the mature grains and the δ^{15}N in the dry matter of all studied plant organs. Moreover, the large amount of N accumulated in the whole grains of the spike, together with the relatively low amount of N available in the different organs supports the concept that N imported into the grains cannot be sustained by one organ alone; rather, different organs may simultaneously export nitrogen to the grains. In spite of that, the role of the flag leaf blade as a potential supplier of N to grains increased, in comparison to other upper parts of the plant, under improved growing conditions (and thus higher GY) as well as in the modern (semidwarf) cultivars compared to the

old cultivars. In contrast, the relative importance of the ear and peduncle increased under water stress conditions (low GY) or in the old genotypes compared to the new genotypes. Such findings indicate that other than the flag leaf (and eventually other vegetative organs), the role of the ear as a supplier of N should be taken into account even though growing conditions may affect the relative potential contribution of the different plant parts. Thus, the potential ear N content could be a positive trait for plant phenotyping, especially under water limiting and/or low fertility conditions. The total N content of the spike at early grain filling should be considered a trait amenable in crop management (e.g., precision agriculture) as well as for breeding (phenotyping). Nevertheless, the challenge is to find high throughput monitoring techniques for this trait. Besides the specific results achieved, the objective of this study was to test methodologies to asses the potential N contribution of different organs to developing grains. It does not pretend to in detail estimate the integrated in time contributions of different organs to grain N throughout grain filling, but to provide a comparative view across organs potentially useful as an affordable phenotyping tool.

AUTHOR CONTRIBUTIONS

RS and JA conceived designed the study; RS and MS carried out the field measurements; RS conducted laboratory work, RS and JA analyzed the data; RS and JA interpreted the results; RS took the principal role in writing the manuscript under supervision of JL. All authors have contributed to the revision of the manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01986/full#supplementary-material

REFERENCES

Araus, J. L., Amaro, T., Casadesús, J., Asbati, A., and Nachit, M. M. (1998). Relationships between ash content, carbon isotope discrimination and yield in durum wheat. *Aust. J. Plant Physiol.* 25, 835. doi: 10.1071/PP98071

Araus, J. L., Brown, H. R., Febrero, A., Bort, J., and Serret, M. D. (1993). Ear photosynthesis, carbon isotope discrimination and the contribution of respiratory CO_2 to differences in grain mass in durum

wheat. *Plant Cell Environ.* 16, 383–392. doi: 10.1111/j.1365-3040.1993.tb00884.x

Araus, J. L., Cabrera-Bosquet, L., Serret, M. D., Bort, J., and Nieto-Taladriz, M. T. (2013). Comparative performance of δ^{13}C, δ^{18}O and δ^{15}N for phenotyping durum wheat adaptation to a dryland environment. *Funct. Plant Biol.* 40, 595. doi: 10.1071/FP12254

Araus, J. L., and Tapia, L. (1987). Photosynthetic gas exchange characteristics of wheat flag leaf blades and sheaths during grain

filling: the case of a spring crop grown under mediterranean climate conditions. *Plant Physiol.* 85, 667–673. doi: 10.1104/pp. 85.3.667

Bahrani, A., and Joo, M. H. (2010). Flag leaf role in N accumulation and remobilization as affected by nitrogen in a bread and durum wheat cultivars. *Am. J. Agric. Environ. Sci.* 8, 728–735.

Bancal, P. (2009). Decorrelating source and sink determinism of nitrogen remobilization during grain filling in wheat. *Ann. Bot.* 103, 1315–1324. doi: 10.1093/aob/mcp077

Bateman, A. S., and Kelly, S. D. (2007). Fertilizer nitrogen isotope signatures. *Isotopes Environ. Health Stud.* 43, 237–247. doi: 10.1080/10256010701550732

Bort, J., Febrero, A., Amaro, T., and Araus, J. (1994). Role of awns in ear water-use efficiency and grain weight in barley. *Agronomie* 14, 133–139. doi: 10.1051/agro:19940209

Cabrera-Bosquet, L., Albrizio, R., Nogues, S., and Luis Araus, J. (2011). Dual Delta 13C/delta 18O response to water and nitrogen availability and its relationship with yield in field-grown durum wheat. *Plant Cell Environ.* 34, 418–433. doi: 10.1111/j.1365-3040.2010.02252.x

Cernusak, L. A., Winter, K., and Turner, B. L. (2009). Plant $\delta^{15}N$ correlates with the transpiration efficiency of nitrogen acquisition in tropical trees. *Plant Physiol.* 15, 1667–1676. doi: 10.1104/pp.109.145870

Cliquet, J. B., Deléens, E., and Mariotti, A. (1990). C and N mobilization from stalk and leaves during kernel filling by C and N tracing in Zea mays L. *Plant Physiol.* 94, 1547–1553. doi: 10.1104/pp.94.4.1547

Coplen, T. B. (2008). *Explanatory Glossary of Terms Used in Expression of Relative Isotope Ratios and Gas Ratios. IUPAC Provisional Recommendations. Inorganic Chemistry Division. Commission on Isotopic Abundances and Atomic Weights.*

Coque, M., Bertin, P., Hirel, B., and Gallais, A. (2006). Genetic variation and QTLs for ^{15}N natural abundance in a set of maize recombinant inbred lines. *Field Crop. Res.* 97, 310–321. doi: 10.1016/j.fcr.2005.11.002

Dalling, M., Boland, G., and Wilson, J. (1976). Relation between acid proteinase activity and redistribution of nitrogen during grain development in wheat. *Aust. J. Plant Physiol.* 3, 721–730. doi: 10.1071/PP9760721

Dawson, T. E., Mambelli, S., Plamboeck, A. H., Templer, P. H., and Tu, K. P. (2002). Stable isotopes in plant ecology. *Annu. Rev. Ecol. Syst.* 33, 507–559. doi: 10.1146/annurev.ecolsys.33.020602.095451

Del Pozo, A., Pérez, P., Gutiérrez, D., Alonso, A., Morcuende, R., and Martínez-Carrasco, R. (2007). Gas exchange acclimation to elevated CO2 in upper-sunlit and lower-shaded canopy leaves in relation to nitrogen acquisition and partitioning in wheat grown in field chambers. *Environ. Exp. Bot.* 59, 371–380. doi: 10.1016/j.envexpbot.2006.04.009

Dupont, F. M., and Altenbach, S. B. (2003). Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. *J. Cereal Sci.* 38, 133–146. doi: 10.1016/S0733-5210(03)00030-4

Ellis, R. P. (2002). Phenotype/genotype associations for yield and salt tolerance in a barley mapping population segregating for two dwarfing genes. *J. Exp. Bot.* 53, 1163–1176. doi: 10.1093/jexbot/53.371.1163

Evans, J. R. (1983). Nitrogen and photosynyhesis in the flag leaf of wheat (*Triticum aestivum* L.). *Plant Physiol.* 72, 297–302.

Evans, R. D. (2001). Physiological mechanisms influencing plant nitrogen isotope composition. *Trends Plant Sci.* 6, 121–126. doi: 10.1016/S1360-1385(01)01889-1

Evans, R. D., and Belnap, J. (1999). Long-term consequences of disturbance on nitrogen dynamics in an arid ecosystem. *Ecology* 80, 150–160. doi: 10.1890/0012-9658(1999)080[0150:LTCODO]2.0.CO;2

Farquhar, G. D., Firth, P. M., Wetselaar, R., and Weir, B. (1980). On the gaseous exchange of ammonia between leaves and the environment: determination of the ammonia compensation point. *Plant Physiol.* 66, 710–714. doi: 10.1104/pp.66.4.710

Feller, U., and Fischer, A. (1994). Nitrogen metabolism in senescing leaves. *Crit. Rev. Plant Sci.* 13, 241–273. doi: 10.1080/07352689409701916

Foulkes, M. J., Slafer, G. A., Davies, W. J., Berry, P. M., Sylvester-Bradley, R., Martre, P., et al. (2011). Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. *J. Exp. Bot.* 62, 469–486. doi: 10.1093/jxb/erq300

Handley, L. L., Odee, D., and Scrimgeour, C. M. (1994). $\delta^{15}N$ and $\delta^{13}C$ patterns in savanna vegetation: dependence on water availability and disturbance. *Funct. Ecol.* 306–314. doi: 10.2307/2389823

Hawkesford, M. J. (2014). Reducing the reliance on nitrogen fertilizer for wheat production. *J. Cereal Sci.* 59, 276–283. doi: 10.1016/j.jcs.2013.12.001

Hogberg, P. (1997). Tansley Review No. 95. ^{15}N natural abundance in soil-plant systems. *New Phytol.* 137, 179–203. doi: 10.1046/j.1469-8137.1997.00808.x

Hortensteiner, S. (2002). Nitrogen metabolism and remobilization during senescence. *J. Exp. Bot.* 53, 927–937. doi: 10.1093/jexbot/53.370.927

Jensen, B. (1994). Rhizodeposition by field-grown winter barley exposed to 14CO2 pulse-labelling. *Appl. Soil Ecol.* 1, 65–74. doi: 10.1016/0929-1393(94)90024-8

Johansson, E. M., Fransson, P. M. A., Finlay, R. D., and van Hees, P. A. W. (2009). Quantitative analysis of soluble exudates produced by ectomycorrhizal roots as a response to ambient and elevated CO_2. *Soil Biol. Biochem.* 41, 1111–1116. doi: 10.1016/j.soilbio.2009.02.016

Jukanti, A. K., Heidlebaugh, N. M., Parrott, D. L., Fischer, I. A., McInnerney, K., and Fischer, A. M. (2008). Comparative transcriptome profiling of near-isogenic barley (Hordeum vulgare) lines differing in the allelic state of a major grain protein content locus identifies genes with possible roles in leaf senescence and nitrogen reallocation. *New Phytol.* 177, 333–349. doi: 10.1111/j.1469-8137.2007.02270.x

Li, X., Wang, H., Li, H., Zhang, L., Teng, N., Lin, Q., et al. (2006). Awns play a dominant role in carbohydrate production during the grain-filling stages in wheat (*Triticum aestivum*). *Physiol. Plant.* 127, 701–709. doi: 10.1111/j.1399-3054.2006.00679.x

Lopes, M. S., Cortadellas, N., Kichey, T., Dubois, F., Habash, D. Z., and Araus, J. L. (2006). Wheat nitrogen metabolism during grain filling: comparative role of glumes and the flag leaf. *Planta* 225, 165–181. doi: 10.1007/s00425-006-0338-5

Lopes, M. S., Nogués, S., and Araus, J. L. (2004). Nitrogen source and water regime effects on barley photosynthesis and isotope signature. *Funct. Plant Biol.* 31, 995. doi: 10.1071/FP04031

Maheswari, M., Nair, T. V. R., and Abrol, Y. P. (1992). Ammonia metabolism in the leaves and ears of wheat (*Triticum aestivum* L.) during growth and development. *J. Agron. Crop Sci.* 168, 310–317. doi: 10.1111/j.1439-037X.1992.tb01014.x

Malagoli, P., Laine, P., Rossato, L., and Ourry, A. (2005). Dynamics of nitrogen uptake and mobilization in field-grown winter oilseed rape (*Brassica napus*) from stem extension to harvest: I. Global N flows between vegetative and reproductive tissues in relation to leaf fall and their residual N. *Ann. Bot.* 95, 853–861. doi: 10.1093/aob/mci091

Masclaux, C., Quillere, I., Gallais, A., and Hirel, B. (2001). The challenge of remobilisation in plant nitrogen economy. A survey of physio-agronomic and molecular approaches. *Ann. Appl. Biol.* 138, 69–81. doi: 10.1111/j.1744-7348.2001.tb00086.x

Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., and Suzuki, A. (2010). Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. *Ann. Bot.* 105, 1141–1157. doi: 10.1093/aob/mcq028

Maydup, M. L., Antonietta, M., Graciano, C., Guiamet, J. J., and Tambussi, E. A. (2014). The contribution of the awns of bread wheat (*Triticum aestivum* L.) to grain filling: responses to water deficit and the effects of awns on ear temperature and hydraulic conductance. *Field Crop Res.* 167, 102–111. doi: 10.1016/j.fcr.2014.07.012

Maydup, M. L., Antonietta, M., Guiamet, J. J., Graciano, C., López, J. R., and Tambussi, E. A. (2010). The contribution of ear photosynthesis to grain filling in bread wheat (*Triticum aestivum* L.). *Field Crop Res.* 119, 48–58. doi: 10.1016/j.fcr.2010.06.014

Maydup, M. L., Antonietta, M., Guiamet, J. J., and Tambussi, E. A. (2012). The contribution of green parts of the ear to grain filling in old and modern cultivars of bread wheat (*Triticum aestivum* L.): evidence for genetic gains over the past century. *Field Crop Res.* 134, 208–215. doi: 10.1016/j.fcr.2012.06.008

Millard, P., and Grelet, G.-A. (2010). Nitrogen storage and remobilization by trees: ecophysiological relevance in a changing world. *Tree Physiol.* 30, 1083–1095. doi: 10.1093/treephys/tpq042

Oweis, T., Zhang, H., and Pala, M. (2000). Water use efficiency of rainfed and irrigated bread wheat in a mediterranean environment. *Agron. J.* 92, 231. doi: 10.2134/agronj2000.922231x

Palta, J. A., and Gregory, P. J. (1997). Drought affects the fluxes of carbon to roots and soil in 13C pulse-labelled plants of wheat. *Soil Biol. Biochem.* 29, 1395–1403. doi: 10.1016/S0038-0717(97)00050-3

Passioura, J. B. (2002). Review: environmental biology and crop improvement. *Funct. Plant Biol.* 29, 537. doi: 10.1071/FP02020

Perez, P., Martinez-Carrasco, R., Molino, I. M., Martin del Rojo, B., and Ulloa, M. (1989). Nitrogen uptake and accumulation in grains of three winter wheat varieties with altered source—sink ratios. *J. Exp. Bot.* 40, 707–710. doi: 10.1093/jxb/40.6.707

Pritchard, E. S., and Guy, R. D. (2004). Nitrogen isotope discrimination in white spruce fed with low concentrations of ammonium and nitrate. *Trees* 19, 89–98. doi: 10.1007/s00468-004-0367-2

Raimanová, I., and Haberle, J. (2010). The effects of differentiated water supply after anthesis and nitrogen fertilization on $\delta^{15}N$ of wheat grain. *Rapid Commun. Mass Spectrom.* 24, 261–266. doi: 10.1002/rcm.4382

Reynolds, M. P., Foulkes, M. J., Slafer, G. A., Berry, P., Parry, M. A, Snape, J. W., et al. (2009). Raising yield potential in wheat. *J. Exp. Bot.* 60, 1899–1918. doi: 10.1093/jxb/erp016

Robinson, D. (2000). Using stable isotope natural abundances (delta15N and delta13C) to integrate the stress responses of wild barley (*Hordeum spontaneum* C. Koch.) genotypes. *J. Exp. Bot.* 51, 41–50. doi: 10.1093/jexbot/51.342.41

Robinson, D. (2001). $\delta^{15}N$ as an integrator of the nitrogen cycle. *Trends Ecol. Evol.* 16, 153–162. doi: 10.1016/S0169-5347(00)02098-X

Robinson, D., Handley, L. L., and Scrimgeour, C. M. (1998). A theory for 15 N/ 14 N fractionation in nitrate-grown vascular plants. *Planta* 205, 397–406. doi: 10.1007/s004250050336

Rossato, L. (2002). Nitrogen storage and remobilization in *Brassica napus* L. during the growth cycle: effects of methyl jasmonate on nitrate uptake, senescence, growth, and VSP accumulation. *J. Exp. Bot.* 53, 1131–1141. doi: 10.1093/jexbot/53.371.1131

Sanchez-Bragado, R., Elazab, A., Zhou, B., Serret, M. D., Bort, J., Nieto-Taladriz, M. T., et al. (2014a). Contribution of the ear and the flag leaf to grain filling in durum wheat inferred from the carbon isotope signature: genotypic and growing conditions effects. *J. Integr. Plant Biol.* 56, 444–454. doi: 10.1111/jipb.12106

Sanchez-Bragado, R., Molero, G., Reynolds, M. P., and Araus, J. L. (2014b). Relative contribution of shoot and ear photosynthesis to grain filling in wheat under good agronomical conditions assessed by differential organ δ13C. *J. Exp. Bot.* 65, 5401–5413. doi: 10.1093/jxb/eru298

Serret, M., Ortiz-Monasterio, I., Pardo, A., and Araus, J. (2008). The effects of urea fertilisation and genotype on yield, nitrogen use efficiency, $\delta^{15}N$ and $\delta^{13}C$ in wheat. *Ann. Appl. Biol.* 153, 243–257. doi: 10.1111/j.1744-7348.2008.00259.x

Simpson, R. J., Lambers, H., and Dalling, M. J. (1983). Nitrogen redistribution during grain growth in wheat (*Triticum aestivum* L.): IV. development of a quantitative model of the translocation of nitrogen to the grain. *Plant Physiol.* 71, 7–14. doi: 10.1104/pp.71.1.7

Smart, D. R., and Bloom, A. J. (2001). Wheat leaves emit nitrous oxide during nitrate assimilation. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7875–7878. doi: 10.1073/pnas.131572798

Swinnen, J., Van Veen, J. A., and Merckx, R. (1994). ^{14}C pulse-labelling of field-grown spring wheat: an evaluation of its use in

rhizosphere carbon budget estimations. *Soil Biol. Biochem.* 26, 161–170. doi: 10.1016/0038-0717(94)90159-7

Tambussi, E. A., Bort, J., Guiamet, J. J., Nogués, S., and Araus, J. L. (2007). The photosynthetic role of ears in C 3 cereals: metabolism, water use efficiency and contribution to grain yield. *Crit. Rev. Plant Sci.* 26, 1–16. doi: 10.1080/07352680601147901

Tambussi, E. A., Nogués, S., and Araus, J. L. (2005). Ear of durum wheat under water stress: water relations and photosynthetic metabolism. *Planta* 221, 446–458. doi: 10.1007/s00425-004-1455-7

Vicente, R., Pérez, P., Martínez-Carrasco, R., Gutiérrez, E., and Morcuende, R. (2015a). Nitrate supply and plant development influence nitrogen uptake and allocation under elevated CO_2 in durum wheat grown hydroponically. *Acta Physiol. Plant.* 37, 114. doi: 10.1007/s11738-015-1867-y

Vicente, R., Pérez, P., Martínez-Carrasco, R., Usadel, B., Kostadinova, S., and Morcuende, R. (2015b). Quantitative RT-PCR platform to measure transcript levels of C and N metabolism-related genes in durum wheat: transcript profiles in elevated $[CO_2]$ and high temperature at different levels of N supply. *Plant Cell Physiol.* 56, 1556–1573. doi: 10.1093/pcp/pcv079

Vitousek, P. M. (1994). Beyond global warming: ecology and global change. *Ecology* 75, 1861–1876. doi: 10.2307/1941591

Waters, S. P., Peoples, M. B., Simpson, R. J., and Dalling, M. J. (1980). Nitrogen redistribution during grain growth in wheat (*Triticum aestivum* L.). *Planta* 148, 422–428. doi: 10.1007/BF02395309

Yousfi, S., Dolores Serret, M., Jose Marquez, A., Voltas, J., and Luis Araus, J. (2012). Combined use of $\delta^{13}C$, $\delta^{18}O$ and $\delta^{15}N$ tracks nitrogen metabolism and genotypic adaptation of durum wheat to salinity and water deficit. *New Phytol.* 194, 230–244. doi: 10.1111/j.1469-8137.2011.04036.x

Yousfi, S., Serret, M. D., and Araus, J. L. (2009). Shoot $\delta^{15}N$ gives a better indication than ion concentration or Δ 13 C of genotypic differences in the response of durum wheat to salinity. *Funct. Plant Biol.* 36, 144. doi: 10.1071/FP08135

Yousfi, S., Serret, M. D., and Araus, J. L. (2013). Comparative response of $\delta^{13}C$, $\delta^{18}O$ and $\delta^{15}N$ in durum wheat exposed to salinity at the vegetative and reproductive stages. *Plant. Cell Environ.* 36, 1214–1227. doi: 10.1111/pce.12055

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zhang, H., Turner, N. C., and Poole, M. L. (2010). Source–sink balance and manipulating sink–source relations of wheat indicate that the yield potential of wheat is sink-limited in high-rainfall zones. *Crop Pasture Sci.* 61, 852. doi: 10.1071/CP10161

Transcriptomic Analysis for Different Sex Types of *Ricinus communis* L. during Development from Apical Buds to Inflorescences by Digital Gene Expression Profiling

Meilian Tan[1], Jianfeng Xue[1], Lei Wang[1], Jiaxiang Huang[2], Chunling Fu[1] and Xingchu Yan[1]*

[1] Key Laboratory of Biology and Genetic Improvement of Oil Crops, Ministry of Agriculture, Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan, China, [2] Castor Oil Research Institute of Jiaxiang, Zibo, China

The castor plant (*Ricinus communis* L.) is a versatile industrial oilseed crop with a diversity of sex patterns, its hybrid breeding for improving yield and high purity is still hampered by genetic instability of female and poor knowledge of sex expression mechanisms. To obtain some hints involved in sex expression and provide the basis for further insight into the molecular mechanisms of castor plant sex determination, we performed DGE analysis to investigate differences between the transcriptomes of apices and racemes derived from female (JXBM0705P) and monoecious (JXBM0705M) lines. A total of 18 DGE libraries were constructed from the apices and racemes of a wild monoecious line and its isogenic female derivative at three stages of apex development, in triplicate. Approximately 5.7 million clean tags per library were generated and mapped to the reference castor genome. Transcriptomic analysis showed that identical dynamic changes of gene expression were indicated in monoecious and female apical bud during its development from vegetation to reproduction, with more genes expressed at the raceme formation and infant raceme stages compare to the early leaf bud stage. More than 3000 of differentially expressed genes (DEGs) were detected in *Ricinus* apices at three developmental stages between two different sex types. A number of DEGs involved in hormone response and biosynthesis, such as auxin response and transport, transcription factors, signal transduction, histone demethylation/methylation, programmed cell death, and pollination, putatively associated with sex expression and reproduction were discovered, and the selected DEGs showed consistent expression between qRT-PCR validation and the DGE patterns. Most of those DEGs were suppressed at the early leaf stage in buds of the mutant, but then activated at the following transition stage (5-7-leaf stage) of buds in the mutant, and ultimately, the number of up-regulated DEGs was equal to that of down-regulation in the small raceme of the mutant. In this study, a large number of DEGs and some suggestions involved in sex expression and reproduction were discovered using DGE analysis, which provides large information and valuable hints for next insights into the molecular mechanism of sex determination. It is useful for other further studies in *Ricinus*.

Keywords: castor bean, digital expression profile, pistillate, sex determination, transcriptome analysis

INTRODUCTION

Castor (*Ricinus communis* L.), an important industrial oilseed crop belonging to the Euphorbiaceae family, grows as an indeterminate annual or perennial depending on climate and soil types in tropical, sub-tropical, and warm temperate regions (Anjani, 2012). Because of the high fatty acid content in its seeds (more than 45%) and the rich ricinoleic acid content in its oil (80–90%), castor is a versatile raw material in industrial chemistry: for example, castor oil can be used for the production of lubricants, nylon, hydraulic and brake fluids, paints, dyes, coatings, inks, cold resistant plastics, waxes and polishes, pharmaceuticals, perfumes, and biodiesel (Jeong and Park, 2009; Halilu et al., 2013). To date, several high-yield varieties and hybrids have been developed (Baldanzi and Pugliesi, 1998; Amaral, 2003; Savy Filho, 2005; Pranavi et al., 2011), but to meet the tremendous global demand for castor oil, cultivars with even higher yield and oil content are needed (Anjani, 2012). However, hybrid breeding for improved yield and high purity in *Ricinus* is still constrained by the genetic instability of females and the unknown mechanism of sex expression.

In castor plant, the standard type of inflorescence is gradient monoecious raceme (female flowers at the apex and staminate flowers on the lower portion) (Shifriss, 1960). However, a wide variation of inflorescence patterns occurs in natural cultivation, including other kinds of racemes such as strictly pistillate (bearing only female flowers), male (only staminate flowers), apically interspersed (monoecism with interspersed male flowers in the apical pistillate region), and entirely interspersed (female and male flowers uniformly interspersed) (George and Shifriss, 1967). In addition, inflorescence setting with one or a few hermaphrodite flowers occurs occasionally (Jacob, 1963). Accordingly, for castor individuals, there are several sex models, including normal monoecism, sex reversal, interspersed sexuality, and strictly female (Shifriss, 1956; Jacob and Atsmon, 1965; George and Shifriss, 1967).

In the past, scientists focused on the sexuality of *Ricinus* to interpret the inheritance and instability of sex variation. Shifriss (1956) described the sex tendency, sex patterns, inheritance and reversion of *Ricinus*, and proposed a hypothesis: Gene *F* controls a genetically stable series of sex variants ranging from female (*f*) to strongly male inbreds, and an unknown factor could affect male tendency, sex reversal, and sex instability by suppressing gene *F* or mutating itself. Later, he described sex variations of *Ricinus* in two genetic systems, which he tentatively named as "conventional" and "unconventional" (Shifriss, 1960). Monoecious variants and rare recessive female mutants were ascribed to the conventional form, sex reversals and non-reverted females belonged to the unconventional form. Monoecism is governed by two major groups of genes: qualitative genes that determine flower type, and polygenes regulating gradient differentiation and racial differences in sex tendency. In addition, gene modifiers such as *id* and *th* were also considered to affect the pattern of sex differentiation. He postulated that genetic instability of spontaneous mutants, so-called position-effect variegation, may result from a rearrangement undergoing two basic kinds of genetic changes: mutation into new hereditary

potentialities, and transformation between an "active" and "inactive" state. Moreover, further evidence regarding the former findings (Shifriss, 1956, 1960) revealed that femaleness transmits more effectively to progeny through female inflorescences of sex-reversal plants than through reverted monoecious inflorescences of the same plant (Jacob and Atsmon, 1965).With regard to interspersed sexuality, Shifriss (1956, 1960) believed that the interspersed pattern of sex differentiation is determined by hereditary factors for femaleness and genes for interspersed staminate flowers (*id*). George and Shifriss (1967) investigated more deeply into the inheritance of interspersed inflorescence patterns, and concluded that the level of expressivity of interspersed staminate flowers depends on the dosage of two independent genes (*id1* and *id2*), their loci, and the environment.

In addition to genetic regulation, sex expression or variation of *Ricinus* is simultaneously affected to some degree by fluctuations in the environment, including temperature, vegetative activity, nutrition level, pruning, and seasonal variations (Shifriss, 1956), as well as by plant hormones (Shifriss, 1961; Philipos and Narayanaswamy, 1976; Kumar and Rao, 1980; Mohan Ram and Sett, 1980; Varkey and Nigam, 1982; Tan et al., 2011). Despite a great deal of progress, the hypotheses described above are vulnerable in the absence of cytological information, and the molecular mechanisms of sex variation and genes determining sex expression in *Ricinus* remain poorly understood.

Genome-wide analyses have dramatically improved the efficiency of gene discovery. Next-generation sequencing (NGS) technologies provide new approaches for global measurements of gene expression. Due to its high efficiency and low cost, NGS has become an attractive alternative method for more efficient study of the genome, epigenome, and transcriptome (Oshlack et al., 2010; McIntyre et al., 2011). Many plant species have benefited from this technology, and large-scale genome sequences and transcriptome data are available in both model and non-model species (Goff et al., 2002; Yu et al., 2002; Kaplan et al., 2006; Ramsey et al., 2008; Huang et al., 2009; Wang et al., 2011; Xu et al., 2013). Cucumber, as a model plant for study of floral sex expression, its sex-controlling genes, sex-modifying plant hormones, and interactions with environmental conditions are clearly understood (Galun, 1962; Trebitsh et al., 1987, 1997; Malepszy and Niemirowicz-Szczytt, 1991; Yin and Quinn, 1995; Yamasaki et al., 2000; Kater et al., 2001; Mibus and Tatlioglu, 2004; Li et al., 2009b), and a number of candidate genes required for sex determination have been identified by transcriptome profile analysis (Guo et al., 2010; Wu et al., 2010). In *Ricinus*, however, a gap still exists between interpreting the molecular mechanism of sex expression and isolating the sex-determining genes. The complete sequencing of the castor bean genome (Chan et al., 2010) provides tremendous opportunities for genomic analysis and facilitates identification of genes of biological interest. Digital gene expression (DGE) analysis, a powerful and recently developed tool, is based on ultra-high-throughput sequencing of millions of signatures in the genome, allowing identification of specific genes and direct quantitation of transcript abundance (Bentley, 2006; Hong et al., 2011). DGE analysis detects gene expression more quantitatively than microarray assays (`t Hoen et al., 2008) while providing similar

assessments of relative transcript abundance. In comparison to high-throughput mRNA sequencing (RNA-seq), DGE more accurately detects expression differences in poorly expressed genes and exhibits much less transcript length bias (Hong et al., 2011). Thus, DGE has been widely utilized to discriminate differences in transcriptional responses among different tissues and organs in plants in the contexts of biotic and abiotic stress, metabolite biosynthesis, and developmental biology (Tian et al., 2013; Wei et al., 2013; Zhang et al., 2013; Guo et al., 2014; Yu et al., 2014; Zhao et al., 2014).

In this study, we performed DGE using the Illumina HiSeq 2000 to investigate the differences between the transcriptomes of apices and racemes from female (JXBM0705P) and monoecism (JXBM0705M) of *Ricinus*. In an attempt to get some hints associated with sex expression in *Ricinus*, we conducted (after high-throughput tag sequencing) an integrated bioinformatic analysis to identify expression patterns of genes and critical pathways in the female and monoecious lines at three stages of apex development. Assessment of the changes in gene expression between female and monoecious apices yielded sets of up-regulated and down-regulated genes associated with sex expression. Based on these differentially regulated genes (DEGs), some DEGs putatively related to sex expression were selected to confirm by quantitative RT-PCR analysis. Thus, comparison of gene expression patterns at three developmental stages of monoecious and female apices provides important hints for further insight into the molecular mechanisms underlying *Ricinus* sex variation.

MATERIALS AND METHODS

Plant Material

Seeds of nearly isogenic monoecious (JXBM0705M) and female (JXBM0705P) castor lines were kindly provided by Professor Huang (Castor Oil Research Institute of Jiaxiang, Zibo, Shandong Province). JXBM0705M is a wild monoecious line that bears male and female flowers (**Figure 1A**), and JXBM0705P, a pistillate line that bears only female flowers (**Figure 1B**), was developed by consecutively selecting spontaneous mutants of JXBM0705M. Seeds were sown and cultivated in the Yangluo experimental field of the Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, under standard field conditions, from spring to autumn of 2012. A previous study revealed that apices are leaf buds prior to the 5-leaf stage (i.e., five fully expanded leaves) and change to floral buds during the 6–9-leaf stage (Wang et al., 2012). According to this criterion, and combined with anatomical observation of a series of apical buds of the cultivar, in order to investigate global changes in the transcriptome during the development of buds from leaf apices to racemes, apical buds at the 3–4-leaf stage (ABML1 for monoecism and ABPL1 for female) and the 5–7-leaf stage (ABML2 for monoecism and ABPL2 for female) (**Figure 1C**), and small racemes 2–3 cm in size (RML for monoecism and RPL for female, **Figure 1D**), were collected in triplicate to construct 18 DGE libraries. Excised apices and infant racemes were immediately frozen in liquid nitrogen and stored at −80°C until use.
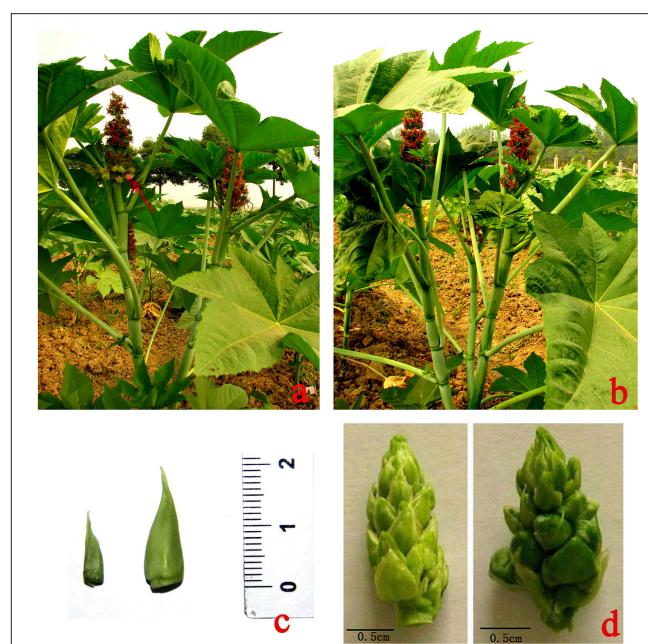
FIGURE 1 | A *Ricinus* monoecious line (JXBM0705M) and its isogenic female (JXBM0705P); sizes of collected buds and inflorescences for DEG analysis. (A) monoecious line (JXBM0705M); arrow indicates blooming male flower. **(B)** pistillate line (JXBM0705P); **(C)** collected buds at 3–4-leaf stage (left) and 5–7-leaf stage (right); **(D)** infant female (left) and monoecious (right) inflorescences, 2 cm in length. Scale bars, 0.5 cm.

RNA Preparation, Library Construction, and Sequencing

Total RNA was extracted from buds and racemes using the GeneJET Plant RNA Purification Mini Kit (Thermo Fisher Scientific, Waltham, USA). RNA was dissolved in diethylpyrocarbonate (DEPC)-treated water and stored at −70°C. The quality of total RNA was checked on 1% agarose gels, and purity was evaluated by OD260/280 ratio using a NanoDrop ND 1000 spectrophotometer (NanoDrop, Wilmington, DE, USA). Concentration and integrity of RNA were verified using an Agilent 2100 Bioanalyzer. The extracted RNA was prepared for next sequencing and quantitative RT-PCR validation.

A total of 8 μg of total RNA was subjected to oligo (dT) magnetic bead adsorption to purify mRNA. Oligo (dT) was then used as a primer for the synthesis of first- and second-strand cDNA. The 5′-ends of tags could be generated using two types of endonucleases, namely, *Nla*III or *Dpn*II. Bead-bound cDNA was subsequently digested with the restriction enzyme *Nla*III, which recognizes and cleaves CATG sites. Fragments other than 3′-cDNA fragments connected to oligo (dT) beads were washed away, and Illumina adaptor 1 was ligated to the sticky 5′-end of the digested bead-bound cDNA fragments. The junction of the Illumina adaptor 1 and the CATG site constitutes the recognition site for *Mme*I, an endonuclease with separate recognition and cleavage sites. *Mme*I cleaves 17 bp downstream of the CATG site, yielding tags with adaptor 1. After removing 3′ fragments by magnetic bead precipitation, adaptor 2 was ligated to the 3′-ends of tags, yielding tags with different adaptors at both ends

to form a tag library. After linear PCR amplification, fragments were purified by PAGE gel electrophoresis. During the quality control steps, the Agilent 2100 Bioanalyzer and ABI Step One Plus Real-Time PCR System were used for quantitation and quality verification of the sample library. Finally, the library was sequenced on an Illumina HiSeq 2000.

Data Analysis and Mapping of DGE Tags

Millions of raw 49 bp sequences were generated. Image analysis, base calling, generation of raw tags, and tag counting were performed using the Illumina pipeline (`t Hoen et al., 2008). Prior to mapping of tags to the reference database, empty tags (i.e., those with no tag sequence between the adaptors), adaptors, low-quality tags (tags containing one or more unknown nucleotides "N"), and tags with a copy number of 1 were removed from raw sequences to obtain clean tags.

To evaluate the normality of the entire DGE dataset, the distribution of clean tag expression was analyzed. The saturation analysis of the 18 libraries was also performed to estimate whether sequencing depth was sufficient for transcriptome coverage. To identify the gene expression patterns in apical buds and small racemes of two Ricinus lines, all clean tags were annotated by mapping to the castor genome (Chan et al., 2010) using the SOAP2 software (Li et al., 2009a), with a maximum of one nucleotide mismatch allowed. All tags mapped to reference sequences were filtered, and the remaining tags were designated as ambiguous tags. Mapping events on both sense and antisense sequences were included in data processing. For gene expression analysis, in order to identify genes that were differentially expressed between monoecious and female apices at the same developmental stages, as well as genes that exhibited distinctive expression between different developmental stages in the same line, we converted the number of raw clean tags in each library to TPM (number of transcripts per million tags) (`t Hoen et al., 2008; Morrissy et al., 2009) to obtain normalized gene expression levels. Then, nine pairs of DGE profiles of different sample libraries (ABML1 vs. ABPL1, ABML2 vs. ABPL2, RML vs. RPL, ABML1 vs. ABML2, ABML2 vs. RML, ABML1 vs. RML, ABPL1 vs. ABPL2, ABPL2 vs. RPL, and ABPL1 vs. RPL, where a is the control and b is the experimental group in "a vs. b") were compared to assess the diversity of gene expression. A rigorous algorithm for identifying DEGs (differentially expressed genes) between two samples, false discovery rate (FDR), was applied to determine the threshold P-value in multiple tests and analyses. Criteria for DEGs were as follows: FDR, ≤ 0.001, and $|\log_2\text{Ratio}|$, ≥ 1 (Audic and Claverie, 1997; Benjamini and Yekutieli, 2001). Following, we performed functional annotation to assign the unambiguously mapped genes identified in all libraries to Gene ontology (GO) terms. Gene annotation was conducted using Blast2GO (Conesa et al., 2005). Gene ontology (GO) was used to determine the possible functions of all DEGs by searching the GO database (http://www.geneontology.org/), and Web Gene Ontology Annotation Plot (WEGO) was also used for GO classification of genes identified in each DGE library (Ye et al., 2006). Moreover, the GO distribution of DEGs in comparisons of each pair of libraries was determined and compared. To further characterize gene function, pathway

enrichment analysis of the DGE results was conducted by BLAST search of the KEGG database (http://www.kegg.jp/kegg/). Clustering analysis of differential gene expression patterns was also performed using MultiExperiment Viewer (MeV) (Chu et al., 2008; Howe et al., 2011). A P-value of 0.05 was selected as the threshold for a gene set to be considered significantly enriched.

Quantitative RT-PCR Analysis

Quantitative RT-PCR analysis was used to verify the DGE results. The RNA samples used for the qRT-PCR assays were identical to those used for the DGE experiments. Three biological replicates and two technical replicates were performed for each sample. The first-strand cDNA was synthesized from the total RNA (1 ug) from each sample using the PrimeScript RT reagent kit with gDNA Eraser (perfect Real Time) (TaKaRa, RR047A). The primers designed for qPCR analysis was listed in Additional File S1. *Actin* gene (forward primer: 5′-TGCTGACAGAATGAGCAAGG-3′; reverse primer: 5′-AATCCACATCTGCTGGAAGG-3′) was used as an internal control gene for normalization. Quantitative PCR reactions used ABI7500 quantitative PCR, which was performed using SYBR® Premix Ex Taq™ II(Tli RNaseH Plus) (TaKaRa, RR820A) kit according to the manufacturer's instructions. Reaction conditions were 95°C for 5 s in order to activate enzyme reaction. Two step cycles were then used: 95°C for 5 s, then 60°C for 30 s, 40 cycles; solubility curve conditions were 95°C for 15 s, 60°C for 1 min, 95°C for 30 s, 60°C for 15 s. The specificity of the SYBR green PCR signal was further confirmed by melting curve analysis and agarose gel electrophoresis, and the relative expression levels of genes were calculated with the $2^{-\triangle\triangle Ct}$ method.

Hormone Measurements

Contents of several phytohormones including auxin or indole-3-acetic acid (IAA, the major form of auxin in plants), abscisic acid (ABA), jasmonic acid (JA), and gibberellins (GAs) were analyzed to compare and verify the physiological differences between the monoecism and the pistillate. The samples used for the hormone measurement were the same as those used for the DGE experiments (Three development stages of apices from leaf bud to raceme for the monoecism and female, respectively; ABML1, ABML2, and RML for monoecism, ABPL1, ABPL2, and RPL for female.) Three biological replicates were designed for each sample. Quantification of endogenous IAA, ABA, JA, and GAs (GA$_1$, GA$_3$, GA$_4$, and GA$_7$) were performed as described previously (Chen et al., 2012). The data of detection was analyzed using Microsoft Excel 2010 and SAS V8 softwares.

RESULTS

Analysis of DGE Libraries and Tag Mapping

The 18 DGE libraries were sequenced and generated approximately 6 million raw tags for each library, and more than 96% of the raw tags were clean (**Table 1**). There produced approximately 5.7 million clean tags per library, of which 148,526–174,340 were distinct. The number of unique distinct clean tag sequences ranged from 134,541 to 159,229 (**Table 1**).

TABLE 1 | Categorization and abundance of tags.

Summary		ABML1	ABPL1	ABML2	ABPL2	RML	RPL
Raw data	Total	5,919,796±211,429	5,888,377±122,404	6,004,837±92,543	6,000,410±135,373	5,939,917±123,287	6,045,455±233,096
	Distinct tags	371,983±4765	374,582±40,008	391,584±12,149	440,819±11,500	369,807±8508	447,852±47,302
Clean tags	Total number	5,697,507±209,466	5,661,387±132,174	5,767,119±88,697	5,729,206±121,449	5,722,947±123,544	5,771,163±203,535
	Distinct tag numbers	150,905±2990	148,526±12,470	155,040±6806	170,590±3920	153,959±4430	174,340±13,817
All tags mapping to genes	Total number	5,393,163±203,811	5,397,141±104,629	5,476,967±99,781	5,477,295±134,778	5,490,070±156,338	5,516,938±151,205
	Total % of clean tags	94.66±0.83	95.35±1.85	94.98±1.72	95.6±0.58	95.92±0.82	95.62±1.79
	Distinct tag numbers	135,847±3278	134,541±8871	139,351±5476	154,604±2420	141,188±2952	159,229±9689
	Distinct tags (% of clean tags)	90.02±0.98	90.69±1.97	89.91±1.71	90.65±1.4	91.72±1.12	91.42±1.89
Unambiguous tag mapping to gene	Total number	168,945±7762	162,305±15,683	163,621±16,927	169,228±5762	172,373±6399	173,617±11,236
	Total % of clean tags	2.96±0.07	2.86±0.21	2.83±0.26	2.96±0.16	3.01±0.09	3.01±0.1
	Distinct tag numbers	7665±306	7323±585	8164±599	8722±416	8223±295	9013±514
	Distinct tags (% of clean tags)	5.08±0.15	4.94±0.31	5.26±0.2	5.11±0.15	5.34±0.04	5.18±0.13
All tag-mapped genes	Number	50,722±577	38,867±1085	45,392±4807	54,090±1159	44,709±1911	41,658±629
	% of ref genes	41.99±0.48	32.17±0.9	37.57±3.98	44.78±0.96	37.01±1.59	34.48±0.52
Unambiguous tag-mapped genes	Number	4867±191	4582±227	5050±333	5215±197	5144±137	5292±144
	% of ref genes	4.03±0.16	3.79±0.19	4.18±0.28	4.32±0.16	4.26±0.12	4.38±0.12
Mapping to genome	Total number	31,478±3790	30,864±8907	33,146±7186	27,758±2213	26,842±2537	30,301±6426
	Total % of clean tags	0.55±0.07	0.54±0.15	0.57±0.12	0.48±0.04	0.47±0.05	0.52±0.1
	Distinct tag numbers	1494±115	1607±173	1879±390	1715±412	1230±113	1548±187
	Distinct tags (% of clean tags)	0.99±0.07	1.08±0.12	1.21±0.2	1±0.22	0.8±0.05	0.89±0.05
Unknown tags	Total number	272,866±43,697	233,382±100,607	257,006±93,700	224,153±28,803	206,034±41,217	223,923±104,814
	Total % of clean tags	4.79±0.76	4.11±1.7	4.45±1.59	3.92±0.55	3.61±0.77	3.85±1.7
	Distinct tag numbers	13,565±1338	12,378±3813	13,810±2853	14,271±2285	11,541±1925	13,563±4350
	Distinct tags (% of clean tags)	8.99±0.91	8.23±1.98	8.89±1.67	8.35±1.18	7.48±1.08	7.69±1.85

The RPL libraries and ABPL2 libraries contained the two highest numbers of distinct clean tags; the other libraries had similar numbers.

Our results showed that mRNAs transcribed from major genes were often present at fewer than 10 copies, and only a small proportion of genes were expressed at higher levels (**Figure 2**; Additional Files S2, S3). The distributions of total and distinct clean tag copy numbers had highly similar properties in libraries from all six sample types, with most tags coming from highly expressed genes (**Figure 2**). Among the distinct clean tags, more than 56% were present at 2–5 copies, 33.81–36.84% were present at 5–100 copies, and fewer than 6.27% had copy numbers higher than 100. However, more than 70% of total clean tags had counts above 100 in each library, indicating that the overall DGE data among the 18 libraries was normally distributed.

A reference gene database that included 120,799 *R. communis* Unigene sequences was preprocessed for tag mapping. Among the sequences, genes with a CATG site accounted for 90.03% (Additional File S3). Distinct clean tags (4.94–5.34%) were mapped unambiguously in the unigene database, whereas 7.48–8.99% of distinct clean tags were not mapped to the unigene virtual tag database, in libraries from the six sample types (**Table 1**). Around 95% of total clean tags were mapped onto the *R. communis* genome with a perfect match or 1 bp mismatch to sense or antisense genes, and approximately 90% of distinct clean tags were successfully mapped. Ultimately, tag mapping onto the *R. communis* genome generated 50,722 tag-mapped genes for ABML1, 38,867 for ABPL1, 45,392 for ABML2, 54,090 for ABPL2, 44,709 for RML, and 41,658 for RPL (**Table 1**; Additional File S4).

In this research, the number of detected genes increased with the amount of sequencing amount until the number of tags reached 3 million or higher (Additional File S5), 20.61–25.73% of distinct clean tags perfectly matched antisense transcripts, and 43.14–53.23% of distinct clean tags perfectly matched sense strand-specific transcripts (Additional Files S3, S4). In total,

more than 88% of genes were transcribed from both strands, indicating the importance of RNA-mediated gene regulation in bud transformation from leaf to inflorescence.

Global Gene Expression of Buds at Different Stages and Analysis of Differential Gene Expression

In this study, we identified 5526 and 5203 genes in the transcriptomes of monoecious and pistillate apical buds, respectively, at the 3–4-leaf stage; 5779 and 6118 genes at the 5–7-leaf stage; and 5870 and 5988 genes in racemes (**Figure 3A**).

Our results showed that whether monoecism or female, the transcriptome of their terminal buds during the three developmental stages was dynamic. In total, 7495 and 7610 genes were expressed in monoecious and pistillate apical buds over the three stages, respectively. Of which, 4066 and 4013 were constitutive, 1881 and 1924 were specific to a single stage, and 1548 and 1637 were expressed at two stages, correspondingly for monoecism and female (**Figures 3B,C**; Additional File S6). The complex changes of gene expression indicated that the transition from leaf apical bud to raceme is an involute process in castor bean. More genes were expressed in the buds closer to raceme formation and incipient racemes than in early apical buds (**Figure 3**), and more stage-specific transcripts were detected in the 5–7-leaf stage bud and the small raceme than in the early leaf bud (**Figure 3D**; Additional File S6). These observations suggest that identical molecular pathways are involved in apical bud development from vegetation to reproduction in the monoecism and female, the developmental process may require expression of a larger number of unique genes involved in regulatory processes and related pathways.

Comparison of the monoecious line with the female line revealed that 1386 DEGs were up-regulated and 1946 DEGs were down-regulated at the 3–4-leaf stage, but the majority of genes were up-regulated during the subsequent two stages: 1952 up-regulated and 1408 down-regulated genes at the 5–7 leaf stage, and 1667 up-regulated and 1451 down-regulated genes at the initial raceme stage (**Table 2**; **Figures 4A–C**; Additional File S7). The number of up-regulated DEGs in the female was greater than the number of down-regulated DEGs during the developmental transition from leaf bud to infant raceme, which might indicate that females require a number of genes involved in silencing of male flower expression. Similarly, it was showed that up-regulated genes outnumbered down-regulated genes in these comparisons between different development stages, except for the RPL vs. ABPL2 (**Table 2**; **Figures 4D–I**; Additional File S7), which also revealed that the expression of many genes tended to increase; a few genes were activated, whereas others were suppressed, during bud development from vegetation to reproduction.

Functional Annotation and DEG Clustering Analysis

Go analysis revealed that these well-annotated sequences belonged to three main categories (cellular component, molecular function, and biological process) and mainly
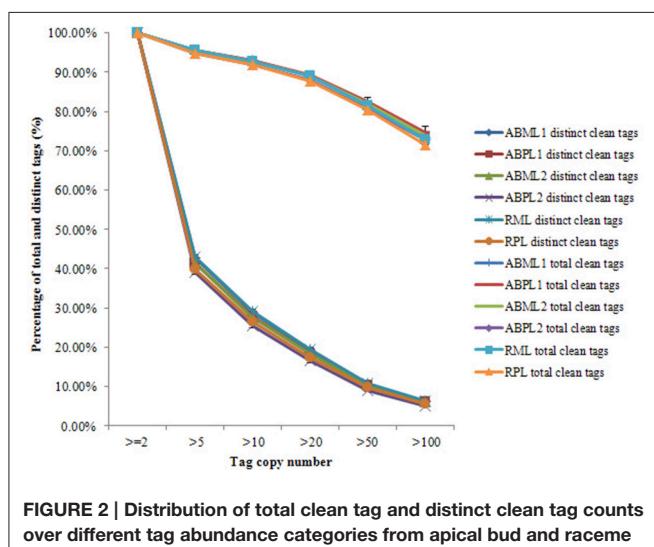
FIGURE 2 | Distribution of total clean tag and distinct clean tag counts over different tag abundance categories from apical bud and raceme libraries from the monoecious and pistillate lines.
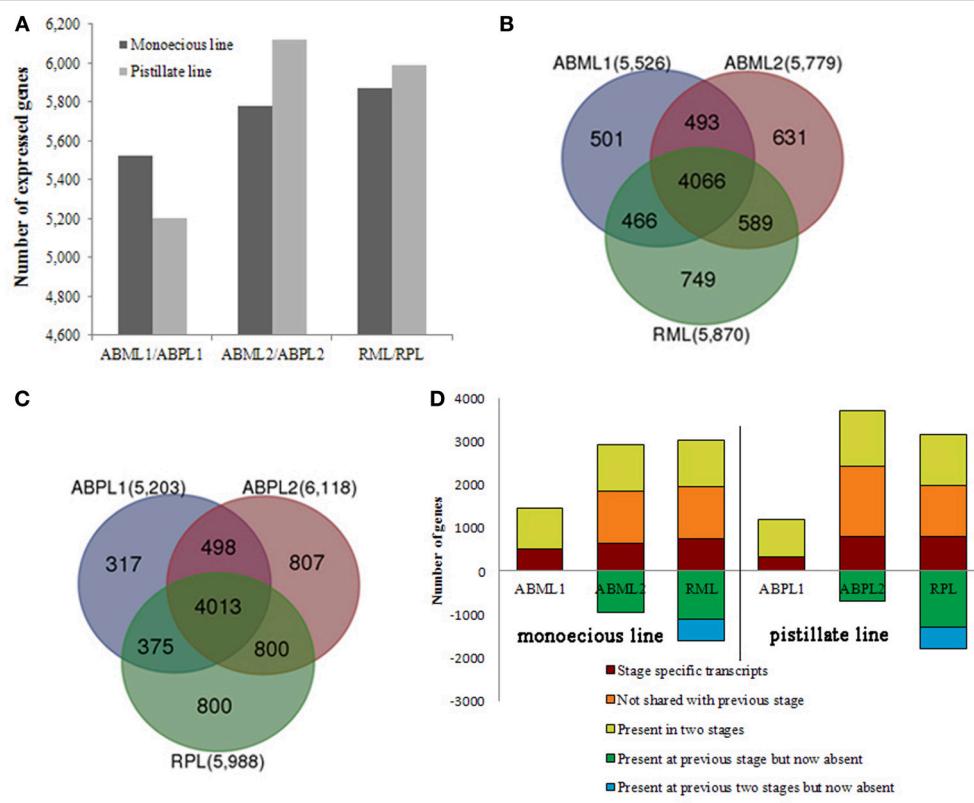
FIGURE 3 | Transcriptome analysis of apical buds and racemes of the monoecious line (ML) and pistillate line (PL). (A) Transcriptome sizes of the monoecious and pistillate lines at three developmental stages. **(B,C)** Venn diagram showing the overlaps between three stages (two apical bud stages and raceme stage) of ML and PL. The number in parentheses after each stage designation is the total transcripts detected in that stage(s). **(D)** Analysis of transcriptome changes from apical buds to raceme development of the ML and PL. Transcripts shared by three stages of ML and PL are not shown; numbers above the x-axis represent transcripts present in the indicated stage that are stage-specific (dark red), not shared with the previous stage (orange), or shared with the prior stage but missing in at least one other stage (yellow-green). Numbers below the x-axis represent transcripts present in the prior stage that were not detected in the current stage (green and blue).

TABLE 2 | Gene expression levels across different sample libraries.

	ABML1/ ABML2	ABML2/ RML	ABML1/ RML	ABPL1/ ABPL2	ABPL2/ RPL	ABPL1/ RPL	ABML1/ ABPL1	ABML2/ ABPL2	RML/ RPL
Significantly up-regulated	17	48	68	32	75	103	2	15	6
Up-regulated	1873	1991	2083	2613	2055	2535	1384	1937	1661
Not DEGs	3372	3184	3207	3043	2940	2863	3089	3664	3855
Down-regulated	1471	1757	1494	1118	2190	1293	1929	1401	1443
Significantly down-regulated	13	14	12	4	33	9	17	7	8
Total expression genes	6746	6994	6864	6810	7293	6803	6421	7024	6973

Significantly up-regulated: log_2Ratio, ≥ 1, and probability, ≥ 0.8; Up-regulated: log_2Ratio, ≥ 1, and probability, <0.8; Not DEGs: genes not differentially expressed ($0 \leq |log_2Ratio| < 1$); Down-regulated: log_2Ratio, ≤ -1, and probability, <0.8; Significantly down-regulated: log_2Ratio, ≤ -1, and probability, ≥ 0.8.

distributed into 43 categories, including the most predominant pathways such as "Intracellular membrane-bounded organelle," "Adenyl ribonucleotide binding," "Transition metal ion binding," "Intrinsic to membrane," "Cellular protein modification process," "Plastid," "Membrane and cell part," "Binding and protein binding," "Metabolic process," and "Transcription, DNA-dependent." The distributions among the pairwise comparisons were similar, except for small differences in the numbers of

genes in each category and the total numbers of main category (Additional Files S8, S9). In addition, all of the annotated genes were mapped to terms in the KEGG database to search for significantly enriched genes involved in spliceosome, cell cycle, homologous recombination, or signaling pathways (Additional File S10).

In general, the 7687 DEGs in all comparisons were clustered as the union of DEGs. A total of 617 transcripts occurring
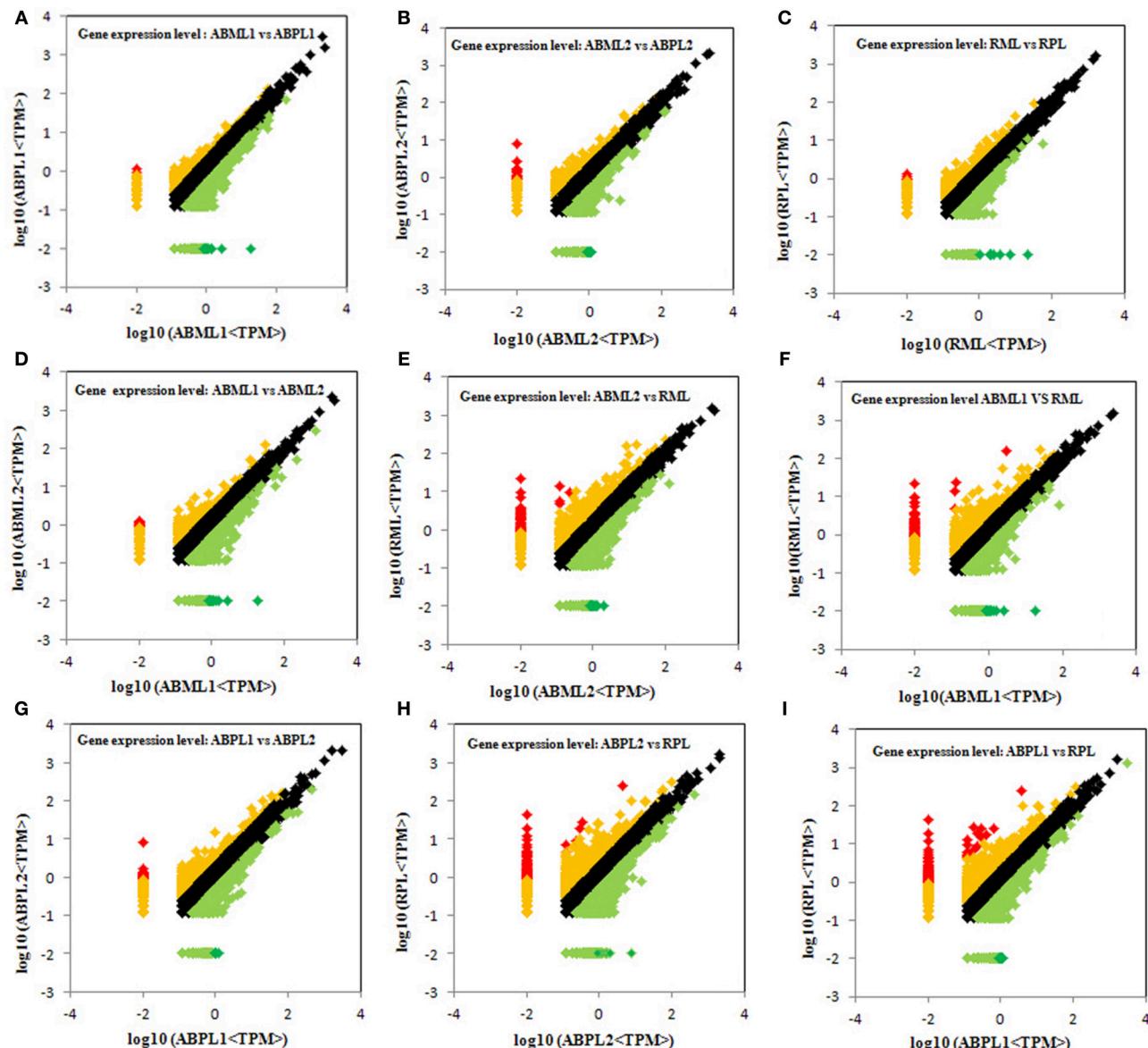
FIGURE 4 | Comparison of gene expression levels across all libraries. All genes mapped to the reference sequence were examined for differential expression across the libraries. ABML1 and ABPL1: apical bud of 3–4-leaf stage of monoecious and female lines, respectively; ABML2 and ABPL2: apices of 5–7-leaf stage of monoecious and pistillate lines; RML and RPL: small raceme (2–3 cm long) of monoecious and female lines. Red prism represented significantly up-regulated genes (log2Ratio, ≥ 1; probability, ≥ 0.8), yellow represented up-regulated genes (log2Ratio, ≥ 1; probability, <0.8), black showed not DEGs (log2Ratio, ≥ 0 and <1), light green indicated down-regulated genes (log2Ratio, ≤ -1; probability, <0.8) and green prism demonstrated significantly down-regulated genes (log2Ratio, ≤ -1; probability, ≥ 0.8).

simultaneously in the three comparisons of three development stages between the monoecism and female were identified, and the comparison of the ratio values of these genes were used for clustering as the intersections of DEGs (**Figure 5**). Among the nine major clusters, 62 genes grouped into Cluster A were down-regulated in the three comparisons of developmental stages between the monoecism and the female, whereas the 97 genes in Cluster D were up-regulated in the three comparisons. GO analysis of these clustered genes indicated that the known genes were mainly involved in membrane or

vesicle structures of intracellular or cytoplasmic components, activity and binding functions, response to light or hormone stimulus, transport, signal transduction, metabolic processes, and developmental processes involved in differentiation and reproduction (Additional File S11).

A heat map was generated of DEGs between different growth stages for both the monoecious and female buds (**Figure 6**). There were 621 DEGs in the four comparisons that clustered as the intersections of DEGs (Additional File S12). As shown in **Figure 6**, the DEGs in "ABML1 vs. ABML2" and "ABPL1
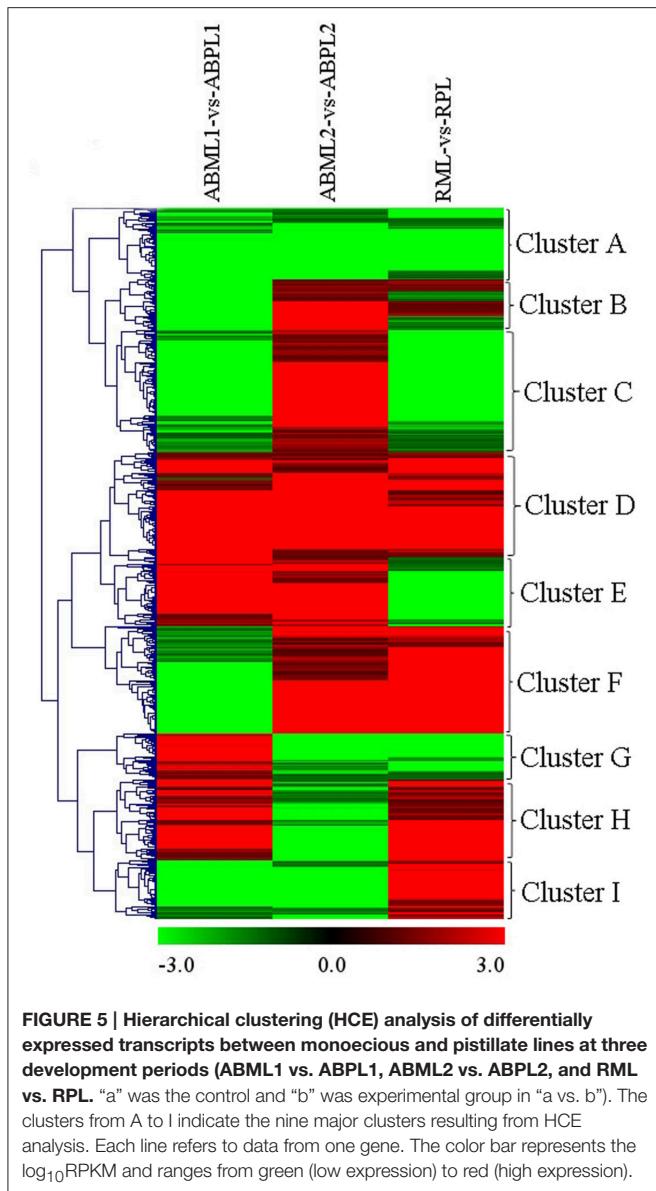
FIGURE 5 | Hierarchical clustering (HCE) analysis of differentially expressed transcripts between monoecious and pistillate lines at three development periods (ABML1 vs. ABPL1, ABML2 vs. ABPL2, and RML vs. RPL. "a" was the control and "b" was experimental group in "a vs. b"). The clusters from A to I indicate the nine major clusters resulting from HCE analysis. Each line refers to data from one gene. The color bar represents the $\log_{10}$RPKM and ranges from green (low expression) to red (high expression).

FIGURE 6 | Hierarchical cluster analysis of DEGs between different developmental stages for the monoecious and female lines (ABML1 vs. ABML2, ABML2 vs. RML, ABPL1 vs. ABPL2, and ABPL2 vs. RPL. "a" was the control and "b" was experimental group in "a vs. b"). Each line refers to data from one gene. The color bar represents the $\log_{10}$RPKM and ranges from green (low expression) to red (high expression). The clusters from A to H indicate the eight major clusters resulting from HCE analysis.

vs. ABPL2" were closely correlated to the DEGs in "ABML2 vs. RML" and "ABPL2 vs. RPL," respectively, indicating that gene expression differences between monoecious and female lines were not obvious during the same development processes. However, this analysis revealed a greater number of DEGs between different development processes, with a relatively distant relationship, which suggests that transformation of apical bud from leaf to flower is a complex process.

Genes Associated with Sex Expression and Reproduction in Castor Bean

Based on prior knowledge of the putative involvement in sex determination and the expression levels and functions of DEGs between the monoecious and female buds in three development stages, several subgroups were assumed to be putatively related

to sex expression and reproduction (**Table 3**). We identified seven genes involved in response to hormone stimulus, whose expression changed significantly in mutant females over the three developmental stages (**Table 3**). These genes included those that encode dynamin-2A, auxin response factor, PCI domain-containing protein, ATP-binding protein, spermidine synthase, Xaa-Pro amino peptidase, and a conserved hypothetical protein. These genes are associated with tissue or organ developmental processes, and several of them take part in signal transduction, auxin transport, phytohormone biosynthesis, and metabolism (such as polyamine and abscisic acid).

In this study, we detected eight transcription factors that exhibited significantly different expression patterns between the pistillate and monoecious lines during the three developmental stages. These transcriptional regulators were identified as DNA-dependent transcription factors, including MADS box proteins, DNA-binding proteins, the Axial regulator YABBY5, transcription initiation factors, auxin response factor, and

TABLE 3 | Some selected differentially expressed genes detected by digital expression profiling in apical bud of castor bean.

Functional group	Gene ID	Gene annotation	$\log_2$ratio		
			ABML1/ABPL1	ABML2/ABPL2	RML/RPL
RESPONSE TO HORMONE STIMULUS					
	gi|255573875|ref|XM_002527811.1|	Dynamin-2A	−8.05	−3.54	−8.54
	gi|255538885|ref|XM_002510462.1|	Auxin response factor (transcription DNA-dependent)	−3.58	1.00	−3.66
	gi|255546086|ref|XM_002514057.1|	PCI domain-containing protein	−3.66	3.58	4.84
	gi|255543197|ref|XM_002512616.1|	ATP-binding protein	1.15	−6.23	4.84
	gi|255550142|ref|XM_002516076.1|	Spermidine synthase (auxin transport, polyamine biosynthetic process)	−5.71	1.36	−1.54
	gi|255568254|ref|XM_002525056.1|	Xaa-pro amino peptidase (auxin transport)	−1.77	1.50	−1.78
	gi|255560413|ref|XM_002521176.1|	conserved hypothetical protein (ABA and carotenoid metabolic process)	−4.56	4.56	3.58
TRANSCRIPTION, DNA-DEPENDENT					
	gi|255558565|ref|XM_002520262.1|	MADS box protein	−2.52	−4.90	−2.80
	gi|255583281|ref|XM_002532359.1|	DNA-binding protein	3.58	3.66	3.58
	gi|255572532|ref|XM_002527155.1|	Axial regulator YABBY5	−1.18	8.07	5.69
	gi|255538455|ref|XM_002510247.1|	DNA-binding protein	−2.08	1.31	4.14
	gi|255558973|ref|XM_002520464.1|	transcription initiation factor	3.58	−3.58	3.58
	gi|255565967|ref|XM_002523926.1|	RNA polymerase sigma factor rpoD1	−5.14	−4.54	4.14
	gi|255558989|ref|XM_002520472.1|	protein with unknown function	−1.15	−1.80	3.58
	gi|255538885|ref|XM_002510462.1|	Auxin response factor	−3.58	1.00	−3.66
SIGNAL TRANSDUCTION					
	gi|255555566|ref|XM_002518774.1|	gcn4-complementing protein	−5.16	−1.07	−4.56
	gi|255550049|ref|XM_002516030.1|	protein with unknown function	−4.92	3.62	−1.57
	gi|255585732|ref|XM_002533502.1|	Histidine-containing phosphotransfer protein	−5.39	4.87	−3.62
	gi|255584316|ref|XM_002532848.1|	Cyclic nucleotide-gated ion channel	−3.58	3.66	−4.14
SEX DIFFERENTIATION					
	gi|255557726|ref|XM_002519847.1|	arginine/serine-rich splicing factor	−4.56	3.66	3.58
POLLINATION					
	gi|255560126|ref|XM_002521035.1|	acid phosphatase	−2.42	1.27	1.90
PROGRAMMED CELL DEATH					
	gi|255539309|ref|XM_002510674.1|	*Ricinus communis* cysteine protease	−4.22	3.66	−4.54
DEVELOPMENTAL PROCESS INVOLVED IN REPRODUCTION					
	gi|255577146|ref|XM_002529411.1|	1,4-alpha-glucan branching enzyme	−1.02	4.20	−3.62
	gi|255588742|ref|XM_002534658.1|	DNA replication helicase dna2	−1.45	5.71	−1.96
	gi|255568769|ref|XM_002525310.1|	sentrin/sumo-specific protease	−1.66	2.32	1.29
	gi|255578470|ref|XM_002530054.1|	U4/U6 small nuclear ribonucleoprotein Prp4	−6.24	4.20	4.22
	gi|255544821|ref|XM_002513426.1|	sorting and assembly machinery (sam50) protein	−3.66	4.20	4.14
HISTONE DEMETHYLATION/METHYLATION					
	gi|255565532|ref|XM_002523711.1|	eukaryotic translation initiation factor 2c	−3.58	1.29	1.67
	gi|255575536|ref|XM_002528623.1|	set domain protein	2.78	−4.20	−4.14
	gi|255553976|ref|XM_002517983.1|	DNA (cytosine-5)-methyltransferase	−4.89	3.58	−3.62
	gi|255558643|ref|XM_002520301.1|	s-adenosyl-methyltransferase mraw	1.23	3.66	5.33

the RNA polymerase sigma factor rpoDI. Auxin response factor is respond to hormone stimulus and participates in hormone-mediated signaling pathways, also plays a role in tissue and organ developmental processes. Furthermore, one subgroup is composed of four genes related to signal transduction, including Gcn4-complementing protein, histidine-containing phosphotransfer protein, an unknown protein, and cyclic nucleotide-gated ion channel, which exhibited consistent expression changes: down-regulation at the early leaf-bud stage, up-regulation at the 5–7-leaf stage, and down-regulation at the young raceme stage. Another two genes listed above (PCI domain-containing protein and ATP-binding protein) play roles in signal transduction in addition to response to hormone stimulus.

One gene (*arginine/serine-rich splicing factor*) associated with sex differentiation and another gene (*acid phosphatase*) related to pollination were identified as differentially expressed between monoecious and pistillate apical bud development. In female apical buds, both genes were distinctly down-regulated at the early leaf-bud stage, but up-regulated at the 5–7-leaf and initial raceme stages. In addition, expression of a PCD (programmed cell death)-related gene that encodes cysteine proteinase was altered in pistillate apices, with down-regulation at the early leaf-bud and initial raceme stages and up-regulation during the period of transition from leaf bud to floral bud. Several genes involved in reproduction exhibited the same expression patterns in pistillate terminal bud: sentrin/sumo-specific protease, the U4/U6 small nuclear ribonucleoprotein Prp4, sorting and the assembly machinery (sam50) protein were down-regulated at the early leaf stage, but up-regulated at the 5–7-leaf and small raceme stages, (**Table 3**); notable exceptions were 1,4-alpha-glucan branching enzyme and DNA replication helicase (*dna2*). Histone modification and DNA methylation or demethylation, play an important role in the regulation of gene expression or silencing during developmental processes. We identified four genes involved in histone methylation and demethylation (**Table 3**), including eukaryotic translation initiation factor 2c, set domain protein, DNA (cytosine-5)-methyltransferase, and s-adenosyl-methyltransferase (*mraw*). These genes were also significantly differentially expressed between the monoecious and female apical bud development.

DEGs Were Confirmed by Quantitative RT-PCR

To confirm the differentially expressed genes in the monoecious and female apical buds, 14 genes were selected for quantitative RT-PCR analysis at the three apical bud developmental stages. Representative genes selected for qPCR validation were those involved in response to hormone stimulus, transcription factor, signal transduction, sex differentiation, pollination, reproduction and histone demethylation/methylation because those phenomena were putative to associate with plant sex determination(Guo et al., 2010; Wu et al., 2010). The expression of the 13 genes (Dynamin-2A, Auxin response factors, ATP-binding protein, Spermidine synthase, auxin transport, conserved hypothetical protein, MADS box protein, two unknown proteins, arginine/serine-rich splicing factor, acid phosphatase, DNA replication helicase dna2, and eukaryotic translation initiation factor 2c) indicated by qRT-PCR was basically congruent with the DGE analysis patterns (**Figure 7**). Only one gene (s-adenosyl-methyltransferase mraw) did not show consistent expression between the qRT-PCR and DGE data sets.

Differential Hormone Levels in the Apices and Racemes Between the Monoecism and Female

DGE analysis and qRT-PCR verification showed that some genes which response to hormone stimulus were expressed differentially in apices and racemes between the monoecism and

female, and possibly associated with castor sex expression. To validate whether the levels of phytohormone in the apices and inflorescences of the monoecious line and the pistillate line vary or not, the IAA, ABA, JA, and GAs contents of the two lines at three developments were quantitatively measured. The result showed that the auxin in the tow castor lines indicated a similar variation pattern in that IAA was equivalent or increased from the 3–4-leaf stage to the 5–7-leaf stage, and later significantly reduced ($P < 0.01$) (**Figure 8A**; Additional File S13). It was worthwhile to note that the auxin level in pistillate line was remarkably higher than that in the monoecism at each of the three development stages. Four kinds of GAs were performed in the measurement, but only GA_4 were detected with a relatively low level, and other three GA were not detected. The GA_4 level in the apices except little racemes also showed significant differences between the monoecism and female (**Figure 8A**; Additional File S13). The level of GA_4 in the female was significantly lower ($P < 0.01$) than that in the monoecism at the stage of early leaf bud, but markedly higher ($P < 0.01$) at the 5–7-leaf stage and similar at the stage of young inflorescence.

The ABA levels in the monoecism and female revealed an identical change pattern from early leaf apices to infant racemes, with a notable consecutive reduction (**Figure 8B**; Additional File S13). The ABA level in female was significantly higher ($P < 0.01$) than that in monoecism at early leaf apices, but notable lower ($P < 0.01$) at following 5–7-leaf stage, and comparative level at little raceme stage. In addition, the variation in endogenous JA levels showed a similar pattern as IAA in the two castor lines during the three development stages (**Figure 8B**; Additional File S13). However, opposite to IAA, the JA level in the pistillate line was extremely significant lower ($P < 0.01$) than that in the monoecism at the two apical development stages, and significantly higher ($P < 0.01$) than that in monoecism at the stage of little raceme.

DISCUSSION

Sex expression in plants is controlled by genetic factors and non-genetic fluctuations (Shifriss, 1956; Ainsworth, 1999), and phytohormones play important roles in sex determination in some species (Yin and Quinn, 1995; Ainsworth, 1999; Tanurdzic and Banks, 2004). In cucumber, ethylene, a true sex hormone, was proven to be a "female" hormone that exerts a strong feminizing effect (Dellaporta and Calderon-Urrea, 1993); Auxin also promotes cucumber femaleness, potentially playing an indirect role in sex determination by promoting the action of ethylene (Yamasaki et al., 2000). The level of gibberellic acid (GA), which acts downstream of ethylene or through a parallel pathway (Ainsworth, 1999), is correlated with male tendency. Sex expression in castor bean is also regulated by plant hormones (Shifriss, 1961; Tan et al., 2011), but in contrast to cucumber, ethylene and ethylene-like substances (NIA 10637) result in masculinization and can transform female flowers into male ones in monoecious plants (Philipos and Narayanaswamy, 1976); moreover, GA causes feminization in castor, and spraying monoecious inbreds with GA can markedly increase female tendency (Shifriss, 1961). In addition, previous work showed

FIGURE 7 | Quantitative RT-PCR validation of differential expressed genes. TPM, Transcription per million mapped reads. ABML1 and ABPL1, apical buds at 3–4-leaf stage of monoecious and pistillate lines; ABML2 and ABPL2, apical buds at 5–7-leaf stage of monoecious and pistillate lines; RML and RPL, infant raceme stage of monoecious and female lines. Relative expression levels were calculated using Actin as an internal control. Error bars indicate SE ($n = 3$).

that the activity of auxin-like substances causes feminization of castor bean, a process influenced by kinetin and morphactin (Kumar and Rao, 1981). Morphactin causes maleness in *R. communis* (Rkey, 1978; Varkey and Nigam, 1982), whereas daminozide promotes femaleness with reduction in the position of bearing the first female flower and the ratio of male and female flowers (Chauhan et al., 1987). To date, however, the key sex hormone in castor bean remains unknown. In this

FIGURE 8 | Levels of IAA, GAs, ABA, and JA hormones in the apices and racemes of the monoecious line and pistillate line. ML, monoecious line; PL, pistillate line; ABML1 and ABPL1, apical buds at 3–4-leaf stage of monoecious and pistillate lines; ABML2 and ABPL2, apical buds at 5–7-leaf stage of monoecious and pistillate lines; RML and RPL, infant raceme stage of monoecious and female lines; F.W., fresh weight. **(A)** IAA and GA_4 contents of apices and racemes for ML and PL; **(B)** JA and ABA levels in apices and racemes of ML and PL. Error bars indicate SE ($n = 3$); Asterisks indicate $p < 0.05$ (*) and $p < 0.01$ (**) between PL and ML in Duncan's test analysis.

study, we identified several genes involved in the response to hormone stimulus (**Table 3**; **Figure 7**), including auxin transport and auxin response factor genes and other genes involved in plant hormone-mediated biosynthetic and metabolic process, such as polyamine biosynthesis and abscisic acid and carotenoid metabolism. In the pistillate line, three genes (*auxin response factor*, *spermidine synthase*, and *Xaa-Pro amino peptidase*) related to auxin response and transport, were down-regulated in early leaf apices and young racemes, but up-regulated during the transformation from leaf apices to floral buds. In addition, the result of hormone measurement detected that auxin level in the female was higher than that in the monoecism at each of the three development stages and extremely significant higher at the 5–7-leaf stage. The levels of JA, ABA, and GA in the apices also displayed significant differences between the pistillate line and the monoecious line, except the GA and ABA level in the infant racemes (**Figure 8**). Our results suggested that DEGs involved in the pathways of plant hormone response and hormone-mediated biosynthetic and metabolic processes, may participate in or affect the regulation of sex expression and reproduction by changing the hormone level and transport, which will be elucidated in further research. Our result was consistent with previous reports (Heslop-Harrison, 1957; Rkey, 1978). Those authors showed that variation in sex expression in plants is associated with variations in the hormone level (Heslop-Harrison, 1957), and that auxins promote femaleness (Wittwer and Hillyer, 1954). Interference in auxin transport, or quick degradation of the hormone, can decrease the auxin level, resulting in altered sex expression (Thomson and Leopold, 1974; Gaganias and Berg, 1977). The fact that lateral branches of castor bean plants produce only male flowers after decapitation supports the view that inhibition of auxin transport can result in maleness (Rao, 1969). Furthermore, our previous research (Wang et al., 2012) demonstrated that IAA and ABA levels in the inflorescence and flower vary remarkably between the pistillate and the monoecism, also suggesting that some phytohormones such as IAA etc. and related genes are important for sex determination in *R. communis*. In plants,

phytohormones modulate various growth and developmental events through signal transduction pathways (Bleecker and Kende, 2000; Davies, 2004). Notably, in this study, three genes involved in signal transduction exhibited expression patterns similar to those of genes involved in auxin response and transport (i.e., down-regulated in early leaf apices and small racemes, but up-regulated during the transformation stage from leaf to floral bud) in the pistillate line. However, other genes involved in signal transduction were down-regulated throughout all three stages. Subsequent investigations of these genes and crosstalk between other hormones and auxin may reveal their specific roles in sex expression or sex determination.

Castor is a monoecious species, with pistillate flowers borne on short pedicels along the upper portion of the raceme and staminate flowers borne similarly on the lower portion (Brigham, 1967). Castor flowers are unisexual and lack petals. Male and female flowers differ radically in general morphology and size: male flowers are devoid of rudiments of organs of the opposite sex, whereas pistillate flowers occasionally exhibit vestiges of stamen (Wu et al., 1998). The occasional appearance of rudimentary androecium in *Ricinus* female flowers suggests that like unisexual flowers of other species (Atsmon and Galun, 1960; Cheng et al., 1983; Bracale et al., 1991; Malepszy and Niemirowicz-Szczytt, 1991; Veit et al., 1993), castor flowers (especially pistillate flowers) pass through a "bisexual stage" during which development of all floral organs is initiated. Castor sex determination probably involves PCD of opposite sex organs in the flower and inflorescence, similar to what is observed in several other plant species such as maize, cucumber, and campion (Ye et al., 1991; Dellaporta and Calderon-Urrea, 1993; Kater et al., 2001); these plants follow a sex determination pathway involving arrest of preformed sexual organs in bisexual primordia in which some PCD-related genes participate (DeLong et al., 1993). In this study, we detected a PCD-related gene (cysteine protease) that was significantly up-regulated in female apices at the 5–7-leaf stage, but obviously down-regulated at the two other stages (early leaf apices and initial inflorescence). This

result was consistent with recent findings that PCD-related genes, including cysteine protease, are more highly expressed at the peak of anther abortion in CMS and GMS lines of cotton (Lorrain et al., 2003; Wei et al., 2013). It is possible that the cysteine protease might play a role in *Ricinus* apical development, and thus be associated directly or indirectly with sex expression. Our study identified hundreds of genes potentially involved in sex determination, including one (*arginine/serine-rich splicing factor*) associated with sex differentiation; this gene was up-regulated during apical transformation and initial raceme formation in female apical buds.

Transcription factors such as the maize DELLA protein D8 and the melon zinc finger protein CmWIP1 have been functionally associated with the sex determination process (Peng et al., 1999; Martin et al., 2009). Moreover, several transcription factors were identified as differentially expressed during sex determination in cucumber (Guo et al., 2010; Wu et al., 2010), including zinc finger transcription factors, Aux/IAA transcription factor, auxin-induced protein, BRI1, BRI1-associated receptor kinase, MYC transcription factor, BEL1-like homeodomain protein, bHLH proteins, WRKY DNA-binding protein, and NAC domain protein, *IF-2*, and so on. Consistent with the results in cucumber, we identified eight DNA-dependent transcription genes in castor bean that were significantly differentially expressed between the three apical developmental stages in the pistillate line. Of these, auxin response factor, MADS box protein, *IF-2*, and two DNA-binding proteins were also confirmed to be differentially expressed during sex determination. These transcription factors detected in this study were mainly binding proteins, response factors, initiation and regulation factors, and even MADS box proteins; the relationship between these transcription factors and plant sex determination remains to be determined, and should be investigated in future work. In addition, we also identified a few genes involved in histone demethylation/methylation and reproductive processes. DNA methylation and histone modification play essential roles in genome management and control gene expression or silence (Cheng and Blumenthal, 2008). DNA methyltransferase, the main enzyme involved in DNA methylation, was detected in female apices; it was obviously up-regulated at the stage of transformation from leaf to floral apices, and down-regulated at the other two developmental stages, relative to monoecious apices (**Table 3**). DNA methyltransferase is not only associated with DNA methylation, but is also associated with many important biological activities, including cell proliferation and senescence (Berger, 2007). The increase in methyltransferase gene expression

during the transitional stage may change the level of DNA methylation, leading to the subsequent repression of male flower gene expression and the appearance of only female flowers in female racemes. Genes associated with the reproduction process, identified as DEGs here and in future research, will contribute to elucidation of the molecular mechanisms underlying *Ricinus* sex determination and sex patterns.

DATA DEPOSITION

The Illumina pair-end reads of *Ricinus communis* L. obtained in the study are available at NCBI SRP064760.

AUTHOR CONTRIBUTIONS

YX conceived and designed the study, amended the manuscript. TM participated in its design and performed the experiments, analyzed the data and drafted the manuscript. XJ and WL prepared the samples, RNA extraction, cDNA library construction and helped with the sequencing. FC collected and prepared the samples. HJ provided the castor seeds and took part in field cultivation and bud collections. All authors read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.01208

REFERENCES

Ainsworth, C. C. (1999). *Sex Determination in Plants*. Oxford: BIOS Scientific Publishers.

Amaral, J. G. C. (2003). Genetic *Variability* for *Agronomic Characteristics* between self *Pollinated Lines* of *Castor* (Ricinus communis L.) cv. AL *Guarany*. Ph.D. dissertation (Agronomy), College of Agronomic Sciences, São Paulo State University, Botucatu.

Anjani, K. (2012). Castor genetic resources: a primary gene pool for exploitation. *Ind. Crop Prod.* 35, 1–14. doi: 10.1016/j.indcrop.2011.06.011

Atsmon, D., and Galun, E. (1960). A morphogenetic study of staminate, pistillate and hermaphrodite flowers in *Cucumis sativus* L. *Int. Soc. Plant Morphol.* 10, 110–115.

Audic, S., and Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Res.* 7, 986–995.

Baldanzi, M., and Pugliesi, C. (1998). Selection for non-branching in castor, *Ricinus communis* L. *Plant Breed.* 117, 392–394. doi: 10.1111/j.1439-0523.1998.tb01960.x

Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29, 1165–1188. doi: 10.1214/aos/1013699998

Bentley, D. R. (2006). Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* 16, 545–552. doi: 10.1016/j.gde.2006.10.009

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature* 447, 407–412. doi: 10.1038/nature05915

Bleecker, A. B., and Kende, H. (2000). Ethylene: a gaseous signal molecule in plants. *Annu. Rev. Cell Dev. Biol.* 16, 1–18. doi: 10.1146/annurev.cellbio.16.1.1

Bracale, M., Caporali, E., Galli, M. G., Longo, C., Marziani-Longo, G., Rossi, G., et al. (1991). Sex determination and differentiation in *Asparagus officinalis* L. *Plant Sci.* 80, 67–77. doi: 10.1016/0168-9452(91)90273-B

Brigham, R. D. (1967). Inheritance of two female-sterile characters in dwarf-internode castor (*Ricinus communis* L.). *Crop Sci. Soc. Am.* 7, 648–650. doi: 10.2135/cropsci1967.0011183X000700060027x

Chan, A. P., Crabtree, J., Zhao, Q., Lorenzi, H., Orvis, J., Puiu, D., et al. (2010). Draft genome sequence of the oilseed species *Ricinus communis*. *Nat. Biotechnol.* 28, 951–956. doi: 10.1038/nbt.1674

Chauhan, S. V. S., Saxena, B. K., and Kinoshita, T. (1987). Effect of daminozide (B9) on sex-expression and seed setting in castor bean, *Ricinus communis* L. *Jpn. J. Breed.* 37, 262–266. doi: 10.1270/jsbbs1951.37.262

Chen, M. L., Fu, X. M., Liu, J. Q., Ye, T. T., Hou, S. Y., Huang, Y. Q., et al. (2012). Highly sensitive and quantitative profiling of acidic phytohormones using derivatization approach coupled with nano-LC-ESI-Q-TOF-MS analysis. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 905, 67–74. doi: 10.1016/j.jchromb.2012.08.005

Cheng, P. C., Greyson, R. I., and Walden, D. B. (1983). Organ initiation and the development of unisexual flowers in the tassel and ear of Zea mays. *Am. J. Bot.* 70, 450–462. doi: 10.2307/2443252

Cheng, X., and Blumenthal, R. M. (2008). Mammalian DNA methyltransferases: a structural perspective. *Structure* 16, 341–350. doi: 10.1016/j.str.2008.01.004

Chu, V. T., Gottardo, R., Raftery, A. E., Bumgarner, R. E., and Yeung, K. Y. (2008). MeV+R: using MeV as a graphical user interface for Bioconductor applications in microarray analysis. *Genome Biol.* 9:R118 doi: 10.1186/gb-2008-9-7-r118

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

Davies, P. J. (2004). *Plant Hormones: Biosynthesis, Signal Transduction, Action*! Dordrecht; Boston, MA; London: Kluwer academic publishers.

Dellaporta, S. L., and Calderon-Urrea, A. (1993). Sex determination in flowering plants. *Plant Cell* 5, 1241–1251. doi: 10.1105/tpc.5.10.1241

DeLong, A., Calderon-Urrea, A., and Dellaporta, S. L. (1993). Sex determination gene TASSELSEED2 of maize encodes a short-chain alcohol dehydrogenase required for stage-specific floral organ abortion. *Cell* 74, 757–768. doi: 10.1016/0092-8674(93)90522-R

Gaganias, A. A., and Berg, A. (1977). R.: the effect of morphactin (methyl 2-ehloro-9-hydroxyfluorene-9–carboxylate) on basipetal transport of indol-3-yl acetic acid in hypocotyl sections of Phaseolus vulgaris L. *Ann. Bot.* 41, 1135–1148.

Galun, E. (1962). Study of the inheritance of sex expression in the cucumber. The interaction of major genes with modifying genetic and non-genetic factors. *Genetica* 32, 134–163. doi: 10.1007/BF01816091

George, W. L., and Shifriss, O. (1967). Interspersed sexuality in *Ricinus*. *Genetics* 57, 347–356.

Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R., Dunn, M., et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science* 296, 92–100. doi: 10.1126/science.1068275

Guo, Q. Q., Ma, X. J., Wei, S. G., Qiu, D. Y., Wilson, I. W., Wu, P., et al. (2014). *De novo* transcriptome sequencing and digital gene expression analysis predict biosynthetic pathway of rhynchophylline and isorhynchophylline from Uncaria rhynchophylla, a non-model plant with potent anti-Alzheimer's properties. *BMC Genomics* 15:676. doi: 10.1186/1471-2164-15-676

Guo, S. G., Zheng, Y., Joung, J. G., Liu, S. Q., Zhang, Z. H., Crasta, O. R., et al. (2010). Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. *BMC Genomics* 11:384. doi: 10.1186/1471-2164-11-384

Halilu, A. D., Aba, D. A., and Ogunwole, J. O. (2013). Genetic variability, genetic gain and relationships of yield and yield components in castor (*Ricinus communis* L.). *Res. Rev. BioSci.* 7, 181–186.

Heslop-Harrison, J. (1957). The experimental modification of sex expression in flowering plants. *Biol. Rev.* 32, 38–90. doi: 10.1111/j.1469-185X.1957.tb01576.x

Hong, L. Z., Li, J., Schmidt-Küntzel, A., Warren, W. C., and Barsh, G. S. (2011). Digital gene expression for non-model organisms. *Genome Res.* 21, 1905–1915. doi: 10.1101/gr.122135.111

Howe, E. A., Sinha, R., Schlauch, D., and Quackenbush, J. (2011). RNA-Seq analysis in MeV. *Bioinformatics* 27, 3209–3210. doi: 10.1093/bioinformatics/btr490

Huang, S., Li, R., Zhang, Z., Li, L., Gu, X., Fan, W., et al. (2009). The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* 41, 1275–1281. doi: 10.1038/ng.475

Jacob, K. M. (1963). A trisomic male castor bean plant. *J. Heredity* 54, 292–296.

Jacob, K. M., and Atsmon, D. (1965). Sex inheritance in Ricinus communis L.: evidence for a genetic change during the ontogeny of female sex reversals. *Genetica* 36, 253–259. doi: 10.1007/BF01557157

Jeong, G. T., and Park, D. H. (2009). Optimization of biodiesel production from castor oil using response surface methodology. *Appl. Biochem. Biotech.* 156, 431–441. doi: 10.1007/s12010-008-8468-9

Kaplan, B., Davydov, O., Knight, H., Galon, Y., Knight, M. R., Fluhr, R., et al. (2006). Rapid transcriptome changes induced by cytosolic Ca2+ transients reveal ABRE-related sequences as Ca2+-responsive cis elements in Arabidopsis. *Plant Cell* 18, 2733–2748. doi: 10.1105/tpc.106.042713

Kater, M. M., Franken, J., Carney, K. J., Colombo, L., and Angenent, G. C. (2001). Sex determination in the monoecious species cucumber is confined to specific floral whorls. *Plant Cell* 13, 481–493. doi: 10.1105/tpc.13.3.481

Kumar, N., and Rao, P. (1980). Influence of kinetin and morphactin on changes in sex expression, carbohydrate and nitrogen fractions in castor (*Ricinus communis* L.). *Proc. Plant Sci.* 89, 457–464.

Kumar, N. R., and Rao, P. G. (1981). Activity of auxin like substances in relation to feminization of castor bean (*Ricinus communis* L.) influenced by kinetin and morphactin. *Turrialba* 31, 201–208.

Li, R. Q., Yu, C., Li, Y. R., Lam, T. W., Yiu, S. M., Kristiansen, K., et al. (2009a). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25, 1966–1967. doi: 10.1093/bioinformatics/btp336

Li, Z., Huang, S. W., Liu, S. Q., Pan, J. S., Zhang, Z. H., Tao, Q. Y., et al. (2009b). Molecular isolation of the M gene suggests that a conserved-residue conversion induces the formation of bisexual flowers in cucumber plants. *Genetics* 182, 1381–1385. doi: 10.1534/genetics.109.104737

Lorrain, S., Vailleau, F., Balagué, C., and Roby, D. (2003). Lesion mimic mutants: keys for deciphering cell death and defense pathways in plants. *Trends Plant Sci.* 8, 263–271. doi: 10.1016/S1360-1385(03)00108-0

Malepszy, S., and Niemirowicz-Szczytt, K. (1991). Sex determination in cucumber (*Cucumis sativus*) as a model system for molecular biology. *Plant Sci.* 80, 39–47. doi: 10.1016/0168-9452(91)90271-9

Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R., Morin, H., et al. (2009). A transposon-induced epigenetic change leads to sex determination in melon. *Nature* 461, 1135–1138. doi: 10.1038/nature08498

McIntyre, L. M., Lopiano, K. K., Morse, A. M., Amin, V., Oberg, A. L., Young, L. J., et al. (2011). RNA-seq: technical variability and sampling. *BMC Genomics* 12:293. doi: 10.1186/1471-2164-12-293

Mibus, H., and Tatlioglu, T. (2004). Molecular characterization and isolation of the F/f gene for femaleness in cucumber (*Cucumis sativus* L.). *Theor. Appl. Genet.* 109, 1669–1676. doi: 10.1007/s00122-004-1793-7

Mohan Ram, H. Y., and Sett, R. (1980). Induction of male flowers in a pistillate line of *Ricinus communis* L. by silver and cobalt ions. *Planta* 149, 413–415.

Morrissy, A. S., Morin, R. D., Delaney, A., Zeng, T., McDonald, H., Jones, S., et al. (2009). Next-generation tag sequencing for cancer gene expression profiling. *Genome Res.* 19, 1825–1835. doi: 10.1101/gr.094482.109

Oshlack, A., Robinson, M. D., and Young, M. D. (2010). From RNA-seq reads to differential expression results. *Genome Biol.* 11:220. doi: 10.1186/gb-2010-11-12-220

Peng, J., Richards, D. E., Hartley, N. M., Murphy, G. P., Devos, K. M., Flintham, J. E., et al. (1999). 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400, 256–261. doi: 10.1038/22307

Philipos, H., and Narayanaswamy, S. (1976). Induced potential for male sex expression in*Ricinus communis* L. by ethyl hydrogen-1-propylphosphonate. *Protoplasma* 87, 71–77. doi: 10.1007/BF01623959

Pranavi, B., Sitaram, G., Yamini, K. N., and Dinesh Kumar, V. (2011). Development of EST–SSR markers in castor bean (*Ricinus communis* L.) and their utilization for genetic purity testing of hybrids. *Genome* 54, 684–691. doi: 10.1139/g11-033

Ramsey, S. A., Klemm, S. L., Zak, D. E., Kennedy, K. A., Thorsson, V., Li, B., et al. (2008). Uncovering a macrophage transcriptional program by integrating evidence from motif scanning and expression dynamics. *PLoS Comput. Biol.* 4:e1000021. doi: 10.1371/journal.pcbi.1000021

Rao, P. G. (1969). Sex expression in *Ricinus communis* L. *Sci. Cult.* 35, 326–327.

Rkey, V. A. (1978). *M.: Experimental Studies on Effects of Plant Growth Regulators on Vegetative Morphology and Floral Morphogenesis in Ricinus communis L.* Ph.D dissertation (Agronomy), Kaupur University.

Savy Filho, A. (2005). "Castor bean breeding," in *Improvement of Cultivated Species*, ed A. Borém (Viçosa: Federal University of Viçosa).

Shifriss, O. (1956). Sex instability in ricinus. *Genetics* 41, 265–280.

Shifriss, O. (1960). Conventional and unconventional systems controlling sex variations in *Ricinus*. *J. Genet.* 57, 361–388. doi: 10.1007/BF02987242

Shifriss, O. (1961). Gibberellin as Sex Regulator in *Ricinus communis*. *Science* 133, 2061–2062. doi: 10.1126/science.133.3470.2061

Tan, M., Yan, M., Wang, L., and Yan, X. (2011). Effect of chemical treatments on sex expression of *Ricinus communis*. *Chin. Agr. Sci. Bull.* 27, 164–169.

Tanurdzic, M., and Banks, J. A. (2004). Sex-determining mechanisms in land plants. *Plant Cell* 16(Suppl.), S61–S71. doi: 10.1105/tpc.016667

`t Hoen, P. A. C., Ariyurek, Y., Thygesen, H. H., Vreugdenhil, E., Vossen, R. H. A. M., de Menezes, R. X., et al. (2008). Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Res.* 36:e141. doi: 10.1093/nar/gkn705

Thomson, K. S., and Leopold, A. C. (1974). *In-vitro* binding of morphactins and 1-N-naphthylphthalamic acid in corn coleoptiles and their effects on auxin transport. *Planta* 115, 259–270. doi: 10.1007/BF00390522

Tian, D. Q., Pan, X. Y., Yu, Y. M., Wang, W. Y., Zhang, F., Ge, Y. Y., et al. (2013). *De novo* characterization of the Anthurium transcriptome and analysis of its digital gene expression under cold stress. *BMC Genomics* 14:827. doi: 10.1186/1471-2164-14-827

Trebitsh, T., Rudich, J., and Riov, J. (1987). Auxin, biosynthesis of ethylene and sex expression in cucumber (*Cucumis sativus*). *Plant Growth Regul.* 5, 105–113. doi: 10.1007/BF00024738

Trebitsh, T., Staub, J. E., and O'Neill, S. D. (1997). Identification of a 1-aminocyclopropane-1-carboxylic acid synthase gene linked to the female (F) locus that enhances female sex expression in cucumber. *Plant Physiol.* 113, 987–995.

Varkey, M., and Nigam, R. K. (1982). A gradual reduction of female structures in the pistillate flowers of Ricinus communis L. (castor-bean) with chlorflurenol (morphactin). *Biol. Plantarum.* 24, 152–154. doi: 10.1007/BF02902863

Veit, B., Schmidt, R. J., Hake, S., and Yanofsky, M. F. (1993). Maize floral development: new genes and old mutants. *Plant Cell* 5, 1205–1215. doi: 10.1105/tpc.5.10.1205

Wang, L., Tan, M., Yan, M., Wang, L., and Yan, X. (2012). Initial research on inflorescence characteristics and flower bud differentiation of *Ricinus communis* L. *Chin. J. Oil Crop Sci.* 34, 544–550.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919

Wei, M. M., Song, M. Z., Fan, S. L., and Yu, S. X. (2013). Transcriptomic analysis of differentially expressed genes during anther development in genetic male sterile and wild type cotton by digital gene-expression profiling. *BMC Genomics* 14:97. doi: 10.1186/1471-2164-14-97

Wittwer, S. H., and Hillyer, I. G. (1954). Chemical induction of male sterility in cucurbits. *Science* 120, 893–894. doi: 10.1126/science.120.3126.893

Wu, S., Liao, G., and Guo, C. (1998). Floral Development in *Ricinus Communis* L. (Euphorbiaceae). *Biol. Bull. NTNU* 33, 105–113.

Wu, T., Qin, Z. W., Zhou, X. Y., Feng, Z., and Du, Y. L. (2010). Transcriptome profile analysis of floral sex determination in cucumber. *J. Plant Physiol.* 167, 905–913. doi: 10.1016/j.jplph.2010.02.004

Xu, Q., Chen, L. L., Ruan, X., Chen, D., Zhu, A., Chen, C., et al. (2013). The draft genome of sweet orange (*Citrus sinensis*). *Nat. Genet.* 45, 59–66. doi: 10.1038/ng.2472

Yamasaki, S., Fujii, N., and Takahashi, H. (2000). The ethylene-regulated expression of CS-ETR2 and CS-ERS genes in cucumber plants and their possible involvement with sex expression in flowers. *Plant Cell Physiol.* 41, 608–616. doi: 10.1093/pcp/41.5.608

Ye, D., Oliveira, M., Veuskens, J., Wu, Y., Installe, P., Hinnisdaels, S., et al. (1991). Sex determination in the dioecious Melandrium. The X/Y chromosome system allows complementary cloning strategies. *Plant Sci.* 80, 93–106. doi: 10.1016/0168-9452(91)90275-D

Ye, J., Fang, L., Zheng, H. K., Zhang, Y., Chen, J., Zhang, Z. J., et al. (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* 34, 293–297. doi: 10.1093/nar/gkl031

Yin, T. J., and Quinn, J. A. (1995). Tests of a mechanistic model of one hormone regulating both sexes in *Cucumis sativus* (Cucurbitaceae). *Am. J. Bot.* 82, 1537–1546. doi: 10.2307/2446182

Yu, J., Hu, S., Wang, J., Wong, G. K., Li, S., Liu, B., et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science* 296, 79–92. doi: 10.1126/science.1068037

Yu, Y., Huang, W., Chen, H., Wu, G., Yuan, H., Song, X., et al. (2014). Identification of differentially expressed genes in flax (*Linum usitatissimum* L.) under saline-alkaline stress by digital gene expression. *Gene.* 549, 113–122. doi: 10.1016/j.gene.2014.07.053

Zhang, S. P., Xiao, Y. N., Zhao, J. R., Wang, F. G., and Zheng, Y. L. (2013). Digital gene expression analysis of early root infection resistance to Sporisorium reilianum f. sp zeae in maize. *Mol. Genet. Genomics* 288, 21–37. doi: 10.1007/s00438-012-0727-3

Zhao, W., Xia, W., Li, J., Sheng, S., Lei, L., and Zhao, S. (2014). Transcriptome profiling and digital gene expression analysis of Fallopia multiflora to discover putative genes involved in the biosynthesis of 2,3,5,4′-tetrahydroxy stilbene-2-O-β-d-glucoside. *Gene.* 547, 126–135. doi: 10.1016/j.gene.2014.06.041

Physiological Mechanisms Underlying the High-Grain Yield and High-Nitrogen Use Efficiency of Elite Rice Varieties under a Low Rate of Nitrogen Application in China

Lilian Wu, Shen Yuan, Liying Huang, Fan Sun, Guanglong Zhu, Guohui Li, Shah Fahad, Shaobing Peng and Fei Wang *

National Key Laboratory of Crop Improvement, Key Laboratory of Crop Ecophysiology and Farming System in the Middle Reaches of the Yangtze River, Ministry of Agriculture, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China

Selecting rice varieties with a high nitrogen (N) use efficiency (NUE) is the best approach to reduce N fertilizer application in rice production and is one of the objectives of the Green Super Rice (GSR) Project in China. However, the performance of elite candidate GSR varieties under low N supply remains unclear. In the present study, differences in the grain yield and NUE of 13 and 14 candidate varieties with two controls were determined at a N rate of 100 kg ha^{-1} in field experiments in 2014 and 2015, respectively. The grain yield for all of the rice varieties ranged from 8.67 to 11.09 t ha^{-1}, except for a japonica rice variety YG29, which had a grain yield of 6.42 t ha^{-1}. HY549 and YY4949 produced the highest grain yield, reflecting a higher biomass production and harvest index in 2014 and 2015, respectively. Total N uptake at maturity (TN$_{PM}$) ranged from 144 to 210 kg ha^{-1}, while the nitrogen use efficiency for grain production (NUEg) ranged from 35.2 to 62.0 kg kg^{-1}. Both TN$_{PM}$ and NUEg showed a significant quadratic correlation with grain yield, indicating that it is possible to obtain high grain yield and NUEg with the reduction of TN$_{PM}$. The correlation between N-related parameters and yield-related traits suggests that promoting pre-heading growth could increase TN$_{PM}$, while high biomass accumulation during the grain filling period and large panicles are important for a higher NUEg. In addition, there were significant and negative correlations between the NUEg and N concentrations in leaf, stem, and grain tissues at maturity. Further improvements in NUEg require a reduction in the stem N concentration but not the leaf N concentration. The daily grain yield was the only parameter that significantly and positively correlated with both TN$_{PM}$ and NUEg. This study determined variations in the grain yield and NUE of elite candidate GSR rice varieties and provided plant traits that could be used as selection criteria in breeding N-efficient rice varieties.

Keywords: daily grain yield, Green Super Rice, grain yield, nitrogen use efficiency

INTRODUCTION

Rice is one of the staple food crops for approximately half of the global population (Godfray et al., 2010), and rice production must increase by 70% by 2050 to satisfy the requirements of the growing world population (Koning et al., 2008; Godfray et al., 2010). Moreover, increased rice production needs to be achieved under the pressures of decreased arable land area, global climate change (Peng et al., 2004), intensified natural disasters (Tao et al., 2013), and the frequent occurrence of diseases and pests (Sheng et al., 2003). Therefore, it is imperative to develop new varieties that have a higher yield potential and improved adaptation to the environment.

Yield potential is defined as the yield of a cultivar grown in environments to which it is adapted, with nutrients and non-limiting water, as well as pests, diseases, weeds, lodging, and other stresses effectively controlled (Evans and Fischer, 1999). Cassman (1999) provided a more functional definition of yield potential, suggesting that this parameter is the yield obtained when an adapted cultivar is grown with the minimal possible stress, which is achieved by using the best management practices. In rice, yield potential has been significantly augmented, reflecting the utilization of semi-dwarf genes, heterosis, and the combination of intersubspecific heterosis and new plant types (Peng et al., 2008). In 2014, the elite super hybrid rice Y-Liang-You900 (YLY900) showed a record high yield of 15.4 t ha^{-1} (Li et al., 2014). However, the main dilemma is that new varieties that have a high yield potential were achieved using surplus nutrient application, suggesting that farmers should apply a higher amount of fertilizers than the minimum required to produce the highest grain yield in rice production (Peng et al., 2002). The performance of these newly developed high-yielding varieties under low nutrient input remains unclear.

Nitrogen (N) is a vital element for plant development and growth, and the application of N fertilizer could significantly increase yield formation (Andrews et al., 2013). From 1960 to 2012, the global N fertilizer consumption increased by 800% and the annual N consumption in China increased from 8 to 35% of the world's N consumption (data from IFA). In China, the average rate of N application in rice production is ~180 kg ha^{-1}, which is 75% higher than the world average rate (Peng et al., 2002, 2006). High N fertilizer input leads to low nitrogen use efficiency (NUE) due to the rapid N losses from ammonia volatilization, denitrification, surface runoff, and leaching in the soil-floodwater system (Vlek and Byrnes, 1986; De Datta and Buresh, 1989). A low NUE results in significant environmental pollution, such as soil acidification (Guo et al., 2010), air pollution (Smil, 1999), and water eutrophication (Diaz and Rosenberg, 2008). To increase the NUE in rice production, scientists have developed a range of optimized crop management practices, such as Site-Specific N Management (SSNM, Dobermann et al., 2002), Real-Time N Management (RTNM, Peng et al., 2006), the San-Ding Cultivation Method (SDCM, Zou et al., 2006), and "Three Controls" Nutrient Management Technology (TCNM, Zhong et al., 2007).

One potential approach to reduce N fertilizer application in rice production is the development of varieties with an improved

NUE (Sun et al., 2014; Hu et al., 2015). Variations in the NUE of different rice genotypes have been determined, and NUE-related traits have been evaluated for their accuracy in reflecting genotypic variation in rice NUE from 1987 to 2003 (Broadbent et al., 1987; De Datta and Broadbent, 1988, 1993; Tirol-Padre et al., 1996; Singh et al., 1998; Inthapanya et al., 2000; Ntanos and Koutroubas, 2002; Koutroubas and Ntanos, 2003). These studies reported significant differences in N uptake capacity and N use efficiency for grain production (NUEg), suggesting candidate parameters reflecting NUE variation, such as WP/Nt (panicle weight/total N uptake), NPI (the product of grain yield at zero N treatment and NUEg), among others. Since 2003, there have only been a few studies on NUE variation in rice. Recently, Ju et al. (2015) compared the grain yield of two N-efficient varieties and two N-inefficient varieties under low N input conditions, reporting that a high grain yield at a low N rate was associated with deeper roots, increased root oxidation activity, and a higher photosynthetic NUE. However, NUE differences among newly developed elite varieties under low N input condition have not been studied.

Zhang (2007) proposed strategies for developing Green Super Rice (GSR) to meet the challenges in rice production. In 2010, the Ministry of Science and Technology of China launched a mega project to develop GSR as proposed by Zhang (2007). One main aspect in this project is to decrease N fertilizer application in rice production through the genetic development of N-efficient varieties. Thus, the objectives of the present study were to evaluate the grain yield and NUE of the newly developed candidate GSR varieties from different breeding institutes under low N supply and to examine the physiological mechanisms underlying the differences in NUE.

MATERIALS AND METHODS

Plant Materials

In 2014, 13 candidate GSR varieties were grown in the middle season with YLY6 (a super hybrid variety) and HHZ (a potential GSR) as control varieties. In 2015, 14 new candidate GSR varieties were grown in the middle season with YLY6 and HHZ as control varieties. HY549 and HLY630 were used in both years. All of the candidate GSR varieties were developed in recent years, achieving a high grain yield in local variety tests, while the two control varieties, YLY6 and HHZ, were widely planted in South China in the last decade. Detailed information concerning these varieties is shown in **Table 1**.

Experimental Design

Field experiments were conducted in Zhougan Village (2014) and Zhangbang Village (2015) of Wuxue County, Hubei Province, China. Prior to the experiments, soil samples from the upper 20-cm layer were collected to analyze the soil chemical properties. In 2014, the soil had a clay loam texture with a pH of 5.60, organic matter of 27.18 g kg^{-1}, total N of 1.83 g kg^{-1}, available P of 4.91 mg kg^{-1}, and available K of 105.8 mg kg^{-1}, while in 2015, the soil had a clay loam texture with a pH of 5.20, organic matter of 26.69 g kg^{-1}, total N of 1.19 g kg^{-1}, available P of 22.56 mg kg^{-1}, and available K of 159.2 mg kg^{-1}. The data for daily rainfall, solar

TABLE 1 | Grain yield and yield components of the varieties in 2014 and 2015 at Wuxue County, Hubei Province, China.

Variety	Abbreviation	Year of release	Type	Institute of release
2014				
Hanyou549	HY549	–	Hybrid	Shanghai Agrobiological Gene Center
Huiliangyou858	HYL858	–	Hybrid	Anhui Academy of Agricultural Sciences
Jinkeyou651	JKY651	2013	Hybrid	Huazhong Agricultural University
9you6hao	9Y6H	–	Hybrid	Chinese Academy of Agricultural Sciences
Huiliangyou630	HLY630	2014	Hybrid	Anhui Academy of Agricultural Sciences
Rongfengyou41	RFY41	–	Hybrid	Huazhong Agricultural University
Rongyou225	RY225	2009	Hybrid	Jiangxi Academy of Agricultural Sciences
Wuyouhang1573	WYH1573	2014	Hybrid	Jiangxi Academy of Agricultural Sciences
Quanyou982	QY982	–	Hybrid	Chinese Academy of Agricultural Sciences
Yangliangyou6	YLY6	2003	Superhybrid	Lixiahe Institute of Agricultural Sciences
Huanghuazhan	HHZ	2005	Inbred	Guangdong Academy of Agricultural Sciences
Guangliangyou5	GLY5	2013	Hybrid	Hubei Academy of Agricultural Sciences
Hanyou73	HY73	2014	Hybrid	Shanghai Agrobiological Gene Center
Zhongzu14	ZZ14	2006	Inbred	Chinese Rice Research Institute
Yungeng29	YG29	2011	Inbred	Yunnan Academy of Agricultural Sciences
2015				
Yongyou4949	YY4949	2015	Hybrid	Ningbo Seed Company
Hanyou549	HY549	–	Hybrid	Shanghai Agrobiological Gene Center
Chuanyou5727	CY5727	–	Hybrid	Sichuan Academy of Agricultural Sciences
Jiyou225	JY225	2014	Hybrid	Jiangxi Academy of Agricultural Sciences
Yyou278	YY278	–	Hybrid	Hunan Hybrid Rice Research Center
Longliangyouhuazhan	LLYHZ	2015	Hybrid	Longping High-Tech
Huiliangyou630	HLY630	2014	Hybrid	Anhui Academy of Agricultural Sciences
Mingliangyou143	MLY143	–	Hybrid	Hunan Hybrid Rice Research Center
Luoyou10	LY10	–	Hybrid	Wuhan University
Huanghuazhan	HHZ	2005	Inbred	Guangdong Academy of Agricultural Sciences
Jinyou959	JY959	–	hybrid	Yunnan Jinrui Seed Industry Company
Zhonghua1	ZH1	–	hybrid	Guangdong Academy of Agricultural Sciences
Heliangyou7185	HLY7185	–	hybrid	Fujian Academy of Agricultural Sciences
Huhan1709	HH17-09	–	hybrid	Shanghai Agrobiological Gene Center
Shanyou108	SY108	2013	hybrid	Guizhou Academy of Agricultural Sciences
Yangliangyou6	YLY6	2003	Superhybrid	Lixiahe Institute of Agricultural Sciences
Hualiangyou1511	HLY1511	–	Hybrid	Huazhong Agricultural University
Wushansimiao	WSSM	2009	Hybrid	Guangdong Academy of Agricultural Sciences

radiation, and minimum and maximum temperatures during the rice growing season were collected at a meteorological station (CR800, Campbell Scientific Inc., Logan, Utah, USA) near the fields, and are shown in **Figure 1**.

The experiment was arranged in a completely randomized block design with four replications. The seedlings were raised in the seedbed with a sowing date of May 23, 2014 and May 25, 2015. Twenty-five-day-old seedlings were transplanted on June 17 and June 19 in 2014 and 2015, respectively, at a hill spacing of 20.0 × 20.0 cm with two seedlings per hill and a plot size of 30 m^2 in 2014 and 25 m^2 in 2015. The fertilizers were manually broadcasted and incorporated 1 day before transplanting for basal application (40 kg N ha^{-1} urea, 40 kg P ha^{-1} calcium superphosphate, 50 kg K ha$^-$ potassium chloride, and 5 kg Zn ha^{-1} zinc sulfate heptahydrate for 2 years). Nitrogen topdressings

were applied at midtillering (20 kg ha^{-1}) and panicle initiation (PI; 40 kg ha^{-1}), and K was topdressed at PI at a rate of 50 kg ha^{-1} during a 2-year experimental period. To minimize seepage between the plots, all of the bunds were covered with plastic film and installed at a depth of 20 cm below the soil surface. A water depth of 5–10 cm was maintained until 7 days prior to maturity when the fields were drained. The weeds were controlled manually and using herbicides. Pests and diseases were controlled using insecticides and fungicides; no obvious water, weed, pest, or disease stresses were observed during the experiment.

Crop Measurements

Twelve hills were sampled from each plot at mid-tillering (MT), PI, and heading (HD). Plant height and stem (main stems plus tillers) numbers were recorded. A tiller with at least one leaf was
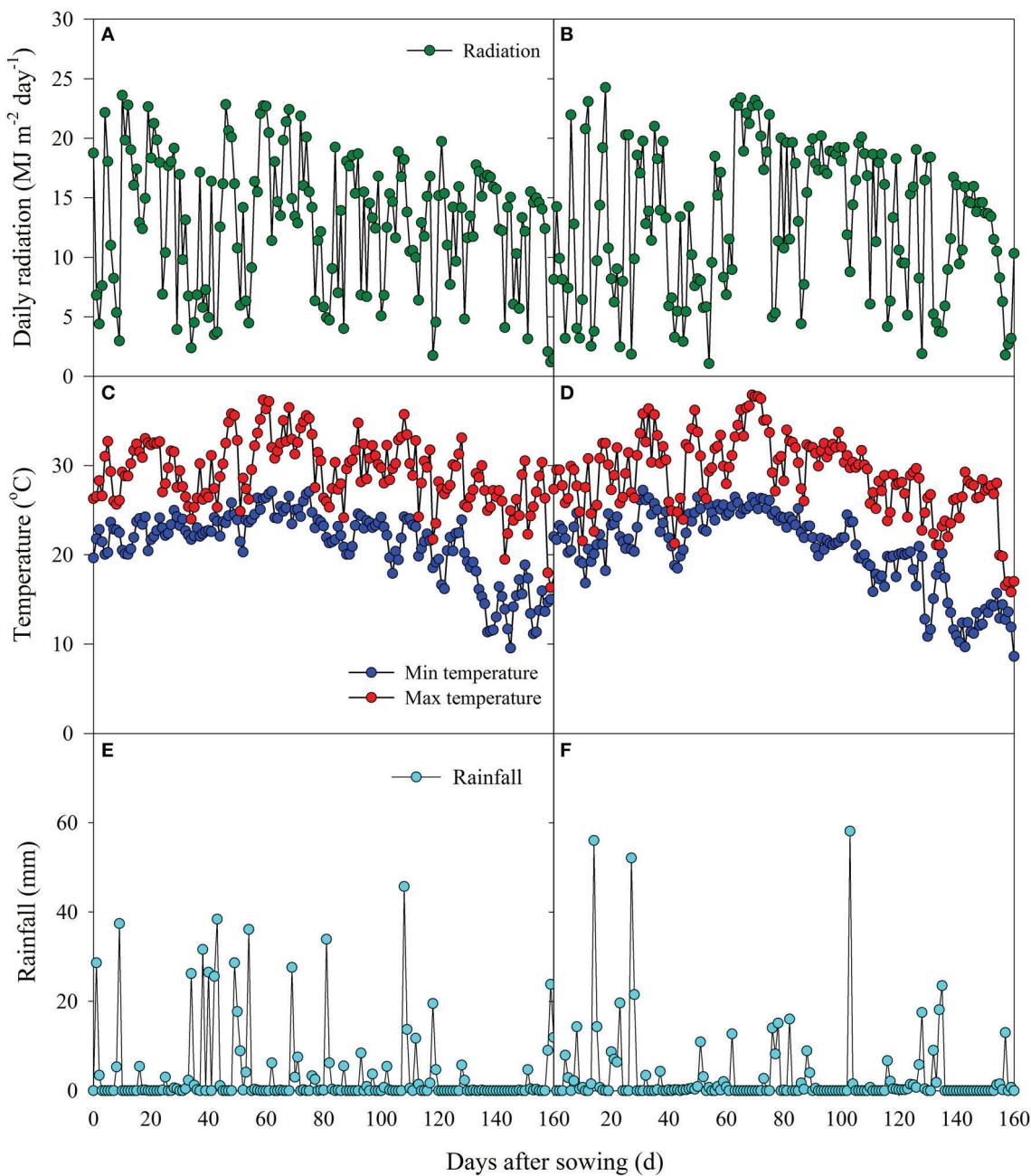
FIGURE 1 | Daily radiation, daily minimum and maximum temperatures, and daily rainfall in the middle growing season of 2014 (A, C, E) and 2015 (B, D, F).

counted as a stem. The plant samples were separated into leaf blade (leaf), culm plus sheath (stem), and panicle. The green leaf area was measured using a leaf area meter (LI-3100, LI-COR, Lincoln, NE, USA) and was expressed as the leaf area index (LAI). The specific leaf weight (SLW) was defined as the ratio of the leaf dry weight to leaf area. The dry weight of each component was determined after oven drying at 80°C to a constant weight. The plant dry weight was the sum of all of the aboveground components.

At physiology maturity (PM), 12 hills were obtained from each subplot to determine the aboveground total biomass and other yield components. Plant height and panicle number were obtained from 12 hills. The plant samples were separated into leaf, stem and panicle. The dry weight of straw was determined after oven drying at 80°C to a constant weight. The panicles were hand-threshed, and the filled spikelets were separated from unfilled spikelets after submerging them in tap water. The empty spikelets were separated from the half-filled spikelets through

winnowing. Three 30-g subsamples of filled spikelets, three 2-g subsamples of empty spikelets, and the total number of half-filled spikelets were obtained to quantify the number of spikelets per m^2. The dry weights of rachis, filled, half-filled, and unfilled spikelets were determined after oven drying at 80°C to constant weight. The aboveground total biomass was calculated as the total dry matter of straw, rachis, and filled, half-filled, and unfilled spikelets. The spikelets per panicle, grain filling percentage (100× filled spikelet number/total spikelet number), and harvest index (HI) (100× filled spikelet weight/aboveground total biomass) were calculated. The grain yield was determined from a 5-m^2 area in each subplot and was adjusted to a standard moisture content of 0.14 g H$_2$O g^{-1} fresh weight. The grain moisture content was measured with a digital moisture tester (DMC-700, Seedburo, Chicago, IL, USA).

The tissue N concentration of each component at HD and PM was determined using an Elemental analyzer (Elementar vario MAX CNS/CN, Elementar Trading Co., Ltd, Germany). The plant N accumulation at HD and PM was calculated as the sum of N in each of the aboveground components. The nitrogen use efficiency for grain production (NUEg) was calculated as the grain yield per unit plant N accumulation. The nitrogen use efficiency in biomass production (NUEb) was determined as the ratio of biomass production to plant N accumulation. The nitrogen harvest index (NHI) was calculated as the percentage of accumulated N in grain to plant N accumulation (Peng et al., 1996).

Statistical Analysis

The data were analyzed using analysis of variance, and the mean values among the varieties were compared based on the least significant difference (LSD) test at the 0.05 probability level.

RESULTS

Growth Duration

The growth duration ranged from 118 to 141 d in 2014 and from 119 to 144 d in 2015. For the majority of varieties, the growth duration ranged from 130 to 141 d in 2014 and from 137 to 144 d in 2015 (**Table 2**). The days from sowing to flowering for these varieties ranged from 78 to 99 d in 2014 and from 84 to 101 d in 2015, and the grain filling period was from 33 to 47 d in 2014 and from 33 to 53 d in 2015. Generally, the growth duration of HY549 and HLY630 was similar in 2014 and 2015. The growth duration of HHZ in 2014 was longer than that in 2015, while the growth duration of YLY6 in 2014 was shorter than that in 2015. Notably, YY4949 had the shortest growth period prior to flowering but had the longest growth period after flowering in 2015 (**Table 2**).

Grain Yield and Yield Components

The grain yield ranged from 6.42 to 10.41 t ha^{-1} in 2014 (**Table 3**). HY549 produced the highest grain yield, and YG29 produced the lowest grain yield. No significant difference in grain yield was observed between YLY6 and HHZ. Compared with the two controls, HY549, HYL858, JKY651, 9Y6H, HLY630, and RFY41 produced a significantly superior grain yield. The grain yields of RY225, WYH1573, QY982, and GLY5 were similar to

TABLE 2 | Growth duration of the varieties in 2014 and 2015 at Wuxue County, Hubei Province, China.

Variety	Sowing-flowering	Flowering-maturity	Total growth duration
2014			
HY549	99	41	141
HYL858	85	45	130
JKY651	99	38	137
9Y6H	90	47	137
HLY630	90	47	137
RFY41	85	52	137
RY225	83	54	137
WYH1573	83	47	130
QY982	85	45	130
YLY6	93	44	137
HHZ	85	45	130
GLY5	90	47	137
HY73	78	40	118
ZZ14	85	45	130
YG29	85	33	118
2015			
YY4949	84	53	137
HY549	95	49	144
CY5727	91	53	144
JY225	85	52	137
YY278	101	43	144
LLYHZ	95	49	144
HLY630	91	46	137
MLY143	98	39	137
LY10	92	45	137
HHZ	86	33	119
JY959	95	49	144
ZH1	95	49	144
HLY7185	87	50	137
HH1709	87	50	137
SY108	90	47	137
YLY6	98	46	144
HLY1511	95	42	137
WSSM	90	47	137

those of the two controls. HY73, ZZ14, and YG29 produced a significantly lower grain yield compared with the two controls (**Table 3**). In 2015, the grain yield ranged from 8.96 to 11.09 t ha^{-1}. Notably, a higher grain yield was observed in YY4949, HY549, and CY5727 than in either HHZ or YLY6. WSSM generated a significantly lower grain yield than HHZ. The average grain yield of HHZ, YLY6, HY549, and HLY630 was 9.82 t ha^{-1} in 2014 and 10.13 t ha^{-1} in 2015 (**Table 3**), and analysis of variance indicated that the difference in the average grain yield of the four common varieties between 2014 and 2015 was not statistically significant.

In 2014, the higher grain yields of HY549, HYL858, JKY651, and 9Y6H primarily reflected the higher biomass production (**Table 3**). HY549 had a smaller number of large panicles,

TABLE 3 | Grain yield and yield components of the varieties in 2014 and 2015 at Wuxue County, Hubei Province, China.

Variety	Grain yield (t ha^{-1})	Panicles m^{-2}	Spikelets per panicle	Spikelets m^{-2} ($\times 10^3$)	Grain filling percentage (%)	Grain weight (mg)	Biomass (t ha^{-1})	Harvest index (%)
2014								
HY549	10.41	180	290	52.2	83.2	23.5	18.6	54.8
HYL858	10.07	209	198	41.4	84.1	27.3	17.0	55.6
JKY651	10.05	232	176	40.9	88.9	26.0	17.4	54.4
9Y6H	10.04	183	202	37.1	90.5	28.8	17.7	54.7
HLY630	9.98	201	194	39.1	88.0	25.9	16.0	55.5
RFY41	9.98	197	194	38.1	83.2	28.3	15.9	56.5
RY225	9.71	204	231	47.1	80.0	23.1	14.9	58.4
WYH1573	9.71	208	242	50.4	89.0	20.8	16.4	57.0
QY982	9.60	207	188	38.9	87.2	29.0	17.3	56.9
YLY6	9.50	179	191	34.2	91.1	29.5	16.7	55.0
HHZ	9.39	220	219	48.1	90.6	19.5	15.0	56.5
GLY5	9.11	181	206	37.3	79.5	27.5	15.5	52.7
HY73	8.76	188	185	34.8	87.8	28.8	16.3	53.7
ZZ14	8.67	262	216	56.9	76.3	18.4	14.5	54.8
YG29	6.42	199	175	35.1	62.3	23.7	13.4	38.3
LSD (0.05)	0.39	16	17	4.0	4.2	0.5	1.3	2.0
2015								
YY4949	11.09	209	265	55.7	93.7	21.5	18.7	59.7
HY549	10.99	240	206	49.5	88.2	23.7	20.3	51.1
CY5727	10.93	244	168	40.9	93.2	27.8	20.3	51.5
JY225	10.65	255	173	44.1	85.2	24.2	16.4	55.5
YY278	10.60	246	180	44.2	92.6	25.3	20.3	50.9
LLYHZ	10.53	244	178	43.4	93.4	23.4	19.2	49.5
HLY630	10.49	241	172	41.6	95.8	25.9	18.1	56.9
MLY143	10.34	265	153	40.7	93.4	26.5	18.6	54.0
LY10	10.26	255	195	49.4	95.5	22.3	18.1	53.6
HHZ	10.04	277	153	42.3	93.1	21.3	16.7	50.3
JY959	9.97	202	222	44.7	90.9	24.9	19.4	51.0
ZH1	9.83	174	230	39.9	92.8	25.8	20.9	45.8
HLY7185	9.74	270	142	38.1	92.2	24.9	16.8	52.3
HH17-09	9.66	293	161	47.0	92.1	19.0	16.0	51.5
SY108	9.64	231	160	36.7	88.8	29.1	18.0	52.8
YLY6	9.61	208	170	35.3	90.4	28.6	18.4	49.5
HLY1511	9.44	214	153	32.7	92.6	28.7	17.1	51.0
WSSM	8.96	283	158	44.6	91.2	19.9	17.2	47.1
LSD (0.05)	0.92	23	36	8.2	2.5	4.0	1.9	2.1

resulting in a significantly larger sink size (spikelets m^{-2}) compared with the other varieties. In 2015, the higher grain yield of YY4949 resulted from a higher harvest index, while the higher biomass production of HY549 and CY5727 contributed to the yield advantage of these two varieties. YY4949 and HY549 had a larger sink size than the other varieties. The higher grain yields of HHZ, YLY6, HY549, and HLY630 in 2015 reflected the higher biomass production (**Table 3**).

Nitrogen Uptake and Use Efficiency

Significant differences were observed among the varieties for total N uptake at the heading stage (TN$_{HD}$), total N uptake during the grain filling period (TN$_{GF}$), total N uptake at maturity (TN$_{PM}$), nitrogen use efficiency for grain production (NUEg), nitrogen use efficiency for grain production (NUEb), and NHI in 2014 and 2015. The TN$_{PM}$ ranged from 144 to 172 kg ha^{-1} in 2014 and from 158 to 210 kg ha^{-1} in 2015. The NUEg of the varieties ranged from 35.2 to 62.0 kg kg^{-1} in 2014 and from 43.1 to 58.4 kg kg^{-1} in 2015. The TN$_{PM}$ of HHZ, YLY6, HY549, and HLY630 in 2015 was higher than that in 2014, resulting in a lower NUEg for these varieties in 2015 than that in 2014 (**Table 4**).

In both years, the leaf N concentration was significantly higher than that in the stem and panicle at the heading

TABLE 4 | Nitrogen uptake at the heading stage (TN$_{HD}$), nitrogen uptake during grain filling period (TN$_{GF}$), nitrogen uptake at maturity (TN$_{PM}$), nitrogen use efficiency for grain production (NUEg), nitrogen use efficiency for biomass production (NUEb), and nitrogen harvest index (NHI) of the varieties in 2014 and 2015 at Wuxue County, Hubei Province, China.

Variety	TN$_{HD}$ (kg ha^{-1})	TN$_{GF}$ (kg ha^{-1})	TN$_{PM}$ (kg ha^{-1})	NUEg (kg kg^{-1})	NUEb (kg kg^{-1})	NHI (%)
2014						
HY549	173	−4.3	169	60.9	111	70.9
HLY858	137	35.4	172	55.6	101	71.6
JKY651	166	−0.7	166	57.2	105	71.6
9Y6H	156	9.0	165	59.0	108	72.9
HLY630	146	10.3	157	57.3	103	75.0
RFY41	130	32.3	162	55.6	98	71.5
RY225	113	36.6	150	58.4	100	74.6
WYH1573	132	29.1	161	58.2	102	69.5
QY982	131	28.8	160	62.0	109	75.8
YLY6	128	24.7	153	60.8	110	75.4
HHZ	121	28.7	149	57.1	100	71.3
GLY5	125	19.8	144	56.8	108	72.8
HY73	111	47.0	158	55.6	103	70.1
ZZ14	125	29.5	154	52.2	95	70.7
YG29	102	44.5	146	35.2	92	50.9
LSD (0.05)	29	36.5	18.4	5.0	7.6	4.5
2015						
YY4949	164	36.0	200	56.4	94	70.8
HY549	167	24.9	192	54.1	106	62.2
CY5727	177	32.8	210	49.9	97	62.8
JY225	164	19.5	184	49.5	89	63.5
YY278	161	34.4	195	53.1	104	63.8
LLYHZ	159	23.5	182	52.2	105	63.5
HLY630	155	21.9	177	58.4	103	70.5
MLY143	167	20.4	187	54.0	100	66.7
LY10	156	37.1	193	50.3	94	64.3
HHZ	165	12.0	177	47.3	94	62.5
JY959	175	17.7	193	51.6	101	65.9
ZH1	187	22.3	209	45.8	100	57.8
HLY7185	150	27.1	177	49.4	94	62.8
HH1709	159	−0.7	158	52.1	101	62.6
SY108	162	21.1	183	52.0	99	62.5
YLY6	174	2.4	177	51.7	105	64.8
HLY1511	153	19.8	172	50.5	99	63.8
WSSM	163	24.6	188	43.1	92	57.5
LSD (0.05)	28	34.4	30	5.2	7.9	3.9

stage, while the grain N concentration was the highest at maturity only in 2014. At the maturity stage in 2015, the N concentration in the leaf was similar to that in the grain. However, HY549 had the lowest leaf N concentration at maturity in both 2014 and 2015, while YY4949 had the highest N concentration in the leaf and stem at the heading stage in 2015 (**Table 5**).

Relationship between NUE and Growth Traits

The data for the varieties in 2014 and 2015 were used for correlation analyses to examine the relationship between the NUE-related parameters and growth analyses (**Figures 2**, **3**; **Table 6**). A significant quadratic relationship was observed between the grain yield and TN$_{PM}$, demonstrating that the grain yield was augmented with an increase in TN$_{PM}$ until 180 kg ha^{-1} (**Figure 2A**). The significant quadratic relationship between the grain yield and NUEg revealed that improvements in NUEg had no influence on the grain yield when NUEg was higher than 45 kg kg^{-1} (**Figure 2B**). Improvements in the N uptake capacity and NUEg were accomplished through breeding for high biomass production and HI, respectively (**Figures 2C,D**). No significant relationship was observed between NUEg and TN$_{PM}$ (**Figure 3A**). NUEg was significantly and positively correlated with NUEb and NHI, but was negatively correlated with the N concentration in the grain, leaf, and stem (**Figures 3B–F**). The quadratic relationship between the NUEg and N concentration in leaf and stem suggested that improvements in NUEg were dependent on a decreased leaf N concentration when the NUEg value was lower than 50 kg kg^{-1}, while further improvements in NUEg were dependent on a decreased stem N concentration (**Figures 3E,F**).

Many growth traits were significantly and positively correlated with TN$_{PM}$, such as the total growth duration, daily grain yield, plant height at heading, leaf area index at heading, crop growth rate before heading, biomass at heading, and grain filling percentage. However, the biomass during the grain-filling period and the panicle N concentration at heading were negatively correlated with TN$_{PM}$ (**Table 6**). A significant and positive correlation was observed between the NUEg and biomass during the grain-filling period, spikelets per panicle, and daily grain yield. Most of the growth parameters affected the NUEb or NHI (**Table 6**).

DISCUSSION

Intervarietal Difference in NUE

Variations in rice NUE have been studied since the research by Broadbent et al. (1987), who reported significant differences in the NUE of 24 rice genotypes at IRRI. Thereafter, many studies were conducted to examine the rice NUE, showing that TN$_{PM}$ and NUEg ranged from 48 to 130 kg ha^{-1} and 35 to 79 kg kg^{-1}, respectively, under irrigated lowland conditions (Tirol-Padre et al., 1996; Singh et al., 1998). Under rainfed lowland conditions, Inthapanya et al. (2000) showed that TN$_{PM}$ ranged from 25.7 to 40.4 kg ha^{-1} and NUEg from 55.1 to 83.8 kg kg^{-1} for 16 genotypes under a N fertilizer rate of 60 kg ha^{-1}. Under Mediterranean direct water-seeded conditions, Koutroubas and Ntanos (2003) observed that the NUEg ranged from 60.9 to 90.9 kg kg^{-1} for two *indica* and three *japonica* rice varieties at an N fertilizer rate of 150 kg ha^{-1}. In wheat, significant differences in NUE have been examined (Le Gouis et al., 2000). The TN$_{PM}$ and NUEg in wheat ranged from 31 to 264 kg ha^{-1} and 27 to 77 kg kg^{-1}, respectively, depending on the N rate, variety, and

TABLE 5 | Nitrogen concentration in various plant organs of the varieties at the heading stage and maturity in 2014 and 2015 at Wuxue County, Hubei Province, China.

Variety	HD			PM			
	Nleaf	Nstem	Npanicle	Nleaf	Nstem	Ngrain	Nother
2014							
HY549	2.65	0.85	1.62	0.76	0.52	1.17	0.81
HLY858	2.78	0.88	1.70	1.00	0.49	1.29	0.85
JKY651	2.67	0.84	1.68	1.21	0.49	1.25	0.75
9Y6H	2.63	0.89	1.39	0.91	0.44	1.24	0.77
HLY630	2.75	0.91	1.45	0.96	0.45	1.31	0.89
RFY41	2.55	0.87	1.56	1.09	0.54	1.29	0.84
RY225	2.51	0.80	1.38	0.86	0.51	1.28	0.90
WYH1573	2.65	0.90	1.54	1.19	0.58	1.20	0.63
QY982	2.62	0.77	1.44	0.90	0.40	1.22	0.73
YLY6	2.41	0.80	1.40	0.84	0.41	1.24	0.74
HHZ	2.65	0.86	1.64	1.16	0.52	1.25	0.82
GLY5	2.31	0.69	1.36	0.88	0.44	1.28	0.86
HY73	2.58	0.77	1.42	1.23	0.44	1.26	0.74
ZZ14	2.81	0.95	1.41	1.07	0.56	1.36	0.90
YG29	2.45	0.68	1.36	1.58	0.56	1.45	1.13
LSD (0.05)	0.45	0.24	0.36	0.19	0.13	0.08	0.08
2015							
YY4949	3.19	1.12	1.28	1.29	0.58	1.26	0.70
HY549	2.49	0.73	1.25	1.17	0.60	1.15	0.93
CY5727	2.72	0.82	1.26	1.29	0.66	1.26	0.87
JY225	2.94	1.00	1.54	1.41	0.81	1.28	0.87
YY278	2.27	0.70	1.16	1.18	0.54	1.20	0.77
LLYHZ	2.73	0.75	1.27	1.16	0.57	1.22	0.76
HLY630	2.59	0.85	1.31	1.22	0.52	1.21	0.70
MLY143	2.47	0.81	1.28	1.27	0.54	1.24	0.68
LY10	2.62	0.79	1.24	1.35	0.65	1.28	0.80
HHZ	2.93	0.88	1.55	1.46	0.57	1.32	0.89
JY959	2.64	0.79	1.36	1.19	0.55	1.28	0.87
ZH1	2.65	0.78	1.31	1.48	0.56	1.27	0.72
HLY7185	2.62	0.98	1.32	1.37	0.68	1.27	0.75
HH1709	2.93	0.96	1.26	1.23	0.62	1.20	0.85
SY108	2.64	0.83	1.29	1.42	0.65	1.20	0.69
YLY6	2.50	0.76	1.33	1.15	0.53	1.25	0.78
HLY1511	2.61	0.83	1.37	1.24	0.59	1.27	0.71
WSSM	2.77	0.85	1.30	1.28	0.78	1.33	0.81
LSD (0.05)	0.25	0.18	0.11	0.15	0.10	0.11	0.12

TABLE 6 | Correlation between NUE-related parameters and growth-related parameters.

Parameters	Total N uptake	NUEg	NUEb	NHI
Total growth duration	0.54**	0.17 ns	0.61**	−0.38*
Daily grain yield	0.36*	0.46**	0.30 ns	−0.10 ns
Stem No. m^{-2} at heading	0.20 ns	−0.41*	0.44**	−0.60**
Plant height at heading	0.39*	0.12 ns	0.26 ns	−0.06 ns
Leaf area index at heading	0.80**	−0.27 ns	0.82**	−0.76**
Specific leaf weight at heading	0.08 ns	−0.30 ns	−0.01 ns	−0.08 ns
Crop growth rate before heading	0.79**	−0.18 ns	0.77**	−0.67**
Biomass at heading	0.70**	−0.11 ns	0.72**	−0.56**
Biomass during grain filling period	−0.79**	0.58**	−0.89**	0.99**
Panicle No m^{-2}	0.29 ns	−0.37*	0.49**	−0.62**
Spikelets per panicle	−0.07 ns	0.42*	−0.34 ns	0.45*
Spikelets m^{-2}	0.17 ns	−0.14 ns	0.05 ns	−0.06 ns
Grain filling percentage	0.59**	0.24 ns	0.72**	−0.48**
Grain weight	0.01 ns	0.26 ns	−0.01 ns	0.20 ns
Leaf N concentration at heading	0.23 ns	−0.09 ns	0.12 ns	−0.23 ns
Stem N concentration at heading	0.07 ns	0.14 ns	0.01 ns	−0.06 ns
Panicle N concentration at heading	−0.45**	0.31 ns	−0.63**	0.61**

* and **indicate significance at 0.05 and 0.01 probability level.

of Tirol-Padre et al. (1996) and Singh et al. (1998), but was lower than the findings of Koutroubas and Ntanos (2003). Koutroubas and Ntanos (2003) reported a grain yield ranging from 6.0 to 8.3 t ha^{-1}, thus the relatively high NUEg reflected a lower TN$_{PM}$, which ranged from 76.2 to 124.2 kg ha^{-1}. Notably, the grain yield in the present study was significantly higher than the values reported in all of the previous studies, indicating that it is feasible to simultaneously achieve high yield and high efficiency.

Relationship between Grain Yield and NUE

Broadbent et al. (1987) evaluated the stability of nine NUE-related parameters using the N^{15} labeling method to rank the genotypes across different seasons, and De Datta and Broadbent (1988) further tested these methods without using isotopically labeled fertilizer to reflect genotypic variations in NUE. These studies showed that the yield and GW/Nt were the most stable parameters reflecting genotypic differences in the NUE of rice. In the present study, the grain yield and daily grain yield were significantly and positively correlated with both TN$_{PM}$ and NUEg (**Figure 2**, **Table 6**; Samonte et al., 2006). This finding is consistent with the evidence that genetic improvements in the yield potential improve both TN$_{PM}$ and NUEg (Fischer, 1981; Bingham et al., 2012; Zhu et al., 2016). However, the correlations between the grain yield and TN$_{PM}$ and between the grain yield and NUEg were quadratic (**Figures 2A,B**; Cassman et al., 1993; Singh et al., 1998). This finding indicated that the increase in grain yield through an increase in TN$_{PM}$ is marginal when TN$_{PM}$ is higher than 150 kg ha^{-1}, and this increase is likely to improve NUEg while maintaining a high grain yield.

Plant Traits Related with NUE

The N uptake efficiency accounted for 64% of the variation in the NUE at zero N rate, while the NUEg was more significant

year (Barraclough et al., 2010; Gaju et al., 2011; Bingham et al., 2012).

In a previous study, we observed that TN$_{PM}$ ranged from 138 to 248 kg ha^{-1}, and NUEg ranged from 28.8 to 58.4 kg kg^{-1} for 14 rice mega varieties developed at different ages. Similarly, both TN$_{PM}$ and NUEg were significantly enhanced through the advancements in genetic breeding (Zhu et al., 2016). In the present study, the TN$_{PM}$ of the elite varieties ranged from 144 to 210 kg ha^{-1}, which was higher than the values observed for rice at a similar N rate, as previously discussed. NUEg ranged from 35.2 to 60.9 kg kg^{-1}, which is consistent with the findings

FIGURE 2 | Correlation between total N uptake at maturity and grain yield (A), NUEg and grain yield (B), biomass at maturity and total N uptake (C), and harvest index (HI) and NUEg (D).

at a higher N rate (Le Gouis et al., 2000). Gaju et al. (2011) also demonstrated the association between the N uptake efficiency and showed that NUE increased with increasing N limitation. Thus, breeders should select varieties with a high N uptake efficiency for low-yield crops, and varieties with high NUEg for high-yield crops, although it is possible to simultaneously improve the N uptake efficiency and NUEg (**Figure 3A**; Moll et al., 1982). The following plant traits were associated with TN_{PM} and NUEg.

TN_{PM} could be estimated from primary plant parameters, as this measurement was significantly correlated with the tiller number, spikelet number, main culm panicle node number reflecting the potential tillers and leaves of a plant (Singh et al., 1998; Samonte et al., 2006). Moreover, Singh et al. (1998) observed that varieties with long growth durations had higher TN_{PM} values compared with varieties with medium growth durations. In addition, deeper roots and greater root oxidation activities are important for N uptake at low N rates in both rice and wheat (Foulkes et al., 2009; Worku et al., 2012; Ju et al., 2015). In the present study, TN_{PM} was significantly and positively correlated with total growth duration, plant height at heading, leaf area index at heading, crop growth rate before

heading, biomass at heading, and grain filling percentage, but it was negatively correlated with biomass accumulation during the grain filling period and panicle N concentration at heading (**Table 6**). Consequently, genetically promoting plant growth prior to heading is important for improvements in the TN_{PM} at low N rates.

N utilization efficiency is dependent on the N efficiency of biomass formation, the effect of N on carbohydrate partitioning, nitrate reduction efficiency, and remobilization of N from senescent tissues and storage functions (Foulkes et al., 2009). NUEg was significantly correlated with HI (**Figure 2D**), as HI was positively and significantly correlated with the dry matter translocation efficiency (Ntanos and Koutroubas, 2002). Mathematically, NUEg is equal to the ratio of the NHI and grain N concentration; thus, the NUEg was positively and significantly correlated with the NHI but was negatively correlated with the grain N concentration (**Figures 3C,D**). In the present study, the NHI ranged from 57.5 to 75.0%, which is consistent with the values reported in the studies of Tirol-Padre et al. (1996) and Singh et al. (1998). Consequently, it might be possible to further increase the NHI of rice to some extent (Sinclair and Vadez, 2002). Significant negative correlations between the grain
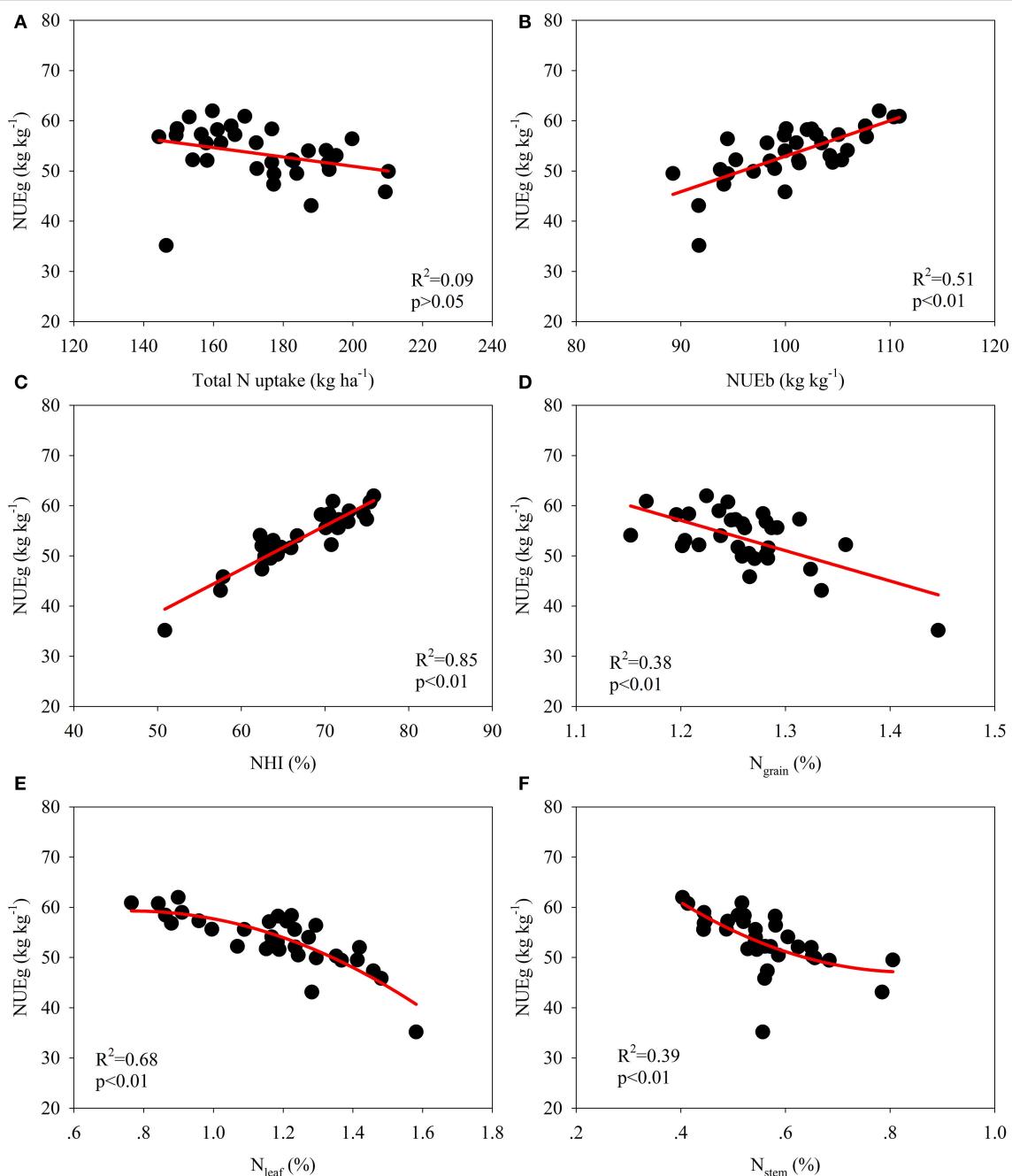
FIGURE 3 | Correlation among NUE-related parameters and their relationship with the N concentration in various plant organs at maturity. (A) Correlation between total N uptake and NUEg, **(B)** Correlation between NUEb and NUEg, **(C)** Correlation between NHI and NUEg, **(D)** Correlation between N grain and NUEg, **(E)** Correlation between N leaf and NUEg, **(F)** Correlation between N stem and NUEg.

N concentration and NUEg have been widely demonstrated among different genotypes in rice and wheat (Singh et al., 1998; Inthapanya et al., 2000; Koutroubas and Ntanos, 2003). Moreover, Cassman et al. (1993) demonstrated that a lower N content grain in rice than that in bread wheat contributes to a higher NUEg in rice, particularly at high yield levels. The straw N concentration explained a large percentage of the genotypic variation in NUEg in the studies of Singh et al. (1998) and

Koutroubas and Ntanos (2003). In the present study, we further demonstrated that the variation in NUEg was dependent on the changes in the leaf N concentration at maturity at low NUEg levels, while further increases in NUEg resulted from decreases in the stem N concentration (**Figures 3E,F**). These results are consistent with the findings in wheat, suggesting that delayed leaf senescence is a key trait for increasing NUEg at low N supply (Foulkes et al., 2009; Gaju et al., 2011). Moreover, NUEg was

significantly and positively correlated with biomass accumulation during the grain-filling period, spikelets per panicle and daily grain yield (**Table 6**).

In conclusion, the present study determined the genotypic variation in NUE among newly developed elite rice varieties in China and demonstrated that genetic improvements in the yield potential under high nutrient input conditions also increased the TN_{PM} and NUEg at a low N supply. The quadratic correlation between the grain yield and TN_{PM} and between the grain yield and NUEg suggests that a further increase in N uptake results in a small increase in grain yield when TN_{PM} is above $160 \, kg \, ha^{-1}$, and it is possible to simultaneously achieve a high grain yield and high NUEg under low N supply. Improvements in the NUE are likely to occur with simultaneous increases in TN_{PM} and NUEg through the improvements in the daily grain yield. Plant traits associated with the rapid crop growth rate prior to heading could be used to increase TN_{PM}, while biomass accumulation and a large panicle are essential for improvements in NUEg. Moreover, further improvements in NUEg depend on the increase in the translocation of N from the stems to delay leaf senescence during the grain-filling period.

AUTHOR CONTRIBUTIONS

LW, SY, LH, and FS conducted the field experiments. SP and FW designed the experiments. SF revised the manuscript. LW analyzed the data, and FW drafted the manuscript.

FUNDING

REFERENCES

Andrews, M., Raven, J. A., and Lea, P. J. (2013). Do plants need nitrate? The mechanisms by which nitrogen form affects plants. *Ann. Appl. Biol.* 163, 174–199. doi: 10.1111/aab.12045

Barraclough, P. B., Howarth, J. R., Jones, J., Lopez-Bellido, R., Parmar, S., Shepherd, C. E., et al. (2010). Nitrogen efficiency of wheat: genotypic and environmental variation and prospects for improvement. *Eur. J. Agron.* 33, 1–11. doi: 10.1016/j.eja.2010.01.005

Bingham, I. J., Karley, A. J., White, P. J., Thomas, W. T. S., and Russell, J. R. (2012). Analysis of improvements in nitrogen use efficiency associated with 75 years of spring barley breeding. *Eur. J. Agron.* 42, 49–58. doi: 10.1016/j.eja.2011.10.003

Broadbent, F. E., De Datta, S. K., and Laureles, E. V. (1987). Measurement of nitrogen utilization efficiency in rice genotypes. *Agron. J.* 79, 786–791. doi: 10.2134/agronj1987.00021962007900050006x

Cassman, K. G. (1999). Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. *Proc. Natl. Acad. Sci. U.S.A.* 96, 5952–5959. doi: 10.1073/pnas.96.11.5952

Cassman, K. G., Kropff, M. J., Gaunt, J., and Peng, S. (1993). Nitrogen use efficiency of rice reconsidered: what are the key constraints. *Plant Soil* 155/156, 359–362. doi: 10.1007/BF00025057

De Datta, S. K., and Broadbent, F. E. (1988). Methodology for evaluating nitrogen utilization efficiency by rice genotypes. *Agron. J.* 80, 793–798. doi: 10.2134/agronj1988.00021962008000050020x

De Datta, S. K., and Broadbent, F. E. (1993). Development changes related to nitrogen-use efficiency in rice. *Field Crops Res.* 34, 47–56. doi: 10.1016/0378-4290(93)90110-9

De Datta, S. K., and Buresh, R. J. (1989). Integrated nitrogen management in irrigated rice. *Adv. Soil Sci.* 10, 143–169. doi: 10.1007/978-1-4613-8847-0_4

Diaz, R. J., and Rosenberg, R. (2008). Spreading dead zones and consequences for marine ecosystems. *Science* 321, 926–929. doi: 10.1126/science.1156401

Dobermann, A., Witt, C., Dawe, D., Gines, H. C., Nagarajan, R., Satawathananont, S., et al. (2002). Site-specific nutrient management for intensive rice cropping systems in Asia. *Field Crops Res.* 74, 37–66. doi: 10.1016/S0378-4290(01)00197-6

Evans, L. T., and Fischer, R. A. (1999). Yield potential: its definition, measurement, and significance. *Crop Sci.* 39, 1544–1551. doi: 10.2135/cropsci1999.3961544x

Fischer, R. A. (1981). Optimization the use of water and nitrogen through breeding for crops. *Plant Soil* 58, 249–278. doi: 10.1007/BF02180056

Foulkes, M. J., Hawkesford, M. J., Barraclough, P. B., Holdsworth, M. J., Kerr, S., Kightley, S., et al. (2009). Identifying traits to improve the nitrogen economy of wheat: recent advances and future prospects. *Field Crops Res.* 114, 329–342. doi: 10.1016/j.fcr.2009.09.005

Gaju, O., Allard, V., Martre, P., Snape, J. W., Heumez, E., Le Gouis, J., et al. (2011). Identification of traits to improve the nitrogen-use efficiency of wheat genotypes. *Field Crops Res.* 123, 139–152. doi: 10.1016/j.fcr.2011.05.010

Godfray, H. C., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., et al. (2010). Food security: the challenge of feeding 9 billion people. *Science* 327, 812–818. doi: 10.1126/science.1185383

Guo, J. H., Liu, X. J., Zhang, Y., Shen, J. L., Han, W. X., Zhang, W. F., et al. (2010). Significant acidification in major Chinese croplands. *Science* 327, 1008–1010. doi: 10.1126/science.1182570

Hu, B., Wang, W., Ou, S., Tang, J., Li, H., and Che, R. (2015). Variation in NRT1.1B contributes to nitrate-use divergence between rice subspecies. *Nat. Genet.* 47, 834–838. doi: 10.1038/ng.3337

Inthapanya, P., Sipaseuth, Sihavong, P., Sihathep, V., Chanphengsay, M., Fukai, S., et al. (2000). Genotype differences in nutrient uptake and utilization for grain yield production of rainfed lowland rice under fertilized and non-fertilized conditions. *Field Crops Res.* 65, 57–68. doi: 10.1016/S0378-4290(99)00070-2

Ju, C., Buresh, R. J., Wang, Z., Zhang, H., Liu, L., Yang, J., et al. (2015). Root and shoot traits for rice varieties with higher grain yield and higher nitrogen use efficiency at lower nitrogen rates application. *Field Crops Res.* 175, 47–55. doi: 10.1016/j.fcr.2015.02.007

Koning, N. B. J., Van Ittersum, M. K., Becx, G. A., Van Boekel, M., Brandenburg, W., Van Den Broek, J. A., et al. (2008). Long-term global availability of food: continued abundance or new scarcity? *Wagen. J. Life Sci.* 55, 229–292. doi: 10.1016/s1573-5214(08)80001-2

Koutroubas, S. D., and Ntanos, D. A. (2003). Genotypic differences for grain yield and nitrogen utilization in Indica and Japonica rice under Mediterranean conditions. *Field Crops Res.* 83, 251–260. doi: 10.1016/S0378-4290(03)00067-4

Le Gouis, J., Beghin, D., Heumez, E., and Pluchard, P. (2000). Genetic differences for nitrogen uptake and nitrogen utilisation efficiencies in winter wheat. *Eur. J. Agron.* 12, 163–173. doi: 10.1016/S1161-0301(00)00045-9

Li, J., Zhang, Y., Wu, J., Shu, Y., Zhou, P., and Deng, Q. (2014). High-yielding cultural techniques of super hybrid rice YLY900 yielded 15.40 t/hm² on a 6.84 hm² scale. *China Rice* 20, 1–4.

Moll, R. H., Kamprath, E. J., and Jackson, W. A. (1982). Analysis and interpretation of factors which contribute to efficiency of nitrogen-utilization. *Agron. J.* 74, 562–564. doi: 10.2134/agronj1982.00021962007400030037x

Ntanos, D. A., and Koutroubas, S. D. (2002). Dry matter and N accumulation and translocation for Indica and Japonica rice under Mediterranean conditions. *Field Crops Res.* 74, 93–101. doi: 10.1016/S0378-4290(01)00203-9

Peng, S., Buresh, R. J., Huang, J., Yang, J., Zou, Y., Zhong, X., et al. (2006). Strategies for overcoming low agronomic nitrogen use efficiency in irrigated rice systems in China. *Field Crops Res.* 96, 37–47. doi: 10.1016/j.fcr.2005.05.004

Peng, S., Garcia, F. V., Laza, R. C., Sanico, A. L., Visperas, R. M., and Cassman, K. G. (1996). Increased N-use efficiency using a chlorophyll meter on high-yielding irrigated rice. *Field Crops Res.* 47, 243–252. doi: 10.1016/0378-4290(96)00018-4

Peng, S., Huang, J., Sheehy, J. E., Laza, R. C., Visperas, R. M., Zhong, X. H., et al. (2004). Rice yields decline with higher night temperature from global warming. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9971–9975. doi: 10.1073/pnas.0403720101

Peng, S., Huang, J., Zhong, X., Yang, J., Wang, G., Zou, Y., et al. (2002). Research strategy in improving fertilizer-nitrogen use efficiency of irrigated rice in China. *Sci. Agric. Sin.* 35, 1095–1103.

Peng, S., Khush, G. S., Virk, P., Tang, Q., and Zou, Y. (2008). Progress in ideotype breeding to increase rice yield potential. *Field Crops Res.* 108, 32–38. doi: 10.1016/j.fcr.2008.04.001

Samonte, S., Wilson, L. T., Medley, J. C., Pinson, S. R. M., McClung, A. M., and Lales, J. S. (2006). Nitrogen utilization efficiency: relationships with grain yield, grain protein, and yield-related traits in rice. *Agron. J.* 98, 168–176. doi: 10.2134/agronj2005.0180

Sheng, C., Wang, H., Sheng, S., Gao, L., and Xuan, W. (2003). Pest status and loss assessment of crop damage caused by the rice borers, *Chilo suppressalis and Tryporyza incertulas* in China. *Entomol. Knowledge* 40, 289–294.

Sinclair, T. R., and Vadez, V. (2002). Physiological traits for crop yield improvement in low N and P environments. *Plant Soil* 245, 1–15. doi: 10.1023/A:1020624015351

Singh, U., Ladha, J. K., Castillo, E. G., Punzalan, G., Tirol-Padre, A., and Duqueza, M. (1998). Genotypic variation in nitrogen use efficiency in medium- and long-duration rice. *Field Crops Res.* 58, 35–53. doi: 10.1016/S0378-4290(98)00084-7

Smil, V. (1999). Nitrogen in crop production: an account of global flows. *Global Biogeochem. Cycles* 13, 647–662. doi: 10.1029/1999GB900015

Sun, H., Qian, Q., Wu, K., Luo, J., Wang, S., and Zhang, C. (2014). Heterotrimeric G proteins regulate nitrogen-use efficiency in rice. *Nat. Genet.* 46, 652–657. doi: 10.1038/ng.2958

Tao, F., Zhang, S., and Zhang, Z. (2013). Changes in rice disasters across China in recent decades and the meteorological and agronomic causes. *Reg. Environ. Change* 13, 743–759. doi: 10.1007/s10113-012-0357-7

Tirol-Padre, A., Ladha, J. K., Singh, U., Laureles, E., Punzalan, G., and Akita, S. (1996). Grain yield performance of rice genotypes at suboptimal levels of soil N as affected by N uptake and utilization efficiency. *Field Crops Res.* 46, 127–143. doi: 10.1016/0378-4290(95)00095-X

Vlek, P. L. G., and Byrnes, B. H. (1986). The efficacy and loss of fertilizer N in lowland rice. *Fertil. Res.* 9, 131–147. doi: 10.1007/BF01048699

Worku, M., Banziger, M., Erley, G., Friesen, D., Diallo, A. O., and Horst, W. J. (2012). Nitrogen efficiency as related to dry matter partitioning and root system size in tropical mid-altitude maize hybrids under different levels of nitrogen stress. *Field Crops Res.* 130, 57–67. doi: 10.1016/j.fcr.2012.02.015

Zhang, Q. (2007). Strategies for developing green super rice. *Proc. Natl. Acad. Sci. U.S.A.* 104, 16402–16409. doi: 10.1073/pnas.0708013104

Zhong, X., Huang, R., Zheng, H., Peng, S., and Buresh, R. J. (2007). Specification for the "Three Controls" nutrient management technology for irrigated rice. *Guangdong Agric. Sci.* 5s, 13–15.

Zhu, G., Peng, S., Huang, J., Cui, K., Nie, L., and Wang, F. (2016). Genetic improvements in rice yield and concomitant increases in radiation- and nitrogen-use efficiency in Middle Reaches of Yangtze River. *Sci. Rep.* 6:21049. doi: 10.1038/srep21049

Zou, Y., Ao, H., Wang, S., and Tang, Q. (2006). Studies on San-Ding cultivation method for super rice: the concept and the principle. *Chin. Agric. Sci. Bull.* 22, 158–162.

Overexpression of *OsDof12* affects plant architecture in rice (*Oryza sativa* L.)

Qi Wu [1,2†], Dayong Li [2†], Dejun Li [3], Xue Liu [2,4], Xianfeng Zhao [2], Xiaobing Li [2], Shigui Li [1*] and Lihuang Zhu [2*]

[1] Rice Research Institute, Sichuan Agricultural University, Chengdu, China, [2] State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China, [3] Key Laboratory of Biology and Genetic Resources of Rubber Tree, Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences, Danzhou, China, [4] Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China

Dof (DNA binding with one finger) proteins, a class of plant-specific transcription factors, are involved in plant growth and developmental processes and stress responses. However, their biological functions remain to be elucidated, especially in rice (*Oryza sativa* L.). Previously, we have reported that *OsDof12* can promote rice flowering under long-day conditions. Here, we further investigated the other important agronomical traits of the transgenic plants overexpressing *OsDof12* and found that overexpressing *OsDof12* could lead to reduced plant height, erected leaf, shortened leaf blade, and smaller panicle resulted from decreased primary and secondary branches number. These results implied that *OsDof12* is involved in rice plant architecture formation. Furthermore, we performed a series of Brassinosteroid (BR)-responsive tests and found that overexpression of *OsDof12* could also result in BR hyposensitivity. Of note, in WT plants the expression of *OsDof12* was found up-regulated by BR treatment while in *OsDof12* overexpression plants two positive BR signaling regulators, *OsBRI1* and *OsBZR1*, were significantly down-regulated, indicating that *OsDof12* may act as a negative BR regulator in rice. Taken together, our results suggested that overexpression of *OsDof12* could lead to altered plant architecture by suppressing BR signaling. Thus, *OsDof12* might be used as a new potential genetic regulator for future rice molecular breeding.

Keywords: *OsDof12*, Dof transcription factor, plant architecture, rice (*Oryza sativa* L.)

Introduction

DOF (DNA binding with one finger) proteins are plant-specific transcription factors (Yanagisawa, 2002; Moreno-Risueno et al., 2007; Noguero et al., 2013). The Dof domain consists of 52 amino acid residues encompassing the $CX_2CX_{21}CX_2C$ motif (Yanagisawa, 2002; Umemura et al., 2004). Dof transcription factors, with the exception in pumpkin, usually regulate the expression of the target genes *via* binding a core DNA motif with an essential sequence of AAAG (Yanagisawa and Schmidt, 1999). Dof proteins are widespread and versatile regulators for various biological processes such as metabolism regulation, phytohormone response, seed germination and development, photoperiodic flowering and plant patterning in plants (Yanagisawa, 2002; Noguero et al., 2013). In carbohydrate metabolism, *ZmDof1* and *ZmDof2* acted antagonistically to control the

expression of *C4 phosphoenolpyruvate carboxylase* (*PEPC*) in maize (*Zea mays*) (Yanagisawa, 2000). In Arabidopsis (*Arabidopsis thaliana*), *AtDof1.1/OBP2* participated in regulation of indole glucosinolate biosynthesis (Skirycz et al., 2006). Overexpression of *OsDof25* in Arabidopsis changed the nitrogen and carbon metabolism (Santos et al., 2012). In tobacco (*Nicotiana tabacum*), NtBBF1 (roiB domain B Factor) can bind to the *rolB* promotor in an auxin-regulated way to modulate its expression, which betters our understanding the mechanism underlying auxin induction (Baumann et al., 1999). Besides, several cases have documented that *Dof* genes are implicated in seed germination and development. The *DOF gene Affecting Germination-1* (*DAG1*) and *DOF gene Affecting Germination-2* (*DAG2*) controlled seed germination in Arabidopsis *via* a maternal switch (Papi et al., 2000; Gualberti et al., 2002). In rice, *OsDof3* might enhance expression of *type3 carboxypeptidase* (*CPD*) under GA control in aleurones (Washio, 2001), and further investigation indicated that OsDof3 interacts with GAMYB to regulate synergically the expression of *RAmy1A* to mediate GA signaling during rice seed germination (Washio, 2003). RPBF (rice prolamin box binding factor) interplayed with the rice basic leucine zipper factor RISBZ1 to maintain proper expression of rice seed storage protein genes during seed development (Kawakatsu et al., 2009).

Moreover, Dof factors are involved in photoperiod flowering. In *Arabidopsis*, Cycling Dof Factor-1 (CDF1) binds to the *COSTANS* (*CO*) and *FLOWERING LOCUS T* (*FT*) promotor regions to block transactivation of this two flowering genes, whereas this inhibition could be released based on the GIGANTEA-FLAVIN-BINDING, KELCH REPEAT, F-BOX1(GI-FKF1) complex mediated degradation of CDF1 under long-day conditions (Imaizumi et al., 2005; Sawa et al., 2007; Song et al., 2012). Fornara et al. (2009) systematically studied a subset of *Dof* family related to *CDF1* and found that *CDF1*, *CDF2*, *CDF3,* and *CDF5* acted redundantly to repress flowering by decreasing the mRNA level of *CO* (Fornara et al., 2009). In rice, *RDD1* (*rice Dof daily fluctuations 1*) was controlled by circadian clock, and repressing the expression of *RDD1* led to delayed flowering time (Iwamoto et al., 2009). Furthermore, several studies have revealed the importance of Dof factors on plant patterning. In Arabidopsis, AtDof5.1 modulated leaf adaxial–abaxial polarity *via* binding to the promotor of *Revoluta* (*REV*) and enhancing expression of *REV* (Kim et al., 2010). *AtDOF4.2* and *AtDOF4.4* were engaged in regulating shoot branching and seed development (Skirycz et al., 2007; Zou et al., 2013). *OBF-binding factor-1*(*OBP1*) is a cell cycle regulator, and overexpressing *OBP1* in Arabidopsis affects cell size and number, rendering dwarfish plant morphology (Skirycz et al., 2008).

To now, of the 36 predicted *Dof* homologous genes in whole Arabidopsis genome, 16 have been confirmed to participate in various biological processes (Noguero et al., 2013). However, in rice, to our knowledge, the studies deciphering Dof factors are rather limited; only 5 of the 30 predicted *Dof* members have been characterized in detail (Washio, 2003; Yamamoto et al., 2006; Li et al., 2008, 2009b; Iwamoto et al., 2009; Santos et al., 2012). In a previous study, we characterized the function of *OsDof12* in rice (Li et al., 2008, 2009b). We found

a pair of sense-antisense transcript at the locus of *OsDof12* (LOC_Os03g07360), denoted as *OsDof12* (sense transcript) and *OsDof12os* (antisense transcript), respectively. *OsDof12* encodes a nuclear-localized protein of 440 amino acids (Li et al., 2008). Expression pattern analysis showed *OsDof12* and *OsDof12os* were co-expressed but reciprocally regulated by each other (Li et al., 2008). Moreover, overexpression of *OsDof12* promoted flowering in rice under long-day conditions by up-regulating rice florigen-encoding gene *Hd3a* and its downstream gene *OsMADS14* (Li et al. 2009b). In wild type rice plants, the transcripts of *OsDof12* were detected in various rice tissues at different development stages (Li et al., 2009b), which suggests *OsDof12* might take part in various biological processes.

In current study, we demonstrated that the transgenic plants overexpressing *OsDof12* indeed display pleiotropic phenotypes such as reduced plant height, shortened leaf length, more erected leaf, smaller panicle size and decreased grain yield. Further molecular biological analyses indicated that these changes in plant architecture could be ascribed to attenuated BR signaling in *OsDof12* overexpression plants.

Materials and Methods

Plant Materials and Growth Conditions

To observe the influence of *OsDof12* on plant architecture, the wild type cultivar Nipponbare (*Oryza sativa* L. ssp. *japonica* cv. Nipponbare) and the previous reported two *OsDof12* overexpressing lines (Line OD2 and Line OD5, Nipponbare background) (Li et al., 2009b) were grown on the research field located in the Experimental Stations of the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, under nature field conditions. After mature stage (approximately 120 days after sowing), we measured the agronomic traits including plant height, internode length, leaf length, and panicle architecture.

For BR sensitivity tests, plants were grown on 1/2 Murashige and Skoog (MS) culture medium in a growth chamber under controlled conditions (16h/light, 30°C and 8h/dark, 26°C) for 7 or 8 days. For skotomorphogenesis analysis, the WT and OD plants were grown on 1/2 MS at 28°C in darkness for 2 weeks. For *OsDof12* induction analysis, 2-week-old WT plants grown on 96-well PCR plates were treated with 100 nM Brassinolide (BL, the most bioactive BR compound, not being synthesized in rice, Yokota, 1997; Fujioka and Yokota, 2003; Kim et al., 2008; Vriet et al., 2013) solution, then the hole plants were sampled at different time points for expression assay. For expression levels analysis on BR related genes, 7-day-old plants of WT and OD were used.

Microscopic Observation

The middle part of the second internode from mature plants were harvested and fixed with 2.5% glutaraldehyde solution overnight at 4°C, and then dehydrated in order by 40, 50, 60, 70, 80, 90, 95, and 100% ethanol solutions. The samples got dried through critical point drying by liquid carbon dioxide. The dry specimens were mounted on a stub, gold-coated with an ion

sputter coater (Hitachi, Tokyo, Japan) and then imaged with a scanning electron microscope (Hitachi, Tokyo, Japan).

BR Treatment

The seeds were dehusked and sterilized with 75% ethanol for 1 min, 3% NaClO for 25 min and then washed five times with sterile water. The seeds were grown on solid 1/2 MS medium containing a series of concentration of BL (WAKO, Japan) for 7 or 8 days. The lengths of second leaf sheath, second leaf, seedling height, root, and coleoptile were measured for further analyses.

BR Measurement

After sterilization, the seeds of WT and OD were sown and grown on 1/2 MS culture medium for 10 days, then about 1 gram of fresh seedlings were collected for measurement of Castasterone (CS, one of the biologically active BR compounds, a likely end product of brassinosteroid biosynthetic pathway in rice, Kim et al., 2008; Vriet et al., 2013) according to the method described previously (Ding et al., 2013).

Laminal Inclination Assay

The seeds were immersed for 2 days, and then the germinated seeds were sown on 1/2 MS supplied with different concentrations of BL. After incubation in a growth chamber for 8 days, the second lamina joints on WT and OD plants were imaged and measured by IMAGEJ according to the method described by Tong et al. (2009).

RNA Extraction and qRT-PCR Analysis

The samples were harvested and stored in liquid nitrogen until RNA extraction. The isolation of total RNA was performed applying Trizol reagent (Invitrogen, California, USA) with corresponding protocol. DNA digestion was accomplished by DnaseI (Takara, Japan), and first-strand cDNA was obtained by GoScript Reverse Transcription System (Promega, http://www.promega.com/). With a rice ubiquitin gene (*UBQ*, *LOC_Os03g13170*) set as the internal control, qRT-PCR was performed using EvaGreenq PCR MasterMix (Abm, Canada) on a real-time PCR System (Bio-Rad CFX96) with the specific primers listed in **Table S2**. The qRT-PCR program consists of 95°C for 3 min and 42 cycles of 95°C for 5 s, 60°C for 10 s. The relative expression level of each examined gene was quantified by a relative quantization method.

Results

Overexpression of *OsDof12* Reduces Plant Height

In our previous study, we developed *OsDof12* overexpressing lines in rice (hereafter named OD) (Li et al., 2009b). The OD lines, OD2 and OD5, and WT plants were grown in a paddy field in Beijing (40°10′N, 116°42′E) under nature field conditions. In order to investigate the effects of *OsDof12* overexpression on rice agronomic traits, we traced the growth performance of OD and WT plants and found that the plant height of OD was almost the same as WT before heading stage. However, after the plant height was stable at mature stage, on average, the WT plant height

reached approximate 83.7 cm while OD plant height was about 65.3 cm and 20% shorter than the WT (**Figures 1A,C**). Rice plant height consists of the internodes length and the panicle length. We measured and compared the internode and panicle lengths of OD plants with those of WT plants, respectively. Statistic data showed that the individual internode and panicle lengths in OD were remarkably shorter than those in WT (**Figures 1B,D**, **2B**), which consequently led to decreased plant height of OD plant. Comparing the percentage length of each internode to the total culm in OD with that in WT, we found that the second and fourth internode lengths of OD were largely and slightly reduced (**Figure 1E**), respectively.

Internode elongation is determined by cell division activity in the intercalary meristem, followed by cell elongation in the elongation zone (Yamamuro et al., 2000). To investigate whether those two factors caused the dwarfish morphology of OD, we observed the longitudinal cell morphology of the OD and WT internodes. After heading stage, we collected and fixed the middle sections of the second internodes, then observed the cell length under scanning electron microscope (SEM). As shown in **Figure 1F**, there was no obvious difference in the longitudinal cell length between WT and OD, thus the reduction in longitudinal cell number on the elongation zone may account for the shortage of the internodes in OD.

OsDof12 Overexpression Plants Produce Smaller Panicles

Besides plant height, we also analyzed the panicle structure and seed size in OD plants. As shown in **Figures 2A,B**, the panicles of OD are shorter and smaller than that of WT. We further investigated the panicle index including numbers of primary branches, secondary branches, and total spikelets, and found that the number of primary branches in OD was remarkably reduced (**Figure 2C**), meanwhile OD plants produced significantly less secondary branches (**Figure 2D**). And, expectedly, there was very significant reduction in total spikelet number per panicle in OD as compared with that in WT (**Figure 2E**). Nevertheless, no obvious differences of seed size and weight appeared between OD and wild type (**Figure S1**). These results imply that overexpression of *OsDof12* could lead to smaller panicle structure with reduced primary branches, secondary branches and total spikelets.

Abnormal Leaf Morphology in *OsDof12* Overexpression Plants

Aside from the alternations on plant height and panicle structure, OD also exhibited some other characteristic phenotypes. In comparison with WT plants, OD plants generated more compacted plant architecture with more erected leaves (**Figure 1A**). As shown in **Figure 3**, obviously the leaf joint angle on OD plants was smaller than that on WT plants. Moreover, we measured the length of flag leaves, penultimate leaves and antepenultimate leaves of OD and WT plants. Statistical analysis indicated that the leaves in OD plants were remarkably shorter than those in WT plants (**Figure S2**).
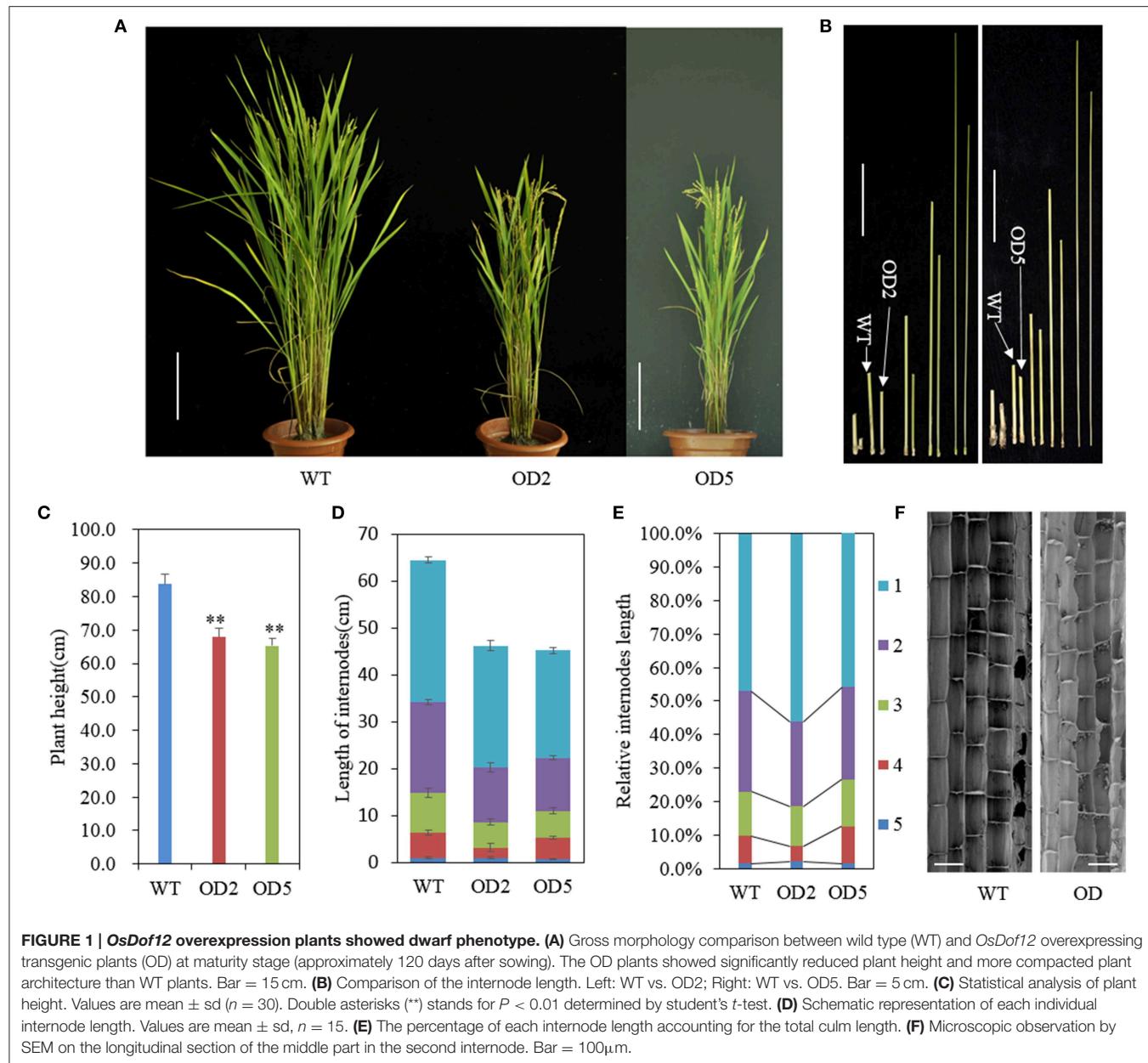
FIGURE 1 | *OsDof12* overexpression plants showed dwarf phenotype. (A) Gross morphology comparison between wild type (WT) and *OsDof12* overexpressing transgenic plants (OD) at maturity stage (approximately 120 days after sowing). The OD plants showed significantly reduced plant height and more compacted plant architecture than WT plants. Bar = 15 cm. (B) Comparison of the internode length. Left: WT vs. OD2; Right: WT vs. OD5. Bar = 5 cm. (C) Statistical analysis of plant height. Values are mean ± sd ($n = 30$). Double asterisks (**) stands for $P < 0.01$ determined by student's *t*-test. (D) Schematic representation of each individual internode length. Values are mean ± sd, $n = 15$. (E) The percentage of each internode length accounting for the total culm length. (F) Microscopic observation by SEM on the longitudinal section of the middle part in the second internode. Bar = 100μm.

Overexpression of *OsDof12* Altered BR Sensitivity

Brassinosteroids (BRs) play pivotal roles in regulating plant growth and development (Müssig and Altmann, 2001; Fujioka and Yokota, 2003; Salas Fernandez et al., 2009; Tong and Chu, 2012; Wang et al., 2012; Zhu et al., 2013). By now, many BR metabolic and signaling-related genes have been well characterized (Yamamuro et al., 2000; Hong et al., 2002, 2003, 2005; Bai et al., 2007; Tanaka et al., 2009; Tong et al., 2009, 2012; Li et al., 2009a; Sakamoto et al., 2011; Thornton et al., 2011; Zhang et al., 2012). Notably, almost all of the corresponding mutants or misexpressors display abnormal plant height and leaf inclination phenotypes. Considering that the plants overexpressing *OsDof12* exhibit reduced plant height and erect leaf morphology, we assumed that *OsDof12* might be involved in BR metabolism or signal transduction. To test this hypothesis, we designed and conducted a series of experiments to evaluate the BR response of OD2 plants.

Firstly, we measured the endogenous BR levels by quantifying the CS content in WT and OD2 plants, respectively. As a matter of fact, no obvious difference in BR levels between WT and OD plants was observed (**Table S1**), suggesting that *OsDof12* overexpression may not affect the metabolism of BRs, which was further confirmed by expression level analysis on BR metabolism related genes (**Figure S3**).

Then, we performed lamina joint bending assay by applying BL on the plants. Notably, the lamina joints of OD plants were remarkably less enlarged than that of WT when both subjected to a mock treatment without BL (**Figure 4A**). When treated with
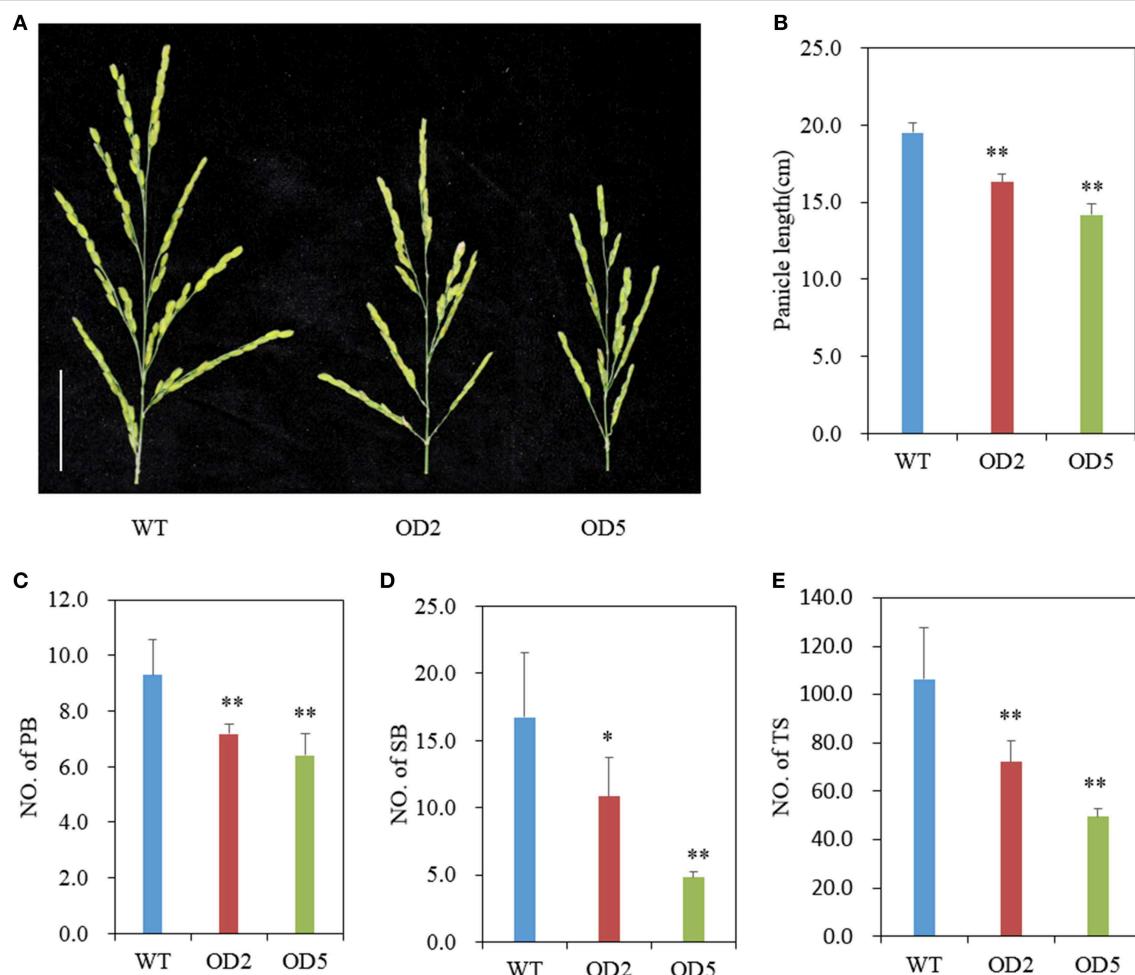
FIGURE 2 | Comparison of panicle structure. (A) Panicle morphology comparison between WT and OD plants. Bar = 5 cm. **(B-E)** Statistical results of panicle length **(B)**, number of primary branches **(C)**, secondary branches **(D)**, and total spikelets per panicle **(E)** in WT and *OsDof12* overexpression plants. PB, primary branches; SB, secondary branches; TS, total spikelets. Values are mean ± sd (*n* = 30). Single asterisk (*) and double asterisks (**) stand for *P* < 0.05 and *P* < 0.01, respectively, determined by student's *t*-test.

the increasing doses of BL, though the angles of lamina joint on both OD and WT plants became relatively larger, the lamina inclination curve of OD ascended much slower than that of WT (**Figure 4B**), indicating that overexpression of *OsDof12* may cause defects on the BR signal transduction pathway leading to impaired bending of leaves.

Besides OD plants exhibiting lamina joint abnormality, we also found that overexpression of *OsDof12* affected other aspects of BR-related morphology. For instance, the second leaf sheath length of OD plants was apparently shorter than that of WT plants. In WT plants, when treated with 1 nM BL, the second leaf sheath length slightly but not significantly increased than those in the mock treatment, however, 10 nM or higher concentrations of BL treatments gave rise to evidently shorter second leaf sheath length in WT plants (**Figure 5A**), suggesting that relative higher level of exogenous BR inhibits the elongation of the leaf sheath. In contrast, various concentrations of BL treatment on OD could hardly altered the elongation of the second

leaf sheath (**Figure 5A**), indicating that the elongation of the second leaf sheath in OD was hyposensitive to exogenous BR treatment. The similar situation also happened to the elongation of the second leaf and seedling height in WT and OD plants, respectively. The second leaf length and seedling height in OD plants was apparently shorter than those in WT. Relative higher concentrations of BL decreased evidently the elongation of the second leaf and seedling height in WT plants (**Figures 5B,C**), while various levels of BL treatment could hardly affect the elongation of the second leaf length and seedling height in OD plants (**Figures 5B,C**), indicating OD plants may have lower sensitivity in response to exogenous BR.

Elongation of the Coleoptiles and Root in Response to BL

The coleoptile length and root elongation are respects to evaluate BR sensitivity of plants (Yamamuro et al., 2000). Thus, we measured the coleoptile length of 8-day-old seedlings. As shown
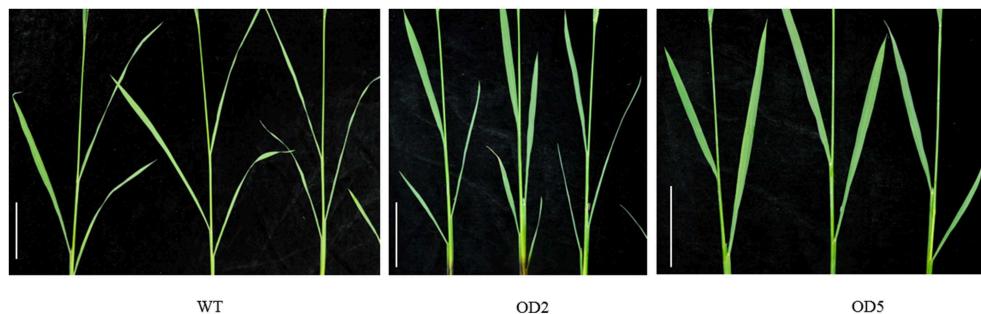
FIGURE 3 | Comparison of the lamina joint zones. The leaf angle in OD2 and OD5 plants is less enlarged than WT plants. Bar = 10 cm.
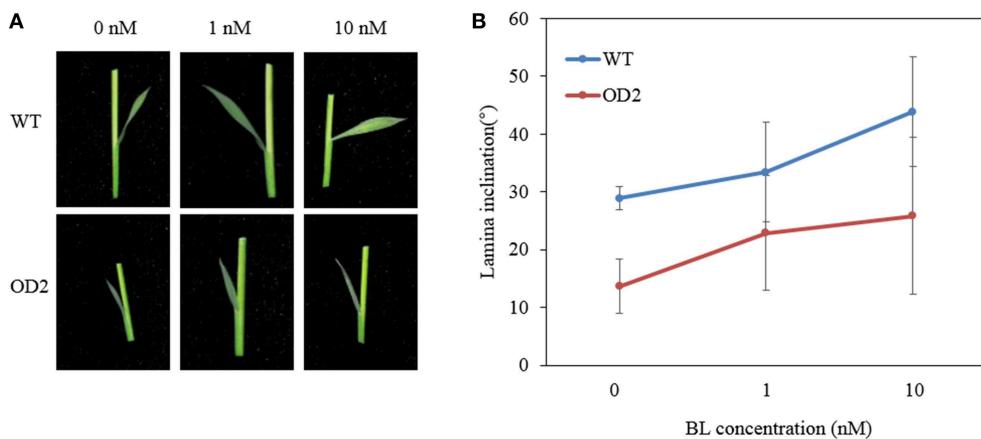
FIGURE 4 | Lamina joint bending analysis. (A) The second lamina joint morphology in WT and OD2 plants grown on BL contained half-strength MS for 7 days. **(B)** Quantification of the second lamina angel bending as shown in **(A)**, values are mean ± sd, $n = 20$.

in **Figure 5D**, the coleoptile elongation of WT was promoted by BL treatment in dose-dependent manner. However, the coleoptile elongation of OD plants showed no difference between BL and mock treatment.

Similar root growth pattern was shared by OD and WT plants when grown in 0 nM or 1 nM BL medium, 1 nM BL slightly increased the root elongation (**Figure 5E**). When treated with 10 nM or 100 nM BL, the WT plants generated apparently shorter root. However, for OD plants, 10 nM BL could not obviously reduce root elongation, only up to 100 nM BL treatment could inhibit the root elongation (**Figure 5E**). The results also suggested that overexpression of *OsDof12* may result in lower BR sensitivity.

Skotomorphogenic Phenotypes of *OsDof12* Overexpressing Plants

The phenotypes of mesocotyl and internode elongation in darkness have been employed as a good criteria to determine whether the dwarf phenotype is related to BR (Hong et al., 2002, 2003). To test whether overexpression of *OsDof12* affects the elongation of mesocotyl and internode, we grew the seeds on half-strength MS under totally dark conditions. Two weeks later, we observed the phenotypes of WT and OD plants and found

that the elongation of mesocotyl and first internode in OD plants was similar to those in WT plants (**Figures 6A–C**), however, the second internode length in OD plants was much less than that in WT plants (**Figures 6A,D**), indicating overexpression of *OsDof12* specifically inhibits the elongation of the internode in dark conditions.

Expression Patterns of BR Signaling Related Genes Were Changed in Transgenic Plants Overexpressing *OsDof12*

As demonstrated above, *OsDof12* may be a regulator for maintaining normal BR signaling. We then test whether *OsDof12* responds to BR treatment. As shown in **Figure 7A**, the *OsDof12* transcripts rapidly accumulated after BL treatment for 1 h; although *OsDof12* gradually descended from 2 to 4 h, it was still higher than that before BR treatment; after BR treatment the transcriptional level peak of *OsDof12* appeared at 8 h. This result suggests *OsDof12* could be induced by exogenous BR, which indicates *OsDof12* might be involved in BR regulation.

Using qRT-PCR, we further analyzed the expression level of six major BR signaling components. As shown in **Figure 7B**, *OsBRI1* and *OsBZR1* were down-regulated, while *OsBU1*, *OsLIC*,

FIGURE 5 | Altered BR sensitivity in overexpressing transgenic plants. (A–E) Statistical analysis of the 2nd leaf sheath length **(A)**, seedling height **(B)**, 2nd leaf length **(C)**, coleoptile length **(D)**, and root length **(E)** in plants treated by various levels of BL. Values are mean ± sd, $n = 15$. Single asterisk and double asterisks respectively stand for $P < 0.05$ and $P < 0.01$ determined by student's *t*-test.

and *DLT* were not affected, suggesting that *OsDof12* is possibly involved in BR signaling pathway.

Discussion

Dof genes have been known involved in various processes at different development stages in plants (Yanagisawa, 2002; Noguero et al., 2013). In previous study, we focused on overexpression of *OsDof12* promoting flowering in rice under long-day conditions. In the present study, we demonstrate that overexpression of *OsDof12* can alter the rice plant architecture. Further evidences indicate *OsDof12* may participate in BR signaling to modulate rice architecture. All these clearly show that *OsDof12* is a pleiotropic regulator for plant growth and development.

Higher plants have developed a complex of metabolic mechanisms, which include biosynthetic and catabolic pathway, to maintain BR homeostasis for normal growth and development (Tanaka et al., 2005; Vriet et al., 2013). The abnormal phenotypes in OD plants are similar with the phenotypes of BR-deficient mutants. Therefore, we investigated this possible clues that

OsDof12 might be involved in BR metabolism. Measurement of the endogenous BR levels indicated that the BR content in plants overexpressing *OsDof12* was not affected. Meanwhile, we quantified expression levels of BR metabolic genes, such as *D2*, *D11*, *OsDWARF*, *OsDWARF1*, *OsDWARF4*, *CYP734A*s and no obvious difference were found between OD and WT plants. These results excluded the possibility that *OsDof12* might modulate plant architecture by affecting BR metabolism pathway.

Notably, a range of BR response tests on *OsDof12* overexpression plants, including lamina joint assay, sheath, root, coleoptile elongation pattern analysis, and skotomorphogenesis analysis all suggest that overexpression of *OsDof12* reduces the BR sensitivity. Besides, the *OsDof12* expression was induced by exogenous BL treatment and the expression patterns of two BR signaling components were suppressed in *OsDof12* overexpression plants, which further suggest that *OsDof12* is involved in coordinating rice BR signaling.

Rice plant architecture is composed of tiller number, internode elongation, leaf angle and panicle structure, and favorable architecture is able to facilitate improving the yield (Wang and Li, 2006, 2008; Yang and Hwa, 2008). Plant height is one of
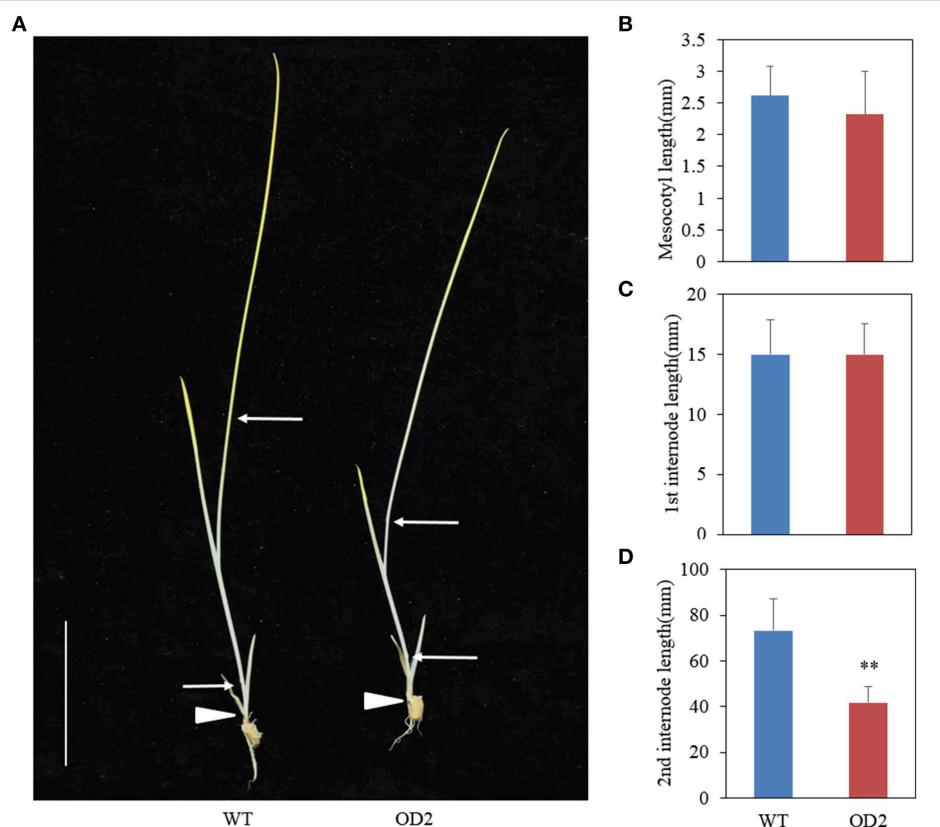
FIGURE 6 | Skotomorphogy analysis. (A) Two-week-old seedlings grown on half-strength MS in dark conditions. The mesocotyl **(B)** and first internode **(C)** length of OD2 is comparable to WT, whereas the second internode **(D)** was obviously shortened in OD2.Values are mean ± sd, n = 12. Double asterisks stand for $P < 0.01$ determined by student's t-test. Arrows and arrowheads indicate the nodes and mesocotyls, respectively. Bar = 25 mm.

the most important agronomic traits (Sakamoto and Matsuoka, 2004; Wang and Li, 2008). Rice dwarf phenotypes have been well categorized into several patterns (Yamamuro et al., 2000), and BR-associated mutants usually display dm-type dwarfism where the second internode is specifically shortened. The *OsDof12* overexpression lines, OD2 and OD5, both generated shortened internodes, among which the relative length of the second internode were shortened significantly, therefore, we speculate that they belong to dm-type dwarfism. As suggested by our microscopic observation in the elongation zone, being different from the two BR mutants, *d61-2* (Yamamuro et al., 2000) and *dlt* (Tong et al., 2009) where the elongation of longitudinal cell is severely hampered, the cell length in OD plants is not affected, which might indicate that the cell number decrease should be the major cause for the dwarfism in OD. These results suggest that the regulatory mechanism underlying cell elongation and plant height appear to be quite complicated, thus the detail mechanism in *OsDOf12* regulation of plant height should be further studied.

Erect leaf pattern is another desirable trait for ideal plant architecture. Generally, the erect leaves in dense plantings allow more light penetrating through the upper leaves layer to make the lower leaves layer capture more light for photosynthesis

and assimilation (Sinclair and Sheehy, 1999). Indeed, several cases have witnessed the effect of erect leaf on improving yield (Morinaka et al., 2006; Sakamoto et al., 2006). For example, the weakest allele of *OsBRI1* (*d61-7*), the counterpart of *Arabidopsis* BR receptor *BRI1*, produced erect leaf and generate greater biomass under high density planting conditions (Morinaka et al., 2006). The erect leaf trait of *OsDof12* overexpression lines is favorable for ideal plant architecture. However, similar to *d61-7* (Morinaka et al., 2006), at least under normal planting density, *OsDof12* overexpression lines would produce decreased spikelets. Therefore, whether the harvest index of *OsDof12* overexpression lines increase or not under different higher planting densities need to be further investigated.

Overexpression of *OsDof12* leads to shorter leaves and less grain yield under normal planting dense, which is unfavorable for utilization in plant breeding. On the contrary, whether the plants down-expressing *OsDof12* induce longer leaves and higher grain yield should be further validated. Indeed, we have tried in this direction but failed to obtain *OsDof12* suppressed lines with obviously distinguishable phenotypes from WT (Li et al., 2009b). We speculated this might result from functional redundancy in rice *Dof* genes, and the functional redundancy in *Dofs* have been demonstrated in Arabidopsis
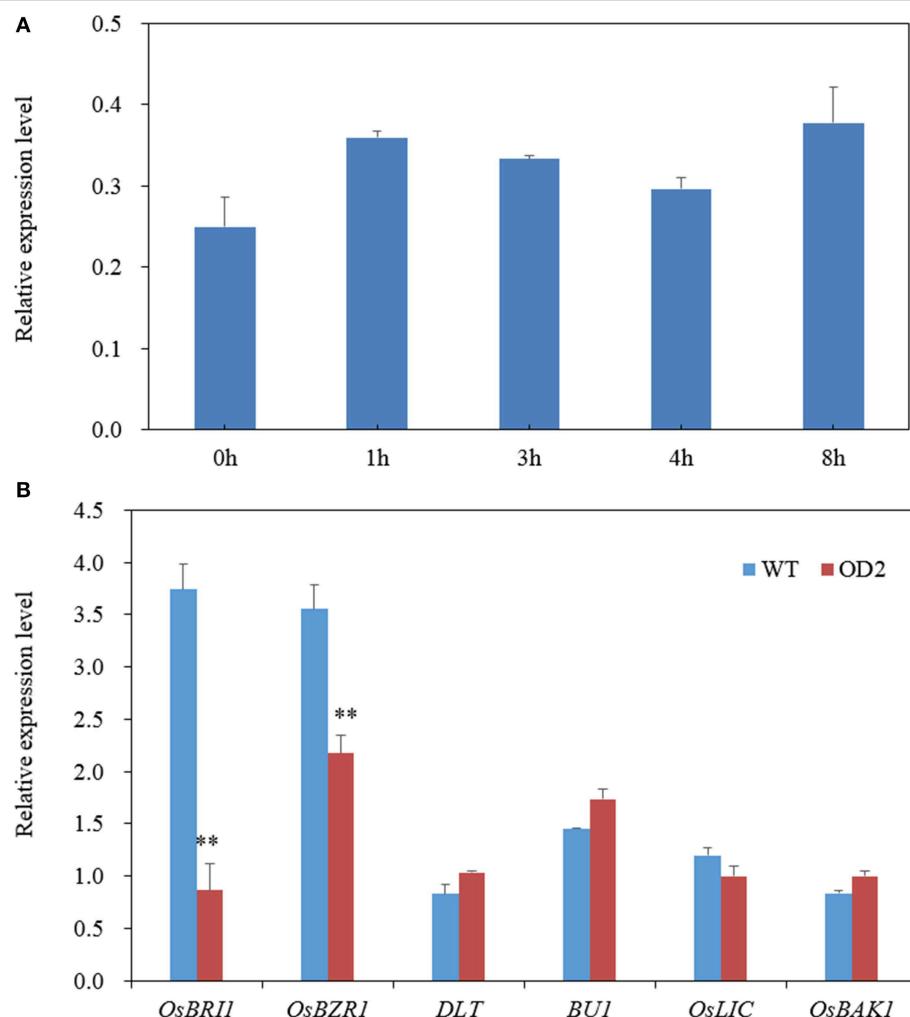
FIGURE 7 | Gene expression analysis. (A) *OsDof12* induction analysis. Two-week-old WT seedlings were subjected to exogenous 100 nM BL treatment, then whole plants were collected at different time point for gene expression analysis. **(B)** qRT-PCR analysis of BR signaling genes. Seven-day-old seedlings grown on half-strength MS were used for expression analysis. Values are mean ± sd, $n = 3$. Double asterisks stand for $P < 0.01$ determined by student's *t*-test.

(Ahmad et al., 2013). Therefore, in our future study, we will try to better understand the full-version function of *OsDof12* by utilizing CRISPR/Cas9 (Belhaj et al., 2015) and Chimeric Repressor gene Silencing Technology (CRES-T) (Mitsuda et al., 2011).

In summary, we show the involvement of *OsDof12* in BR signaling coordination in rice and the effects of overexpression of *OsDof12* on rice plant architecture, and the findings imply *OsDof12* might be a potential genetic module for future rice breeding strategy.

Author Contributions

SL and LZ conceived and designed the experiments. QW and DaL performed most of the experiments. QW, DaL, SL, and LZ wrote the manuscript. DeL and XZ performed phenotypes observation and statistic analysis. XuL and XiL performed microscopic observation. All authors have read and approved the manuscript.

Acknowledgments

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.00833

Figure S1 | Comparison of the grain size between wild type and OD lines. **(A)** There was no obvious difference between WT and *OsDof12* overexpression lines in grain size. Bar = 1.5 mm. **(B)** Statistical data of 1000-grain weight. Values are mean ± sd, *n* = 1000.

Figure S2 | Morphology and statistical analysis of the leaf length. (A) Gross morphology of flag leaves, penultimate leaves and antepenultimate leaves. Bar = 8 cm. **(B)**The length of flag leaves, penultimate leaves and antepenultimate leaves in *OsDof12* overexpression plants was shorter than WT plants. Values are mean ± sd (*n* = 30). Single asterisk and double asterisks stand for *P* < 0.05 and *P* < 0.01 determined by student's *t*-test, respectively.

Figure S3 | Expression analysis of BR metabolism related genes in WT plants and transgenic plants overexpressing *OsDof12*. Expression level of

OsDof12 in OD plants was significantly higher than in WT plants. *D11* and *CYP734A6* was slightly but not obviously down-regulated in OD plants, while expression levels of the other BR metabolism related genes were comparable in both plants. n.d means not detected. Double asterisks stand for *P* < 0.01 determined by student's *t*-test.

Table S1 | Quantification of CS in WT and OD2. Means ± SD of two replicates are shown (ng·g-1 F.W.).

Table S2 | The primers used for qRT-PCR in this study.

References

Ahmad, M., Rim, Y., Chen, H., and Kim, J. Y. (2013). Functional characterization of Arabidopsis Dof transcription factor AtDof4.1. *Russ. J. Plant Physiol.* 60, 116–123. doi: 10.1134/S1021443712060027

Bai, M. Y., Zhang, L. Y., Gampala, S. S., Zhu, S. W., Song, W. Y., Chong, K., et al. (2007). Functions of OsBZR1 and 14-3-3 proteins in brassinosteroid signaling in rice. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13839–13844. doi: 10.1073/pnas.0706386104

Baumann, K., De Paolis, A., Costantino, P., and Gualberti, G. (1999). The DNA binding site of the Dof protein NtBBF1 is essential for tissue-specific and auxin-regulated expression of the rolB oncogene in plants. *Plant Cell* 11, 323–333. doi: 10.1105/tpc.11.3.323

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N. J., and Nekrasov, V. (2015). Editing plant genomes with CRISPR/Cas9. *Curr. Opin. Biotechnol.* 32, 76–84. doi: 10.1016/j.copbio.2014.11.007

Ding, J., Mao, L. J., Wang, S. T., Yuan, B. F., and Feng, Y. Q. (2013). Determination of endogenous brassinosteroids in plant tissues using solid-phase extraction with double layered cartridge followed by high-performance liquid chromatography-tandem mass spectrometry. *Phytochem. Anal.* 24, 386–394. doi: 10.1002/pca.2421

Fornara, F., Panigrahi, K. C. S., Gissot, L., Sauerbrunn, N., Rühl, M., Jarillo, J. A., et al. (2009). Arabidopsis DOF transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. *Dev. Cell* 17, 75–86. doi: 10.1016/j.devcel.2009.06.015

Fujioka, S., and Yokota, T. (2003). Biosynthesis and metabolism of brassinosteroids. *Annu. Rev. Plant Biol.* 54, 137–164. doi: 10.1146/annurev.arplant.54.031902.134921

Gualberti, G., Papi, M., Bellucci, L., Ricci, L., Bouchez, D., Camilleri, C., et al. (2002). Mutations in the Dof zinc finger genes DAG2 and DAG1 influence with opposite effects on the germination of Arabidopsis seeds. *Plant Cell* 14, 1253–1263. doi: 10.1105/tpc.010491

Hong, Z., Ueguchi-Tanaka, M., Fujioka, S., Takatsuto, S., Yoshida, S., Hasegawa, Y., et al. (2005). The Rice brassinosteroid-deficient dwarf2 mutant, defective in the rice homolog of Arabidopsis DIMINUTO/DWARF1, is rescued by the endogenously accumulated alternative bioactive brassinosteroid, dolichosterone. *Plant Cell* 17, 2243–2254. doi: 10.1105/tpc.105.030973

Hong, Z., Ueguchi-Tanaka, M., Shimizu-Sato, S., Inukai, Y., Fujioka, S., Shimada, Y., et al. (2002). Loss-of-function of a rice brassinosteroid biosynthetic enzyme, C-6 oxidase, prevents the organized arrangement and polar elongation of cells in the leaves and stem. *Plant J.* 32, 495–508. doi: 10.1046/j.1365-313X.2002.01438.x

Hong, Z., Ueguchi-Tanaka, M., Umemura, K., Uozu, S., Fujioka, S., Takatsuto, S., et al. (2003). A rice brassinosteroid-deficient mutant, ebisu dwarf (d2), is caused by a loss of function of a new member of cytochrome P450. *Plant Cell* 15, 2900–2910. doi: 10.1105/tpc.014712

Imaizumi, T., Schultz, T. F., Harmon, F. G., Ho, L. A., and Kay, S. A. (2005). FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science* 309, 293–297. doi: 10.1126/science.1110586

Iwamoto, M., Higo, K., and Takano, M. (2009). Circadian clock- and phytochrome-regulated Dof-like gene, Rdd1, is associated with grain size in rice. *Plant Cell and Environ.* 32, 592–603. doi: 10.1111/j.1365-3040.2009.01954.x

Kawakatsu, T., Yamamoto, M. P., Touno, S. M., Yasuda, H., and Takaiwa, F. (2009). Compensation and interaction between RISBZ1 and RPBF during grain filling in rice. *Plant J.* 59, 908–920. doi: 10.1111/j.1365-313X.2009.03925.x

Kim, B. K., Fujioka, S., Takatsuto, S., Tsujimoto, M., and Choe, S. (2008). Castasterone is a likely end product of brassinosteroid biosynthetic pathway in rice. *Biochem. Biophys. Res. Commun.* 374, 614–619. doi: 10.1016/j.bbrc.2008.07.073

Kim, H.-S., Kim, S. J., Abbasi, N., Bressan, R. A., Yun, D.-J., Yoo, S.-D., et al. (2010). The DOF transcription factor Dof5.1 influences leaf axial patterning by promoting Revoluta transcription in Arabidopsis. *Plant J.* 64, 524–535. doi: 10.1111/j.1365-313X.2010.04346.x

Li, D., Wang, L., Wang, M., Xu, Y. Y., Luo, W., Liu, Y. J., et al. (2009a). Engineering OsBAK1 gene as a molecular tool to improve rice architecture for high yield. *Plant Biotechnol. J.* 7, 791–806. doi: 10.1111/j.1467-7652.2009.00444.x

Li, D., Yang, C., Li, X., Gan, Q., Zhao, X., and Zhu, L. (2009b). Functional characterization of rice OsDof12. *Planta* 229, 1159–1169. doi: 10.1007/s00425-009-0893-7

Li, D., Yang, C., Li, X., Ji, B., and Zhu, L. (2008). Sense and antisense OsDof12 transcripts in rice. *BMC Mol. Biol.* 9:80. doi: 10.1186/1471-2199-9-80

Mitsuda, N., Matsui, K., Ikeda, M., Nakata, M., Oshima, Y., Nagatoshi, Y., et al. (2011). CRES-T, an effective gene silencing system utilizing chimeric repressors. *Methods Mol. Biol.* 754, 87–105. doi: 10.1007/978-1-61779-154-3_5

Moreno-Risueno, M., Martínez, M., Vicente-Carbajosa, J., and Carbonero, P. (2007). The family of DOF transcription factors: from green unicellular algae to vascular plants. *Mol. Genet. Genomics* 277, 379–390. doi: 10.1007/s00438-006-0186-9

Morinaka, Y., Sakamoto, T., Inukai, Y., Agetsuma, M., Kitano, H., Ashikari, M., et al. (2006). Morphological alteration caused by brassinosteroid insensitivity increases the biomass and grain production of rice. *Plant Physiol.* 141, 924–931. doi: 10.1104/pp.106.077081

Müssig, C., and Altmann, T. (2001). Brassinosteroid signaling in plants. *Trends Endocrinol. Metab.* 12, 398–402. doi: 10.1016/S1043-2760(01)00477-5

Noguero, M., Atif, R. M., Ochatt, S., and Thompson, R. D. (2013). The role of the DNA-binding One Zinc Finger (DOF) transcription factor family in plants. *Plant Sci.* 209, 32–45. doi: 10.1016/j.plantsci.2013.03.016

Papi, M., Sabatini, S., Bouchez, D., Camilleri, C., Costantino, P., and Vittorioso, P. (2000). Identification and disruption of an Arabidopsis zinc finger gene controlling seed germination. *Genes Dev.* 14, 28–33. doi: 10.1101/gad.14.1.28

Sakamoto, T., Kawabe, A., Tokida-Segawa, A., Shimizu, B., Takatsuto, S., Shimada, Y., et al. (2011). Rice CYP734As function as multisubstrate and multifunctional enzymes in brassinosteroid catabolism. *Plant J.* 67, 1–12. doi: 10.1111/j.1365-313X.2011.04567.x

Sakamoto, T., and Matsuoka, M. (2004). Generating high-yielding varieties by genetic manipulation of plant architecture. *Curr. Opin. Biotechnol.* 15, 144–147. doi: 10.1016/j.copbio.2004.02.003

Sakamoto, T., Morinaka, Y., Ohnishi, T., Sunohara, H., Fujioka, S., Ueguchi-Tanaka, M., et al. (2006). Erect leaves caused by brassinosteroid deficiency increase biomass production and grain yield in rice. *Nat. Biotechnol.* 24, 105–109. doi: 10.1038/nbt1173

Salas Fernandez, M. G., Becraft, P. W., Yin, Y., and Lubberstedt, T. (2009). From dwarves to giants? Plant height manipulation for biomass yield. *Trends Plant Sci.* 14, 454–461. doi: 10.1016/j.tplants.2009.06.005

Santos, L. A., De Souza, S. R., and Fernandes, M. S. (2012). OsDof25 expression alters carbon and nitrogen metabolism in Arabidopsis under high N-supply. *Plant Biotechnol. Rep.* 6, 327–337. doi: 10.1007/s11816-012-0227-2

Sawa, M., Nusinow, D. A., Kay, S. A., and Imaizumi, T. (2007). FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. *Science* 318, 261–265. doi: 10.1126/science.1146994

Sinclair, T. R., and Sheehy, J. E. (1999). Erect leaves and photosynthesis in rice. *Science* 283, 1456–1457. doi: 10.1126/science.283.5407.1455c

Skirycz, A., Jozefczuk, S., Stobiecki, M., Muth, D., Zanor, M. I., Witt, I., et al. (2007). Transcription factor AtDOF4;2 affects phenylpropanoid metabolism in *Arabidopsis thaliana*. *New Phytol.* 175, 425–438. doi: 10.1111/j.1469-8137.2007.02129.x

Skirycz, A., Radziejwoski, A., Busch, W., Hannah, M. A., Czeszejko, J., Kwasniewski, M., et al. (2008). The DOF transcription factor OBP1 is involved in cell cycle regulation in *Arabidopsis thaliana*. *Plant J.* 56, 779–792. doi: 10.1111/j.1365-313X.2008.03641.x

Skirycz, A., Reichelt, M., Burow, M., Birkemeyer, C., Rolcik, J., Kopka, J., et al. (2006). DOF transcription factor AtDof1.1 (OBP2) is part of a regulatory network controlling glucosinolate biosynthesis in Arabidopsis. *Plant J.* 47, 10–24. doi: 10.1111/j.1365-313X.2006.02767.x

Song, Y. H., Smith, R. W., To, B. J., Millar, A. J., and Imaizumi, T. (2012). FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science* 336, 1045–1049. doi: 10.1126/science.1219644

Tanaka, A., Nakagawa, H., Tomita, C., Shimatani, Z., Ohtake, M., Nomura, T., et al. (2009). BRASSINOSTEROID UPREGULATED1, encoding a helix-loop-helix protein, is a novel gene involved in brassinosteroid signaling and controls bending of the lamina joint in rice. *Plant Physiol.* 151, 669–680. doi: 10.1104/pp.109.140806

Tanaka, K., Asami, T., Yoshida, S., Nakamura, Y., Matsuo, T., and Okamoto, S. (2005). Brassinosteroid homeostasis in Arabidopsis is ensured by feedback expressions of multiple genes involved in its metabolism. *Plant Physiol.* 138, 1117–1125. doi: 10.1104/pp.104.058040

Thornton, L. E., Peng, H., and Neff, M. M. (2011). Rice CYP734A cytochrome P450s inactivate brassinosteroids in Arabidopsis. *Planta* 234, 1151–1162. doi: 10.1007/s00425-011-1464-2

Tong, H., Jin, Y., Liu, W., Li, F., Fang, J., Yin, Y., et al. (2009). DWARF AND LOW-TILLERING, a new member of the GRAS family, plays positive roles in brassinosteroid signaling in rice. *Plant J.* 58, 803–816. doi: 10.1111/j.1365-313X.2009.03825.x

Tong, H., Liu, L., Jin, Y., Du, L., Yin, Y., Qian, Q., et al. (2012). DWARF AND LOW-TILLERING acts as a direct downstream target of a GSK3/SHAGGY-like kinase to mediate brassinosteroid responses in rice. *Plant Cell* 24, 2562–2577. doi: 10.1105/tpc.112.097394

Tong, H., and Chu, C. (2012). Brassinosteroid signaling and application in rice. *J. Genet. Genomics* 39, 3–9. doi: 10.1016/j.jgg.2011.12.001

Umemura, Y., Ishiduka, T., Yamamoto, R., and Esaka, M. (2004). The Dof domain, a zinc finger DNA-binding domain conserved only in higher plants, truly functions as a Cys2/Cys2 Zn finger domain. *Plant J.* 37, 741–749. doi: 10.1111/j.1365-313X.2003.01997.x

Vriet, C., Russinova, E., and Reuzeau, C. (2013). From squalene to brassinolide: the steroid metabolic and signaling pathways across the plant kingdom. *Mol. Plant* 6, 1738–1757. doi: 10.1093/mp/sst096

Wang, Y., and Li, J. (2008). Molecular basis of plant architecture. *Annu. Rev. Plant Biol.* 59, 253–279. doi: 10.1146/annurev.arplant.59.032607.092902

Wang, Y., and Li, J. (2006). Genes controlling plant architecture. *Curr. Opin. Biotechnol.* 17, 123–129. doi: 10.1016/j.copbio.2006.02.004

Wang, Z. Y., Bai, M. Y., Oh, E., and Zhu, J. Y. (2012). Brassinosteroid signaling network and regulation of photomorphogenesis. *Annu. Rev. Genet.* 46, 701–724. doi: 10.1146/annurev-genet-102209-163450

Washio, K. (2001). Identification of Dof proteins with implication in the gibberellin-regulated expression of a peptidase gene following the germination of rice grains. *Biochim. Biophys. Acta* 1520, 54–62. doi: 10.1016/S0167-4781(01)00251-2

Washio, K. (2003). Functional dissections between GAMYB and dof transcription factors suggest a role for protein-protein associations in the gibberellin-mediated expression of the RAmy1A gene in the rice aleurone. *Plant Physiol.* 133, 850–863. doi: 10.1104/pp.103.027334

Yamamoto, M. P., Onodera, Y., Touno, S. M., and Takaiwa, F. (2006). Synergism between RPBF Dof and RISBZ1 bZIP activators in the regulation of rice seed expression genes. *Plant Physiol.* 141, 1694–1707. doi: 10.1104/pp.106.082826

Yamamuro, C., Ihara, Y., Wu, X., Noguchi, T., Fujioka, S., Takatsuto, S., et al. (2000). Loss of function of a rice brassinosteroid insensitive1 homolog prevents internode elongation and bending of the lamina joint. *Plant Cell* 12, 1591–1605. doi: 10.1105/tpc.12.9.1591

Yanagisawa, S. (2000). Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. *Plant J.* 21, 281–288. doi: 10.1046/j.1365-313x.2000.00685.x

Yanagisawa, S. (2002). The Dof family of plant transcription factors. *Trends Plant Sci.* 7, 555–560. doi: 10.1016/S1360-1385(02)02362-2

Yanagisawa, S., and Schmidt, R. J. (1999). Diversity and similarity among recognition sequences of Dof transcription factors. *Plant J.* 17, 209–214. doi: 10.1046/j.1365-313X.1999.00363.x

Yang, X. C., and Hwa, C. M. (2008). Genetic modification of plant architecture and variety improvement in rice. *Heredity* 101, 396–404. doi: 10.1038/hdy.2008.90

Yokota, T. (1997). The structure, biosynthesis and function of brassinosteroids. *Trends Plant Sci.* 2, 137–143. doi: 10.1016/S1360-1385(97)01017-0

Zhang, C., Xu, Y., Guo, S., Zhu, J., Huan, Q., Liu, H., et al. (2012). Dynamics of brassinosteroid response modulated by negative regulator LIC in rice. *PLoS Genet.* 8:e1002686. doi: 10.1371/journal.pgen.1002686

Zhu, J. Y., Sae-Seaw, J., and Wang, Z. Y. (2013). Brassinosteroid signalling. *Development* 140, 1615–1620. doi: 10.1242/dev.060590

Zou, H. F., Zhang, Y. Q., Wei, W., Chen, H. W., Song, Q. X., Liu, Y. F., et al. (2013). The transcription factor AtDOF4.2 regulates shoot branching and seed coat formation in Arabidopsis. *Biochem. J.* 449, 373–388. doi: 10.1042/BJ20110060

Identification of Quantitative Trait Loci and Water Environmental Interactions for Developmental Behaviors of Leaf Greenness in Wheat

Delong Yang[1]*, Mengfei Li[1], Yuan Liu[1], Lei Chang[2], Hongbo Cheng[1], Jingjing Chen[1] and Shouxi Chai[2]

[1] Gansu Provincial Key Lab of Aridland Crop Science/School of Life Science and Technology, Gansu Agricultural University, Lanzhou, China, [2] School of Agronomy, Gansu Agricultural University, Lanzhou, China

The maintenance of leaf greenness in wheat, highly responsible for yield potential and resistance to drought stress, has been proved to be quantitatively inherited and susceptible to interact with environments by traditional genetic analysis. In order to further dissect the developmental genetic behaviors of flag leaf greenness under terminal drought, unconditional and conditional QTL mapping strategies were performed with a mixed linear model in 120 F_8-derived recombinant inbred lines (RILs) from two Chinese common wheat cultivars (Longjian 19 × Q9086) in different water environments. A total of 65 additive QTLs (A-QTLs) and 42 pairs of epistatic QTLs (AA-QTLs) were identified as distribution on almost all 21 chromosomes except 5A, explaining from 0.24 to 3.29 % of the phenotypic variation. Of these, 22 A-QTLs and 25 pairs of AA-QTLs were common in two sets of mapping methods but the others differed. These putative QTLs were essentially characteristic of time- and environmentally-dependent expression patterns. Indeed some loci were expressed at two or more stages, while no single QTL was continually active through whole measuring duration. More loci were detected in early growth periods but most of QTL × water environment interactions (QEIs) happened in mid-anaphase, where drought stress was more conducted with negative regulation on QTL expressions. Compared to other genetic components, epistatic effects and additive QEIs effects could be predominant in regulating phenotypic variations during the ontogeny of leaf greenness. Several QTL cluster regions were suggestive of tight linkage or expression pleiotropy in the inheritance of these traits. Some reproducibly-expressed QTLs or common loci consistent with previously detected would be useful to the genetic improvement of staygreen types in wheat through MAS, especially in water-deficit environments.

Keywords: Triticum aestivum, leaf greenness, drought stress, developmental genetics, QTL mapping, water environmental interactions

INTRODUCTION

Wheat (*Triticum aestivum* L.) is one of the most important foodstuff crops in semiarid and arid areas around the world. As current changes in global climate have increased the precipitation variability with frequent episodes of drought (Trenberth, 2011), the wheat production in rainfed regions is strongly constrained by erratic drought stresses (Gregersen et al., 2013). In particular, terminal drought occurring during the reproductive phase in wheat is responsible for poor grain set and development, which finally results in substantial reductions in grain yield (Nawaz et al., 2013). Although, principal explanations for these losses are still complicated, it is critically associated with the drought-induced premature senescence of flag leaf (Verma et al., 2004). Here terminal drought is considered to essentially accelerate leaf chlorophyll degradation and thus impede carbon fixation (Guóth et al., 2009) and assimilate remobilization (Gregersen et al., 2008). In this context, wheat genotypes with a functional staygreen characteristic, i.e., delayed leaf senescence, can maintain photosynthetic capacity and favorable supply of assimilates to grain for a longer duration of time to assure better grain yield (Gong et al., 2005; Christopher et al., 2008; Chen et al., 2010). Therefore, the staygreen attribute of flag leaf under terminal drought is of great importance for determining wheat yield potential and resistance to drought stress (Biswal and Kohli, 2013; Farooq et al., 2014).

To develop the staygreen trait of flag leaf as an effective selection criterion in drought-tolerant breeding in wheat, much effort has been exerted to understand the genetic mechanism of the trait in wheat (Kumar et al., 2010; Naruoka et al., 2012; Barakat et al., 2013) and other cereal crops (Yoo et al., 2007; Kassahun et al., 2010; Wang et al., 2012; Emebiri, 2013). Current genetic gains in this phenotype, as reviewed by Thomas and Ougham (2014), are proved to be polygenic with functional genes and transcription factors by typical approaches of mutation, gene expression profiles and transgenic plants. Alternatively, polygenes with quantitative effects can also be developed by quantitative trait loci (QTLs) analysis (Verma et al., 2004; Yoo et al., 2007; Kassahun et al., 2010; Kumar et al., 2010; Wang et al., 2012; Emebiri, 2013). In wheat, these putative QTLs were almost mapped on all 21 chromosomes, with a widely flexible expression in response to various genetic populations and environments (Verma et al., 2004; Yang et al., 2007; Zhang et al., 2009a,b, 2010; Kumar et al., 2010; Li H. et al., 2012; Naruoka et al., 2012; Barakat et al., 2013; Czyczyło-Mysza et al., 2013). The genetic components estimated from segregation progenies of wheat crosses elucidated that the leaf staygreen trait was governed by only a few of major genes /QTLs with high predominance of additive effects (Silva et al., 2000; Verma et al., 2004; Joshi et al., 2007; Kumar et al., 2010). However, most of present studies indicated that the phenotype was under polygenes control by minor additive effects, which were variable across environments (Li H. et al., 2012; Naruoka et al., 2012; Barakat et al., 2013; Czyczyło-Mysza et al., 2013). In some cases, epistatic effects (Zhang et al., 2009a,b, 2010; Kumar et al., 2012) or QTL × environment interaction (QEI) effects (Yang et al., 2007; Peleg et al., 2009) were also highlighted in the modulation of

its genetic variation. Under these observations, although flag leaf staygreen in wheat was confirmed to be inherited quantitatively, few studies have been undertaken to adequately dissect its genetic components and QEI variability in a same experiment system, especially under the terminal drought.

On the other hand, leaf staygreen *per se* is a complex developmental process (Thomas and Ougham, 2014). The statistical analysis showed that the development of such a quantitative trait occurs through the actions and interactions of polygenes and their environmental interactions, behaving differentially during different growth stages (Atchley and Zhu, 1997). Nevertheless, present QTL information for leaf staygreen in wheat was only achieved at a specific time point without considering sequential effects due to distinct gene expression at different developmental stages. Actually, it is inadequate to discover the real genetic information on the developmental processes of target quantitative traits (Wang et al., 2010). To dissect the dynamics of QTL expression, unconditional and conditional methods have been proposed (Zhu, 1995; Atchley and Zhu, 1997; Wu et al., 1999). Unconditional analysis is a traditional method for studying developmental behavior, which unravels genetic cumulative effects from the original to time *t*, rather than the real effects of gene expression during the ontogeny. Conditional analysis is another method to assess net genetic effects in the period from time (*t*-1) to time *t* on trait development. Being independent of the causal genetic effects and susceptible to the developmental status and environments, conditional analysis is valid to identify dynamic gene expression and new genetic variation arising in specific development periods (Cao et al., 2001; Wu et al., 2010). Present studies have also been verified that conditional effects in early growth periods, as normally cumulative components, could affect later unconditional effects (Wang et al., 2010; Wu et al., 2010; Li S. et al., 2012).Thus, the combination of unconditional and conditional analyses is employed to identify dynamic expressions of developmental QTLs and reveal the comprehensive inheritance of quantitative traits (Wu et al., 2010; Li S. et al., 2012). The strategy has been applied to understanding the genetic basis of crop developmental traits, such as plant height (Wang et al., 2010; Wu et al., 2010), tiller number (Yang et al., 2006; Liu et al., 2010), grain weight (Han et al., 2012; Li S. et al., 2012), grain filling (Takai et al., 2005), and seed quality (Han et al., 2011). All the studies show that the genetic architecture of developmental traits gets more involved in a time-dependent expression of polygenes through additive, epistatic and QEI effects. However, few reports so far on QTLs analysis by this methodology have documented the staygreen of flag leaf in wheat under the terminal drought stress across diverse water environments.

In this study, we used recombinant inbred lines (RILs) in wheat grown under different water environments to explore genetically the developmental behaviors of flag leaf staygreen. Using both the unconditional and conditional genetic models analyzed the trait performances at multiple observation times in the reproductive duration, respectively. The objectives of the work reported here were (1) to identify the QTLs with genetic main effects and QEI effects controlling flag leaf staygreen and (2)

to unravel its dynamics of QTL expression during ontogeny, even under terminal drought stress. The findings might be valuable for well-understanding the genetic basis governing leaf staygreen development, and for wheat genetic improvement of drought tolerance by marker-assisted selection (MAS).

MATERIALS AND METHODS

Plant Materials

A subset of 120 F_8-derived RILs was developed from the cross between two Chinese winter wheat varieties, Longjian 19 and Q9086. The male parent Longjian 19 is an elite drought-tolerance cultivar widely grown in rainfed areas (300~500-mm annual rainfall) in northwestern China and was released by Gansu Academy of Agricultural Sciences, Lanzhou, China. The female Q9086 is a high-yielding cultivar alternative to relatively sufficient water and fertile conditions, but easily susceptible to pre-senescence under terminal drought stress, and was released by Northwest Agriculture & Forestry University, Yangling, China. In addition, the two parents differed significantly in major agronomical and physiological characteristics under terminal drought stress, as described in our previous studies (Yang et al., 2012, 2014; Li et al., 2013; Ma et al., 2014; Hu et al., 2015; Ye et al., 2015).

Field Trials

Two parents together with the RILs were grown at Lanzhou (103°51′ E, 36°04′ N, 1520 m Altitude), Gansu, China, in 2012–2014, denoted in turn as E1, E2, and E3, respectively. The experimental fields in each year were treated under drought stress (DS) and well-watered (WW) conditions. The DS treatment was equivalent to the rainfed condition with a total of 99.6, 113.5, and 101.8 mm rainfall during the growing season (from early October in the sowing year to late June in harvesting year), respectively. The WW treatment was irrigated with 750 m³ ha⁻¹ water supply at the pre-overwintering, jointing, and flowering stages, respectively. The field design of each plot consisted of randomized complete blocks with three replications. Each plot was 2 m long with six rows spaced 20 cm apart with approximately 160 plants per row. Field management was conducted following the local practice in wheat production.

Leaf greenness, a surrogate measure of leaf chlorophyll content, was monitored using a Minolta Chlorophyll Meter SPAD 502 (Konica Minolta, Japan), which has been used extensively in the accurate diagnosis of the staygreen characteristics in many crops (Borrell et al., 2001; Jiang et al., 2004; Harris et al., 2007). In this study, SPAD readings were thus made a direct assessment of the degree of leaf greenness. The flag leaves of 10 main shoots growing uniform of each RIL were tagged from the middle of each plot for the assessment of the dynamic greenness degree. SPAD readings were taken at the central point of target flag leaves at each time of measurement. At the onset of flowering, SPAD values were scored every 6 d until 24 d after flowering, which duration actually covered peak grain-filling stages. Therefore, the foregoing five measurements were designated as S1, S2, S3, S4, and S5, respectively. The trait

means of 10 samples from each plot with three replications were applied to the data analysis.

Data Analysis

According to the development theory proposed by Zhu (1995), the actual SPAD data measuring at the above-mentioned stages (S1–S5) were defined as unconditional values. Using the program of QGA Station based on a mixed model (Yang et al., 2006), the conditional SPAD values were generated by converting the actual SPAD data at every two consecutive occasions, designated as S1|S0, S2|S1, S3|S2, S4|S3, and S5|S4, respectively. When phenotypic values were measuring at the first time point, the unconditional genetic effects were equivalent to those obtained from conditional analysis.

Basic statistics and Pearson phenotypic correlations between the traits were performed by SAS software (SAS Institute, 1996). The broadsense heritability (h_B^2) of the greenness of flag leaf was estimated with the method proposed by Toker (2004). To dissect the quantitative genetic basis of the developmental behavior of the post-anthesis greenness of flag leaf, its unconditional and conditional phenotypic data at five growth stages were subjected to QTL analysis. A genetic linkage map, consisting of 524 simple sequence repeats (SSRs) marker loci mapped on 21 chromosomes, was available. The map was covered 2266.7 cM with an average distance of 4.3 cM between adjacent markers (Hu et al., 2015; Ye et al., 2015). QTL analysis was implemented by the mixed linear model mapping (Wang et al., 1999), using the Windows version computer program QTLNetwork-2.0 (Yang et al., 2008). The genetic model could divide genetic effects into additive effects (A), epistatic effects (AA), and QEI effects (including AE and AAE). An experiment-wise type I error of 0.05 was designated for candidate interval selection and putative QTL detection. The critical F-value to declare putative QTLs and to control genome-wise type I errors was accommodated by 1000 permutation tests. Both the testing and filtration window were set at 10 cM, with a walk speed of 2 cM. QTLs were named according to the rule of 'QTL+trait+research department+chromosome'.

RESULTS

Phenotypic Variation and Trait Correlation

The genotypic means for SPAD values as the greenness of flag leaf in all treatments are summarized in **Table 1**. The phenotypic variations of both RILs and two parents showed a progressive depletion trend with the growth progression across water environments. Moreover, phenotypic values in the DS were significantly lower than those in the WW. The parents, Longjian 19 and Q9086, differed in consecutive traits in response to the water regimes. Under the WW, phenotypic values of Q9086 at most of stages were higher than those of Longjian 19, whereas the case in the DS was reverse and much noticeable. This suggested Longjian 19 was capable of stronger staygreen to withstand drought stress. Across all measuring stages, the mean values of RILs displayed a consistent reduction and were intermediate between those of the two parents. Highly phenotypic variability was found in the population, with coefficients of variation (CV) ranging from 6.31 to 27.00% in the DS and from 4.91 to 15.76% in

TABLE 1 | Phenotypic performance for the greenness (SPAD values) of flag leaf of the parents and RILs at five growth stages in different water environments.

Environ[a].	Stage[b]	Parents		RILs					
		Longjian19	Q9086	Mean	Min.	Max.	CV (%)[c]	Skewness	Kurtosis
E1	S1	46.52/50.05	41.46/54.83	45.10/53.66	34.23/42.10	52.73/54.83	7.21/5.41	0.29/−0.17	−0.22/−0.66
	S2	45.15/50.97	40.03/53.46	42.46/52.72	31.01/41.27	49.31/54.49	6.68/5.70	0.01/−0.15	−0.59/−0.62
	S3	40.50/49.72	31.68/50.14	36.46/49.95	28.20/35.50	43.00/51.17	7.16/4.91	0.07/−0.26	−0.21/0.22
	S4	32.38/44.63	15.71/45.78	23.03/44.13	18.18/29.73	33.62/47.52	14.96/6.68	−0.15/0.17	−0.15/−0.57
	S5	20.27/35.81	12.85/32.15	15.06/33.04	12.78/20.53	22.15/35.33	22.17/15.76	0.60/0.28	0.08/−0.41
E2	S1	50.74/52.16	46.55/55.95	47.32/53.73	36.61/44.42	54.98/55.65	6.31/5.24	0.33/−0.09	−0.21/−0.59
	S2	47.96/50.84	41.28/55.70	43.73/52.85	33.38/43.02	52.76/54.68	6.42/5.87	−0.15/−0.07	−0.24/−0.53
	S3	42.58/49.13	34.64/51.58	38.93/50.68	30.24/39.06	46.80/53.41	6.71/5.45	0.05/−0.34	−0.92/0.10
	S4	38.43/47.50	19.78/48.47	26.22/47.71	21.94/33.57	38.48/50.33	11.93/6.72	0.06/0.08	−0.24/−0.55
	S5	24.90/39.69	14.32/35.86	17.68/36.54	13.52/23.11	25.89/39.18	25.57/14.48	0.61/0.40	0.76/−0.11
E3	S1	46.52/51.43	42.60/54.65	44.57/52.22	33.40/42.94	53.41/55.30	6.99/5.30	0.35/−0.16	−0.10/−0.43
	S2	44.96/51.58	37.38/54.01	40.02/52.41	32.35/40.62	50.08/55.04	7.44/5.91	−0.17/−0.14	−0.18/−0.38
	S3	41.85/49.72	30.77/51.93	37.19/50.76	29.55/37.58	45.09/51.80	7.60/5.54	0.05/−0.41	−0.96/0.43
	S4	33.28/41.63	17.5/42.75	24.48/41.86	21.05/31.64	36.97/48.61	16.90/7.99	0.07/0.07	−0.20/−0.37
	S5	22.43/36.81	12.25/34.15	15.36/33.25	13.15/22.27	24.47/37.53	27.00/13.80	0.55/0.41	0.71/−0.08

The numbers at the left of the slash ("/") are the phenotypic values of traits identified in the drought stress, and the numbers at the right indicate the well-watered condition. [a]E1–E3 represent the location at Lanzhou (103°51′ E, 36°04′ N, 1520 m Altitude), Gansu, China, in 2012–2014, respectively. [b]S1–S5 indicate the first to the fifth measuring stage, respectively. [c]CV(%) means the coefficient of variation.

the WW depending on different stages and water environments. Some lines had more extreme values than the parents, showing substantial transgressive segregation. All skewness and kurtosis values were less than 1.0 occurring at all treatments, suggestive of their continuous distributions and quantitative bases.

Results of variation component analysis (**Table 2**) showed that all the variances for both unconditional and conditional values in RILs reached the 0.05 or 0.01 significant level, except for the interaction variances of both genotype × environment and genotype × environment × water. In contrast, the dominant source of variations for unconditional and conditional traits was the water regime, which accounted for 83.82–93.26% and 83.82–95.42% of total variation, respectively. This indicated that water environments could exert more considerable influence on the ontogeny of flag leaf greenness. Despite the substantial variation, the h_B^2 estimates were reasonable for both unconditional and conditional values across environments, differing from 0.34 to 0.64 and from 0.31 to 0.64, respectively, which were present in a gradual decline trend with developmental stages.

Correlation analysis based on both unconditional and conditional SPAD values in RILs between different growth stages across water environments were given in **Table 3**. All correlations were positive and their correlation coefficients widely varied from 0.13 to 0.98** across environments. Howbeit it appeared obvious that correlation coefficients decreased with the growth course, and even occurred weakly at the last stage. For most stages across water environments, unconditional correlations ($r^2 = 0.21 \sim 0.98^{**}$) were more significant and substantially higher than conditional ones ($r^2 = 0.13 \sim 0.51^{**}$). On the other hand as shown in **Table 4**, most of the correlations

between unconditional and conditional data were poorly positive but rather low, with the exception of significant correlations occurring between the conditional period of S1|S0 and all unconditional stages, and between conditional periods and their corresponding unconditional stages. It was inferred that the greenness ontogeny of flag leaf was highly characteristic of dynamic and environmentally influenced scenarios.

Unconditional QTL Analysis for Leaf Greenness Development

A total of 50 additive QTLs (A-QTLs) with significant *A* effects and/or *AE* effects were detected by unconditional SPAD data at different growth stages across water environments (**Table 5**). These loci were mapped on almost all chromosomes except 1D–4D, 5A, and 6D (**Figure 1**), individually explaining from 0.24 to 1.80% of the phenotypic variation. Half of them had positive *A* effects with 0.28* to 0.94*** conferred by favorable alleles from Q9086, whereas the others with negative *A* effects of -0.26* to -1.10*** were from Longjian 19. By contrast, a majority of putative A-QTLs (34 of 50) were identified at only one specific stage, but the remaining 16 loci were detectable at two to four stages, implying that no QTLs was continually active during the whole period of growth. For example, three of them, *Qspad.acs-1B.1*, *Qspad.acs-2A.3*, and *Qspad.acs-5B.2*, were related to four stages from S1 to S4. Two A-QTLs, *Qspad.acs-1A.2* and *Qspad.acs-6A.2*, were involved in three stages from S1 to S3 and from S2 to S4, respectively. The other 11 were just associated with two stages. Interestingly, these A-QTLs expressed at more than one stage were always conducted in the same sources of *A* effects, whereas effect values from each QTL progressively

TABLE 2 | Analyses of variance (ANOVA) for the greenness of flag leaf in RILs at five growth stages.

Source of variation	df	Mean square				
		S1/S1|S0	S2/S2|S1	S3/S3|S2	S4/S4|S3	S5/S5|S4
Water(W)	1	16612.78**/16612.78**	30119.34**/4272.15**	62388.58**/2735.29**	124672.16**/4943.98**	69656.65**/3916.81**
Environment(E)	2	2920.48**/2920.48**	4249.51**/256.80**	2758.76**/234.53**	14966.08**/322.80**	3945.03**/303.87**
Genotype(G)	119	154.39**/154.39**	159.91**/19.72**	111.63**/21.63**	179.31**/54.81**	101.85**/49.26**
W × E	2	74.38**/74.38**	416.66**/19.04**	57.31**/12.49**	125.62**/67.83**	133.71**/15.13**
W × G	119	49.18**/49.18**	54.73**/7.57*	51.17**/8.31*	87.74**/26.79**	58.28**/27.85**
E × G	238	3.73/3.73	6.90/1.29	4.76/2.28	4.26/4.02	6.01/4.24
W × E × G	238	1.32/1.32	3.92/0.94	1.77/2.04	2.32/2.53	3.27/1.75
Error	1440	2.97/2.97	3.64/3.26	6.22/5.61	6.99/6.86	7.95/5.88
h_B^2		0.64/0.64	0.59/0.50	0.48/0.41	0.47/0.39	0.34/0.31

The numbers at the left of the slash ("/") are mean squares or heritability for the unconditional phenotype of RILs, and the numbers at the right indicate the conditional phenotype of RILs. *P ≤ 0.05 and **P ≤ 0.01. S1–S5 are as shown in **Table 1**. S|S-1 indicates the time interval from S-1 to S.

TABLE 3 | Correlation coefficients of the greenness of flag leaf in RILs between different growth stages in different water environments.

Environment	Stage	S1/S1|S0	S2/S2|S1	S3/S3|S2	S4/S4|S3	S5/S5|S4
E1	S1/S1|S0		0.96**/0.51**	0.84**/0.49**	0.68**/0.43**	0.35*/0.26
	S2/S2|S1	0.98**/0.49**		0.86**/0.45**	0.66**/0.23	0.28/0.21
	S3/S3|S2	0.90**/0.43*	0.92**/0.42**		0.67**/0.23	0.27/0.18
	S4/S4|S3	0.76**/0.37*	0.77**/0.38*	0.80**/0.31*		0.21/0.13
	S5/S5|S4	0.29/0.21	0.27/0.24	0.25/0.25	0.26/0.20	
E2	S1/S1|S0		0.92**/0.36*	0.71**/0.28	0.71**/0.44**	0.28/0.15
	S2/S2|S1	0.95**/0.40*		0.76**/0.24	0.64**/0.47**	0.25/0.24
	S3/S3|S2	0.83**/0.32*	0.86**/0.27		0.60**/0.39*	0.27/0.22
	S4/S4|S3	0.75**/0.33*	0.74**/0.23	0.76**/0.25		0.25/0.22
	S5/S5|S4	0.31/0.23	0.31/0.20	0.33/0.22	0.28/0.16	
E3	S1/S1|S0		0.92**/0.31	0.72**/0.21	0.71**/0.48**	0.27/0.18
	S2/S2|S1	0.94**/0.39*		0.74**/0.23	0.63**/0.41*	0.28/0.22
	S3/S3|S2	0.82**/0.35*	0.84**/0.27		0.61**/0.39*	0.25/0.21
	S4/S4|S3	0.72**/0.34*	0.72**/0.23	0.73**/0.20		0.22/0.15
	S5/S5|S4	0.28/0.20	0.26/0.19	0.26/0.21	0.21/0.17	

The numbers at the left of the slash ("/") are correlation coefficients for the unconditional phenotype of RILs, and the numbers at the right indicate the conditional phenotype of RILs. *P ≤ 0.05 and **P ≤ 0.01. Numbers in the upper right segment apply to the drought stress; those at the lower left are for the well-watered condition. E1–E3 and S1–S5 are as shown in **Table 1**, and S1|S0 to S5|S4 are as shown in **Table 2**.

decreased with advanced stages. The number of A-QTLs per stage differed from 11 at S1 to 24 at the S3, and finally to 5 at S5. This suggested that expressions of A-QTLs governing leaf greenness were highly modulated by the developmental stage and were more activated in early periods. In addition, 27 A-QTLs showed significantly QEI effects, explaining from 0.87 to 5.42% of the phenotypic variations (**Table 5**), implying their genetic susceptibility to environments. Of these, about 90% of additive QEIs (A-QEIs) was highlighted in the period after S2, but few of them reacted at S1. During the period from S1 to S3, nine A-QTLs got involved in negative *AE* effects (−0.25* to −1.03***) with the WW, but in positive effects (0.65*** to 1.12***) with the DS. This indicated that *AE* effects of these loci exposed during the early period could be more up-regulated by the DS. However, the cases for most of remainder loci expressed after S3 were opposite, with

negative *AE* effects (−0.23* to −0.87***) in the DS but positive effects (0.22* to 0.85***) in the WW, suggestive of up-regulating them by the WW in the later duration.

All 38 pairs of epistatic QTLs (AA-QTLs) for the greenness of flag leaf were available with significant *AA* effects, accounting for phenotypic variations of 1.08–3.29% in response to different stages and environments (Table S1). These loci involved in epistasis were distributed on all chromosomes apart from 1B and 5A. Among them, 13 pairs behaved with positive *AA* effects (0.82*** to 1.32***), indicating that parent-type effects were higher than recombinant-type effects. And, 23 pairs showed negative effects (−0.42*** to −1.51∗∗∗) where recombinant-type effects were higher than parent-type effects. The remaining two pairs altered in effect directions responsive to the specific stage, where one pair had positive effects (0.78***) at S4 but

FIGURE 1 | Chromosome locations of additive QTLs for the greenness of flag leaf in wheat RILs. Black squares are common QTLs for both unconditional and conditional chlorophyll content, and green and red squares are specific QTLs for unconditional and conditional chlorophyll content, respectively. QTLs symbolized with the asterisk (*) mean that QTLs can be detected more than two stages.

TABLE 4 | Correlation coefficients between unconditional and conditional greenness of flag leaf in RILs at five growth stages different water environments.

Environment	Stage	S1|S0	S2|S1	S3|S2	S4|S3	S5|S4
E1	S1	1.00**/1.00**	0.31*/0.48**	0.33*/0.39*	0.37*/0.43**	0.25/0.30
	S2	0.96**/0.98**	0.47**/0.61**	0.36*/0.40*	0.38*/0.41*	0.24/0.31
	S3	0.84**/0.90**	0.36**/0.55**	0.51**/0.65**	0.38*/0.45**	0.22/0.30
	S4	0.68**/0.76**	0.34*/0.36**	0.47**/0.50**	0.53**/0.76**	0.24/0.26
	S5	0.36*/0.21	0.26/0.25	0.25/0.28	0.24/0.27	0.84**/0.89**
E2	S1	1.00**/1.00**	0.36*/0.40*	0.28/0.32	0.33*/0.43**	0.29/0.32
	S2	0.92**/0.95**	0.43**/0.50**	0.29/0.29	0.31/0.33	0.25/0.33
	S3	0.71**/0.83**	0.36*/0.37*	0.58**/0.58**	0.23/0.22	0.21/0.27
	S4	0.71**/0.75**	0.37*/0.43**	0.32/0.41*	0.60**/0.75**	0.26/0.30
	S5	0.28/0.23	0.27/0.32	0.21/0.24	0.21/0.26	0.82**/0.86**
E3	S1	1.00**/1.00**	0.31*/0.39*	0.21/0.35*	0.34*/0.48**	0.25/0.18
	S2	0.92**/0.94**	0.29**/0.51**	0.26/0.30	0.30/0.37*	0.23/0.30
	S3	0.72**/0.82**	0.33/0.45**	0.62**/0.63**	0.21/0.21	0.23/0.22
	S4	0.71**/0.72**	0.32/0.42**	0.37/0.51**	0.63**/0.77**	0.28/0.28
	S5	0.27/0.24	0.21/0.29	0.22/0.37	0.21/0.39*	0.85**/0.88**

*The numbers at the left of the slash ("/") are correlation coefficients estimated in the drought stress, and the numbers at the right indicate the well-watered condition. *P ≤ 0.05 and **P ≤ 0.01. E1–E3 and S1–S5 are as shown in **Table 1**, and S1|S0–S5|S4 are as shown in **Table 2**. Underline values mainly highlight in the higher and more significant correlation coefficients.*

negative effects (−0.45***) at S5, and another pair showed positive effects (1.34***) at S1 but negative effects (−0.93*** and −0.75***) at S2 and S3, respectively. Similar to the expression patterns of A−QTLs, putative AA-QTLs were also expressed dynamically and alternatively. For example, the amounts of significant AA-QTLs for each stage subsequently decreased, ranging from 24 at the S1 to 2 at the S5. Moreover, 20 pairs were detected at only one specific stage, while the other pairs were identified more at two to three early consecutive stages (S1 to S2 or S3) except three pairs available after S3. Thus, epistatic reactions to the ontogeny of leaf greenness were short lived without more genetic effects appearing in the later duration, of which AA effects progressively decreased. Apart from AA effects, significant AAE effects were also involved in 24 AA-QTLs, explaining phenotypic variations of 0.82–2.54%. Over 76% of epistatic QEIs (E-QEIs) reacted in the period after S2. Of these, just two pairs involved positive AAE effects (0.66*** and 0.52***, respectively) with the DS. The other 22 pairs entirely specified their QEI with negative effects (−0.33** to −0.95***) in the DS, while positive effects (0.31** to 0.94***) in the WW. This indicated that AAE effects were highly up-regulated by the WW.

Concerning the source of epistatic loci, only 10 significant A-QTLs participated in epistatic interactions and therefore exhibited their pleiotropic functions. Howbeit most of epistatic interactions (nearly 90%) were derived from non-individual QTLs, which were involved in epistatic QTLs without any significant A effects. These loci even constituted QTL-interacting networks at different interaction levels (**Figure 2**) to realize different AA effects. For instance, five A-QTLs interacting with 16 non-individual loci were composed of seven smaller networks by three-locus interactions, respectively. The remaining 13 non-individual QTLs made up three relatively bigger networks

from four or five-locus interactions, respectively. Almost 60% of interactions exhibited negative AA effects with unequal magnitudes at one to three stages. This further suggested that the genetic control of the greenness development of flag leaf was complex and, to a certain extent, reacted as part of QTL networks.

Conditional QTL Analysis for Leaf Greenness Development

Based on conditional mapping, a total of 37 A-QTLs were identified with significant A effects and/or QEI effects for the greenness ontogeny of flag leaf across water environments (**Table 6**). These loci were nearly distributed on the same chromosomes as unconditional ones (**Figure 1**), individually explaining from 0.36 to 1.29% of the phenotypic variation. Of these, nearly half of favorable alleles derived from Q9086 with significantly positive A effects (0.30** to 0.94***), whereas another half came from Longjian 19 with significantly negative effects (−0.31** to −0.88***). The result was exactly consistent with the finding from unconditional mapping, further confirming that favorable alleles were averagely dispersed within the two parents. All of conditional A-QTLs were available just in one specific period. The QTL number detected in each period gradually reduced from 11 at S1 to 2 at S5. This indicated that genes governing leaf greenness were expressed selectively but more in early development period. In addition, 18 loci were noticeably associated with A-QEIs, individually accounting for 0.82–3.80% of the phenotypic variation. Most of them (nearly 85%) happened in the duration from S2|S1 to S5|S4, suggesting that the real gene expression in the period of S1|S0 was less influenced by water environments, whereas it was extremely done thereafter. In this regards, all of A-QTLs interacting with the DS showed significantly negative AE (−0.32** to −0.79***)

TABLE 5 | Unconditional additive and interacting effects of QTL × water environment of identified QTLs for the greenness of flag leaf.

QTL	Flanking markers	Stage[a]	A[b]	AE[c]	$H^2(A)$[b](%)	$H^2(AE)$[c](%)
Qspad.acs-1A.1	Xgwm633-Xgwm164	S1	0.81***		0.73	
		S2	0.72***		0.85	
Qspad.acs-1A.2	Xgwm135-Xwmc304	S1	0.76***		0.91	
		S2	0.65***		0.63	
		S3	0.48***	−0.66***(E2$_{DS}$), −0.72***(E3$_{DS}$), 0.78***(E2$_{WW}$)	0.43	2.24
Qspad.acs-1A.3	Xgwm357-Xgwm633	S2	0.72***	−0.85***(E1$_{WW}$)	0.87	2.54
Qspad.acs-1B.1	Xgwm11-Xwmc626	S1	−0.80***		0.63	
		S2	−1.10***	0.98***(E1$_{DS}$), −0.92***(E3$_{WW}$)	0.92	3.24
		S3	−0.76***		1.05	
		S4	−0.41***		1.31	
Qspad.acs-1B.2	Xmag2064-Xwmc694	S2	−0.76***	0.88***(E1$_{DS}$), 0.83***(E3$_{DS}$), −0.75***(E2$_{WW}$)	0.82	1.64
Qspad.acs-1B.3	Xwmc85-Xcfd65	S2	−0.67***		0.68	
		S3	−0.45***		0.84	
Qspad.acs-1B.4	Xwmc830-Xwmc44	S4	0.59***	−0.31**(E1$_{DS}$)	0.83	1.34
Qspad.acs-1B.5	Xwmc582-Xgwm374	S1	−0.59***		0.56	
Qspad.acs-2A.1	Xgwm339-Xgwm95	S1	0.77***	−0.93***(E2$_{WW}$), −0.99***(E3$_{WW}$)	1.05	1.71
		S3	0.46***	−0.56*** (E1$_{DS}$)	0.69	0.87
Qspad.acs-2A.2	Xgwm249-Xcfa2263	S1	−0.57***		0.59	
Qspad.acs-2A.3	Xgwm558-Xbarc208	S1	0.94***	0.34**(E2$_{WW}$)	1.06	1.14
		S2	0.84***	1.12***(E2$_{DS}$), −1.03***(E2$_{WW}$)	1.16	5.42
		S3	0.70***		1.80	
		S4	0.42***	−0.49***(E1$_{DS}$), 0.40***(E2$_{WW}$), 0.49***(E3$_{WW}$)	0.92	1.91
Qspad.acs-2A.4	Xmag2150-Xgwm339	S2	0.75***		0.95	
Qspad.acs-2A.5	Xwmc296-Xgwm122	S4	0.49***	−0.56***(E1$_{DS}$), 0.29** (E2$_{WW}$)	0.84	1.54
		S5	0.28*	−0.23* (E3$_{DS}$)	0.64	2.74
Qspad.acs-2A.6	Xwmc644-Xmag1730	S1	0.93***		1.29	
Qspad.acs-2B.1	Xcfa2278-Xgwm55	S2	0.64***		0.58	
Qspad.acs-2B.2	Xgwm55-Xbarc128	S2	0.78***		0.77	
		S3	0.61***	0.50***(E2$_{WW}$)	0.72	2.09
Qspad.acs-2B.3	Xgwm388-Xmag3319	S3	0.65***		0.59	
		S4	0.49***		0.55	
Qspad.acs-2B.4	Xwmc223-Xbarc101	S2	0.83***		0.78	
Qspad.acs-3A.1	Xwmc11-Xgwm391	S2	−0.67***	0.78***(E2$_{DS}$)	0.76	1.56
Qspad.acs-3A.2	Xwmc532-Xgwm674	S3	−0.55***	0.65*** (E3$_{DS}$), −0.66*** (E2$_{WW}$), −0.59*** (E3$_{WW}$)	0.52	2.15
		S4	−0.41***		0.40	
Qspad.acs-3A.3	Xcfa2234-Xwmc695	S4	0.50***		0.71	
Qspad.acs-3B.1	Xwmc808-Xbarc102	S3	0.58***	−0.41***(E1$_{DS}$), −0.45***(E2$_{DS}$), 0.47***(E1$_{WW}$)	0.42	1.66
		S4	0.44***		0.63	
Qspad.acs-3B.2	Xmag3356-Xwmc291	S3	0.61***		0.75	
Qspad.acs-3B.3	Xwmc540-Xgwm566	S1	−0.64***		0.75	
Qspad.acs-4A.1	Xmag987-Xbarc78	S5	−0.32**	−0.43*** (E2$_{DS}$), −0.41*** (E3$_{DS}$), 0.30** (E1$_{WW}$)	0.61	2.39
Qspad.acs-4B.1	Xgwm513-Xbarc1142	S4	−0.47***	−0.55***(E1$_{DS}$), 0.25* (E3$_{WW}$)	0.37	1.58
Qspad.acs-4B.2	Xgwm540-Xcfd2	S4	0.62***		0.50	
Qspad.acs-4B.3	Xgwm149-Xgwm495	S3	0.48***		0.38	
Qspad.acs-5B.1	Xmag532-Xgwm499	S2	0.51***		0.44	
Qspad.acs-5B.2	Xwmc734-Xwmc235	S1	−0.88***	−0.25*(E3$_{WW}$)	0.86	0.89
		S2	−0.76***		0.86	
		S3	−0.64***		1.11	
		S4	−0.49**	−0.52*** (E1$_{DS}$), 0.30**(E3$_{WW}$)	0.90	1.76
Qspad.acs-5B.3	Xwmc235-Xcfd10	S3	−0.67***		0.89	
Qspad.acs-5B.4	Xwmc415-Xwmc508	S3	0.45***		0.69	
Qspad.acs-5D.1	Xwmc161-Xgwm565	S4	−0.55***	−0.62*** (E2$_{DS}$), −0.58*** (E3$_{DS}$), 0.31** (E1$_{WW}$)	0.36	2.05

(Continued)

TABLE 5 | Continued

QTL	Flanking markers	Stage[a]	A[b]	AE[c]	$H^2(A)$[b] (%)	$H^2(AE)$[c] (%)
Qspad.acs-5D.2	Xbarc205-Xgwm232	S5	0.33**	-0.39^{***} (E2$_{DS}$), 0.22*(E1$_{WW}$)	0.47	1.91
Qspad.acs-6A.1	Xwmc201-Xwmc684	S2	0.35**	-0.39^{***}(E2$_{WW}$), 0.45***(E3$_{WW}$)	0.49	1.56
Qspad.acs-6A.2	Xbarc171-Xgwm427	S2	-0.59^{***}	0.28*(E2$_{WW}$), 0.35**(E3$_{WW}$)	0.28	1.26
		S3	-0.40^{***}		0.36	
		S4	-0.39^{***}		0.24	
Qspad.acs-6A.3	Xgwm169-Xwmc580	S3	-0.72^{***}	-0.30^{**}(E1$_{DS}$), -0.33^{**}(E3$_{DS}$)	0.60	1.14
Qspad.acs-6A.4	Xbarc113-Xwmc621	S4	-0.69^{***}	-0.72^{***}(E2$_{DS}$), -0.76^{***}(E3$_{DS}$), 0.39***(E3$_{WW}$)	0.56	2.49
Qspad.acs-6B.1	Xgwm508-Xmag1378	S1	0.62***		0.70	
		S3	0.44***		0.38	
Qspad.acs-6B.2	Xcfd13-Xwmc737	S4	0.42***		0.28	
Qspad.acs-6B.3	Xmag3469-Xgwm644	S3	-0.51^{***}		0.32	
Qspad.acs-6B.4	Xgwm644-Xmag1266	S3	-0.40^{***}		0.25	
Qspad.acs-7A.1	Xwmc603-Xwmc116	S3	0.59***	0.73***(E2$_{DS}$), 0.81***(E3$_{DS}$), -0.72^{***}(E2$_{WW}$)	0.56	2.26
Qspad.acs-7A.2	Xwmc273-Xwmc83	S4	-0.45^{***}		0.26	
Qspad.acs-7B.1	Xgwm302-Xbarc258	S3	-0.57^{***}	-0.79^{***}(E1$_{DS}$), -0.72^{***}(E3$_{DS}$)	0.31	2.32
		S5	-0.39^{***}		0.36	
Qspad.acs-7B.2	Xwmc517-Xbarc315	S3	-0.73^{***}	-0.26^{*}(E3$_{DS}$), 0.29**(E3$_{WW}$)	0.60	0.94
Qspad.acs-7B.3	Xwmc311-Xgwm611	S4	-0.62^{***}		0.55	
Qspad.acs-7D.1	Xgwm428-Xbarc111	S3	-0.66^{***}	-0.83^{***}(E2$_{DS}$), -0.80^{***}(E3$_{DS}$), 0. 85***(E1$_{WW}$)	0.54	2.67
Qspad.acs-7D.2	Xgwm37-Xwmc634	S4	-0.43^{***}	-0.53^{***}(E1$_{DS}$), -0.50^{***}(E2$_{DS}$), 0.51***(E2$_{WW}$)	0.25	1.80
		S5	-0.26^{*}	-0.30^{***}(E3$_{DS}$)	0.54	1.56
Qspad.acs-7D.3	Xcfd46-Xwmc438	S3	-0.73^{***}	-0.87^{***}(E1$_{DS}$), -0.80^{***}(E2$_{DS}$)	0.68	2.52

[a]S1–S5 are as shown in **Table 1**. [b]A, the additive effect. A positive value indicates the genetic effect from Q9086 allele, and a negative value represents the genetic effect from Longjian 19 allele; *$P \leq 0.01$, **$P \leq 0.005$, and ***$P \leq 0.001$; $H^2(A)$ (%) indicates the proportion of phenotypic variance explained by additive QTL. [c]AE, the additive QTL × environment interaction effects in drought stress (DS) and the well-watered (WW) conditions in E1–E3 shown in **Table 1**. $H^2(AE)$(%) indicates the phenotypic variance explained by additive QTL × environment interaction.
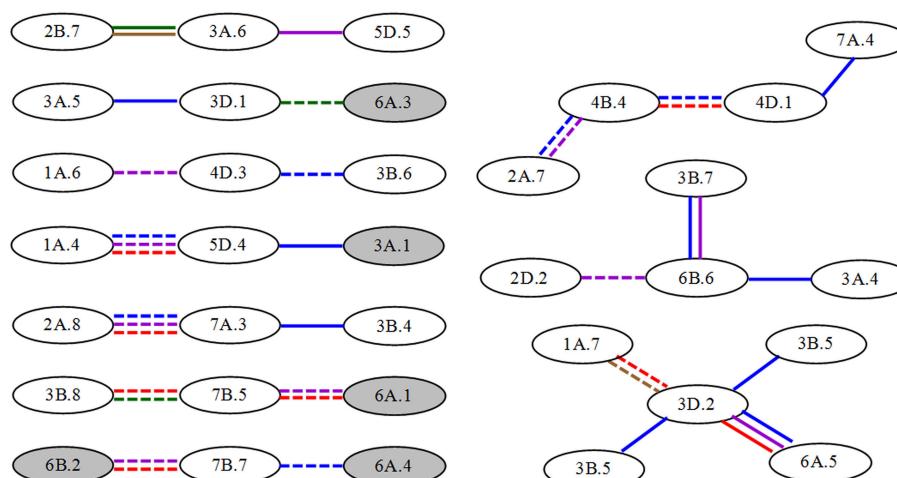
FIGURE 2 | Epistatic QTL network for unconditional greenness of flag leaf at different stages in wheat RILs. Gray and white ellipses indicate A-QTLs and non-individual QTLs, respectively. Solid and dashed lines represent QTL interactions with positive and to negative AA effects, respectively. The different colors such as blue, purple, red, green and brown mean the stages from S1 to S5, respectively.

in different periods across environments. By contrast, most of them reacting to the WW across environments exhibited significantly positive AE (0.25* to 0.80***), with exception of two QTLs, Qspad.acs-2A.1 and Qspad.acs-5B.2, behaving significantly negative AE (-0.25^{*} to -0.99^{***}) in S1|S0. This indicated that the net expression of one locus in the specific period was sensitive to water supply with alternative effect directions, and generally highlighted in up-regulation by the WW.

TABLE 6 | Conditional additive and interacting effects of QTL × environment of identified QTLs for the greenness of flag leaf.

QTL	Flanking markers	Stage[a]	A[b]	AE[c]	H²(A)[b](%)	H²(AE)[c](%)
Qspad.acs-1A.1	Xgwm633-Xgwm164	S1|S0	0.81***		0.73	
Qspad.acs-1A.2	Xgwm135-Xwmc304	S1|S0	0.76***		0.91	
Qspad.acs-1A.6	Xgwm164-Xcfd59	S2|S1	0.64***		0.66	
Qspad.acs-1B.1	Xgwm11-Xwmc626	S1|S0	−0.80***		0.63	
Qspad.acs-1B.2	Xmag2064-Xwmc694	S3|S2	0.59***	0.39***(E2$_{WW}$), 0.36***(E3$_{WW}$)	0.60	1.08
Qspad.acs-1B.5	Xwmc582-Xgwm374	S1|S0	−0.59***		0.56	
Qspad.acs-2A.1	Xgwm339-Xgwm95	S1|S0	0.77***	−0.93***(E2$_{WW}$), −0.99***(E3$_{WW}$)	1.05	1.71
Qspad.acs-2A.2	Xgwm249-Xcfa2263	S1|S0	−0.57***		0.59	
Qspad.acs-2A.3	Xgwm558-Xbarc208	S1|S0	0.94***	0.34**(E2$_{WW}$)	1.06	1.14
Qspad.acs-2A.6	Xwmc644-Xmag1730	S1|S0	0.93***		1.29	
Qspad.acs-2A.8	Xgwm515-Xwmc644	S4|S3	−0.47***	0.25*(E1$_{WW}$)	0.49	0.82
Qspad.acs-2B.4	Xwmc223-Xbarc101	S3|S2	−0.58***		0.50	
Qspad.acs-3A.5	Xcfd193-Xcfa2234	S2|S1	0.61***		0.84	
Qspad.acs-3A.7	Xmag3082-Xbarc19	S3|S2	−0.72***		0.99	
Qspad.acs-3B.3	Xwmc540-Xgwm566	S1|S0	−0.64***		0.75	
Qspad.acs-3B.9	Xgwm108-Xpsp3035	S2|S1	0.53***	0.52***(E2$_{WW}$), 0.54***(E3$_{WW}$)	0.76	2.05
Qspad.acs-3B.10	Xbarc173-Xgwm284	S2|S1	0.69***	−0.43***(E2$_{DS}$)	0.63	1.22
Qspad.acs-4A.1	Xmag987-Xbarc78	S2|S1	−0.36**	−0.76***(E1$_{DS}$), −0.78***(E3$_{DS}$)	0.40	3.80
Qspad.acs-4A.3	Xksum51-Xgwm637	S2|S1	−0.54***		0.73	
Qspad.acs-4B.1	Xgwm513-Xbarc1142	S4|S3	−0.39**	−0.76***(E2$_{DS}$), −0.70***(E3$_{DS}$), 0.75***(E3$_{WW}$)	0.36	3.25
Qspad.acs-4B.5	Xksum244-Xmag2055	S3|S2	−0.31**	−0.77***(E2$_{DS}$), −0.78***(E3$_{DS}$), 0.72***(E2$_{WW}$)	0.41	3.19
Qspad.acs-5B.2	Xwmc734-Xwmc235	S1|S0	−0.88***	−0.25**(E3$_{WW}$)	0.86	0.89
Qspad.acs-5B.8	Xgdm116-Xbarc140	S2|S1	−0.62***		0.74	
Qspad.acs-5D.2	Xbarc205-Xgwm232	S3|S2	0.33**	−0.41***(E1$_{DS}$), 0.32** (E3$_{WW}$)	0.39	2.72
Qspad.acs-5D.6	Xgwm565-Xwmc443	S4|S3	0.41***	−0.32**(E2$_{DS}$), 0.26* (E3$_{WW}$)	0.75	3.52
Qspad.acs-6A.1	Xwmc201-Xwmc684	S2|S1	0.73***		0.86	
Qspad.acs-6A.3	Xgwm169-Xwmc580	S4|S3	−0.48***	−0.41***(E2$_{DS}$), 0.47***(E3$_{WW}$)	0.49	2.53
Qspad.acs-6A.6	Xgwm570-Xwmc553	S3|S2	−0.55***		0.66	
Qspad.acs-6B.1	Xgwm508-Xmag1378	S1|S0	0.62***		0.70	
Qspad.acs-6B.3	Xmag3469-Xgwm644	S5|S4	−0.46***		0.40	
Qspad.acs-6B.7	Xmag1266-Xmag982	S3|S2	−0.43***		0.38	
Qspad.acs-7A.6	Xbarc1034-Xwmc273	S4|S3	0.39***	−0.41***(E2$_{DS}$), 0.21*(E2$_{WW}$)	0.74	2.63
Qspad.acs-7A.7	Xwmc405-Xpsp3050	S3|S2	0.48**	−0.46***(E1$_{DS}$), 0.34**(E1$_{WW}$)	0.36	2.80
Qspad.acs-7B.2	Xwmc517-Xbarc315	S5|S4	0.30**	−0.33**(E2$_{DS}$)	0.57	2.31
Qspad.acs-7B.8	Xbarc315-Xwmc311	S2|S1	−0.63***		0.81	
Qspad.acs-7D.1	Xgwm428-Xbarc111	S3|S2	0.49***	−0.49***(E1$_{DS}$), −0.47***(E2$_{DS}$), 0.47***(E3$_{WW}$)	0.59	1.68
Qspad.acs-7D.3	Xcfd46-Xwmc438	S2|S1	−0.34**	−0.75***(E1$_{DS}$), −0.79***(E3$_{DS}$), 0.80***(E1$_{WW}$)	0.51	3.01

[a]S1|S0–S5|S4 are as shown in *Table 2*. [b]A, the additive effect. A positive value indicates the genetic effect from Q9086 allele, and a negative value represents the genetic effect from Longjian 19 allele; *P ≤ 0.01, **P ≤ 0.005, and ***P ≤ 0.001; H²(A) (%) indicates the proportion of phenotypic variance explained by additive QTL. [c]AE, the additive QTL × environment interaction effects in drought stress (DS) and the well-watered (WW) conditions in E1–E3 shown in *Table 1*. H²(AE)(%) indicates the phenotypic variance explained by additive QTL × environment interaction.

Total of 29 epistatic pairs were mapped on the nearly same chromosomes as unconditional loci other than chromosome 1D, accounting for from 0.66 to 3.29% of the phenotypic variation (Table S2). Of these, 10 pairs appeared significantly positive *AA* effects (0.72*** to 1.32***), whereas the other 19 pairs behaved significantly negative effects (−0.58*** to −1.51***). Apart from one pair detectable in two continual periods (S2|S1 and S3|S2), the other 28 pairs were identified in a specific period, especially in S1|S0—i.e., 25 pairs available. This further confirmed that epistatic effects were short lived but highly predominant in early period. In addition, 14 pairs got involved in E-QEIs, explaining from 0.86 to 4.08% of the phenotypic variation. Nearly 43% of them occurred in S1|S0, the remainders were involved in other periods. Although expressional patterns of AA-QTLs differed from one to another period in response to water environments, all of them exhibited negative effects (−0.54*** to −0.95***) in the DS, but positive effects (0.52*** to 0.98***) in the WW. This indicated that the expressions of conditional AA-QTLs were also enhanced by the WW, in accordance with results of the unconditional analysis.

For the compositions of conditional epistatic loci, most of epistatic interactions were performed by non-individual QTLs, besides seven significant A-QTLs involving epitasis. Thus, QTL networks were involved at different interaction levels (**Figure 3**), but simpler than those of unconditional AA-QTLs. For example, two significant A-QTLs and 10 non-individual QTLs were participated in four three-locus-interaction networks. One A-QTL and other eight non-individual QTLs were involved in two bigger networks from four or five-locus interactions, respectively. In these QTL networks, most of epistatic interactions (nearly 65%) had positive *AA* effects and more occurred in S1|S0, indicative of predominance in parent-type effects and in time-independent expressions in the first period.

Comparative Analysis between Conditional and Unconditional QTLs

Following the above mapping results, 22 A-QTLs (**Tables 5**, **6**) and 25 pairs of AA-QTLs (Tables S1, S2) were common between unconditional and conditional mapping strategies across five measuring stages. With regard to these common loci, 11 A-QTLs and 24 AA-QTLs were detected at the first stage, where both genetic effects and contribution rates for the phenotypic variation were exactly equal to conditional ones. However, the remainder loci expressed at one or more stages, responding to specific stages and water environments, remarkably differed in *A* and *AE* effects, as well as in their contribution rates between two sets of mapping strategies. This suggested that common loci also behaved alternatively for inheritance of leaf greenness, with cumulative or net genetic effects highly modulated by developmental courses and water environments. Opposite to these common loci, 28 A-QTLs and 13 pairs of AA-QTLs were specifically detected by the unconditional mapping, whereas, in this way, 15 A-QTLs and 4 pairs of AA-QTLs were identified only by the conditional mapping. Therefore, by combining these two sets of mapping strategies based on time-dependent evaluation, more novel loci might be available and further exposed dynamic expression of polygenes governing the ontogeny of leaf greenness. Furthermore, an interesting feature was that both common and specifically-expressed A-QTLs detected by two sets of mapping strategies were nearly distributed in cluster occurred in specific neighboring marker intervals in several chromosomes (**Figure 1**). For example, two to seven loci shared neighboring intervals with flanking markers from Xgwm357 to Xwmc304 on chromosome 1A, from Xwmc85 to Xgwm374 on chromosome 1B, from Xwmc 296 to Xmag1730 on chromosome 2A, and so on, respectively. This indicated that specific marker intervals might carry a wealth of genetic information for the ontogeny of leaf greenness.

With regard to the general effects of genetic components from two sets of mapping analysis, the dynamics of QTL expressions was highly visible in the measuring duration of ontogeny (**Figures 4A,B**). The general effects almost appeared negative, but their absolute values greatly altered in different periods, as the trend in noticeable increase before the third phase and thereafter decrease to the minimum in the final period. Except the equal effect values emerging in the first period, the

other effect values of unconditional QTLs were considerably higher than those of conditional ones. It could be perceived that the genetic regulation for the development of leaf greenness by cumulative effects was stronger than that by net effects. On the other hand, both unconditional and conditional *AAE* effects were greatly higher than other genetic components in the first period. During the second period, superior genetic components were associated with *AA* effects in unconditional mapping, but *AE* and *AAE* in conditional mapping. Thereafter, superior genetic components almost tended to the similarity between two sets of mapping strategies, highlighting in *AA* and/or *AE* effects. This suggested that, in each developmental period, QTLs governing the ontogeny of leaf greenness were also expressed dynamically, ascribed to the specific effect strength of genetic components.

On the other hand, the general contribution rates explaining the phenotypic variation were also further illustrated the dynamic characteristics of QTL expressions for ontogeny of leaf greenness (**Figures 4C,D**). The trend appeared decline in whole measuring duration, whereas drift magnitudes were less in unconditional QTLs than conditional ones. During the early period (S1–S3), unconditional QTLs showed higher general contribution rates (69.06–75.24%), and then decreased to the bottom (17.82%) at last stage. However, the maximum (71.23%) of conditional loci happened in S1|S0, and then sharply declined to the minimum (3.28%) in S5|S4. This indicated that cumulative genetic effects could be maintained longer and stronger activation, whereas net genetic effects were weaker and short lived. By contrast to the contribution rates of genetic components, the predominant performances were mainly attributed to *AA* effects at S1–S2 or *AE* effects at S3–S5 by the unconditional mapping, and *AA* effects in S1|S0 or *AE* effects in S2|S1 to S5|S4 by the conditional mapping. Owing to the above analysis, both genetic effects and their contribution rates altered dynamically with a similar variation trend in progressive reduction during the whole measuring period, but mainly expressed in the early duration. Furthermore, the performances of *AA* or *AE* effects were highly predominant to determine the developmental genetics of leaf greenness.

DISCUSSION

According to the theory of developmental genetics, functional genes will be expressed dynamically in response to different growth stages. Furthermore, the expression pattern essentially occurs through the actions and interactions of polygenes during the ontogeny and is also flexibly modified by environments (Atchley and Zhu, 1997; Wu et al., 1999). Thus, the combination of unconditional with conditional mapping is verified to be an efficient approach to dissect dynamic expressions of developmental QTLs and reveal the inheritance of quantitative traits (Wu et al., 2010; Li S. et al., 2012). The strategy has been successfully applied to developmental QTLs analysis for agronomic and physiological traits in many crops (Liu et al., 2010; Wu et al., 2010; Han et al., 2011; Li S. et al., 2012). So far, although a multitude of previous studies have identified a wealth of QTLs for leaf staygreen and its associated traits by the method of unconditional mapping (Verma et al., 2004; Kumar et al., 2010; Li H. et al., 2012; Naruoka et al., 2012;
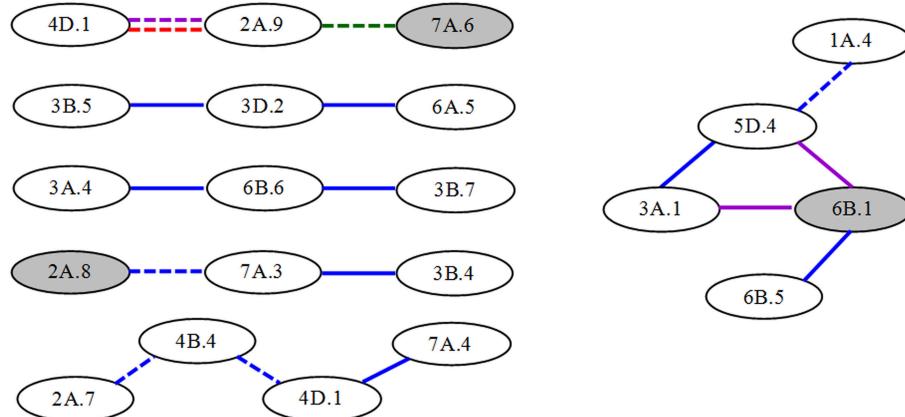
FIGURE 3 | Epistatic QTL network for conditional greenness of flag leaf at different stages in wheat RILs. Gray and white ellipses indicate A-QTLs and non-individual QTLs, respectively. Solid and dashed lines represent QTL interactions with positive and to negative *AA* effects, respectively. The different colors such as blue, purple, red and green mean the stages from S1|S0 to S4|S3, respectively.
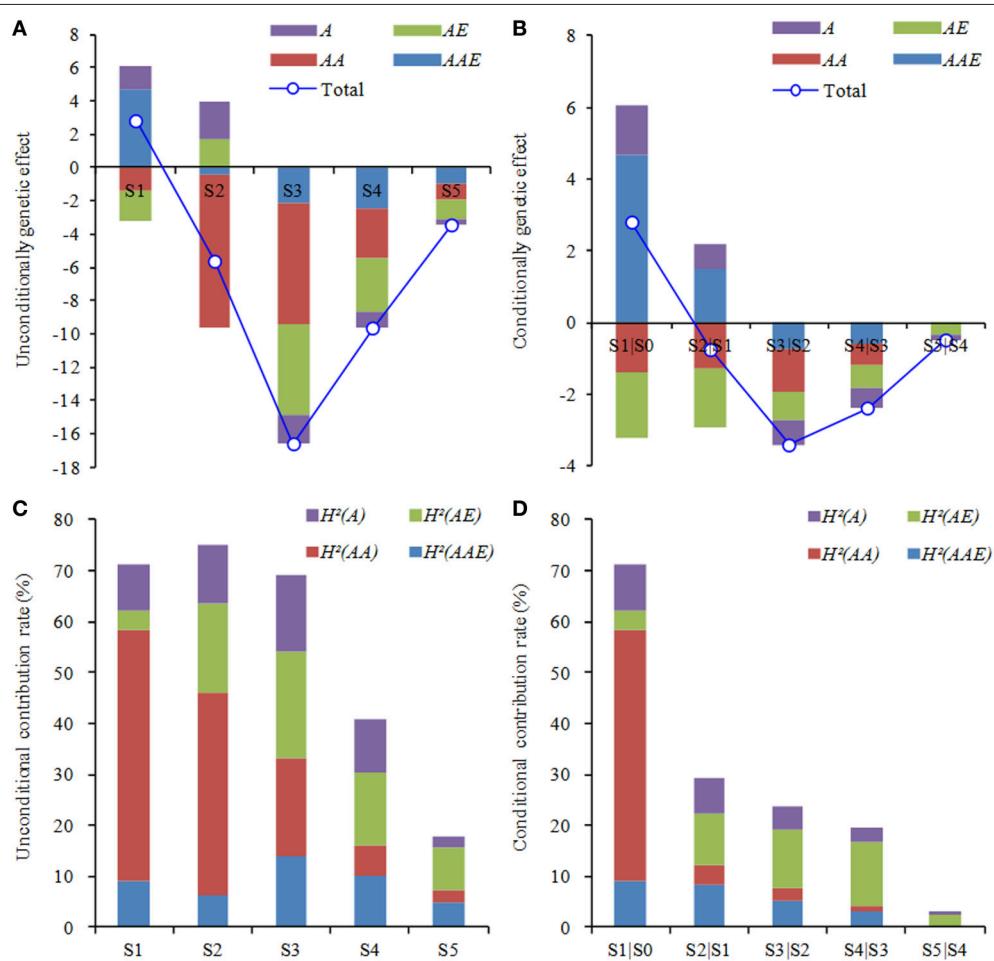
FIGURE 4 | Components of genetic effects (A,B) and contribution (C,D) to development of the greenness of flag leaf in wheat RILs. *A*, the additive effect; *AE*, the additive QTL × environment interaction effects; *AA*, the epistatic effect; and *AAE*, the interaction effect of epistatic QTL × environment. $H^2(A)$, $H^2(AE)$, $H^2(AA)$, and $H^2(AAE)$ indicate the phenotypic variance explained by corresponding genetic effect. S1–S5 are as shown in **Table 1**, and S1|S0 to S5|S4 are as shown in **Table 2**.

Barakat et al., 2013; Czyczyło-Mysza et al., 2013), it seems rather obscure on genetic information of the ontogeny of leaf staygreen in wheat, because of the complexity. In this study, by employing unconditional and conditional mapping, a total of 65 A-QTLs and 42 pairs of AA-QTLs for the ontogeny leaf staygreen after flowering were detected on almost all 21 chromosomes except 5A across water environments, individually explaining from 0.24 to 3.29% of the phenotypic variation, indicative of typical quantitative traits controlled by minor-effect polygenes. Regardless of cumulative or net genetic effects, all loci were highly characteristic of time-dependent expressions. In this context, most of them were associated with specific development stage, while no locus was continually detectable over measuring time, except few loci active in two or more stages. Obviously, more QTLs were detected in earlier development stages and showed higher performances of genetic effects (**Figure 4**). The selective expressions of QTLs might be favorable to the complicated genetic regulations responsible for the ontogeny (Wu et al., 1999, 2010; Li S. et al., 2012). Likewise, the result could interpret why the phenotypic values of leaf greenness varied in subsequently decreased trend during the measuring period (**Table 1**). It was confirmed that the maintenance of leaf greenness might be greatly dependent upon early-expressed QTLs. And in another aspect, some major QTLs could be easily neglected in case only evaluating them by the phenotypic data at a specific stage, especially in the later period. Similar findings were also observed in wheat plant height (Wu et al., 2010), soybean pod number (Sun et al., 2006), and rice tiller trait (Yang et al., 2006).

By contrast, a total of 22 A-QTLs and 25 pairs of AA-QTLs were common between two sets of mapping methods. Besides, specifically- and reproducibly-expressed QTLs were more detected by the unconditional analysis (**Table 5**, Table S1). Each reproducibly-expressed A-QTL or AA-QTL showed significant genetic effects with same effect directions. For instance, one A-QTL, *Qspad.acs-1B.1*, showed negative *A* effects (-0.41^{***} to -1.10^{***}) across four stages (S1–S4). One pair, *Qspad.acs-3D.2 × Qspad.acs-6A.5*, exhibited positive *AA* effects (1.00^{***} to 1.32^{***}) across three stages (S1–S3). This could be explained by the fact that unconditional QTLs were attributed to the cumulative expression from the initial time to stage *t* (Zhu, 1995). On the other hand, conditional QTLs *per se* could interpret the real gene expression in the specific period from stage *t*-1 to *t* (Zhu, 1995; Atchley and Zhu, 1997). Here almost all conditional QTLs were expressed only in a specific period (**Table 6**, Table S2). Other than QTLs detected at the first stage, conditional loci were significantly distinct from unconditional ones—e.g., some loci were observed with unconditional effects but without any conditional effect, and vice versa. Even though there existed some common QTLs between two mapping strategies, their expression profiles were variable in response to the specific measuring stage. For example, *Qspad.acs-1B.2* showed the unconditional *A* effect with -0.76*** at S2, but the conditional effect with 0.59*** in S3|S2. This indicated that parental contribution of favorable alleles at the same map position was also variable along with the development of leaf greenness. Therefore, the above evidence clearly suggests that QTL expressions for the ontogeny of leaf greenness are

time-dependent. By combining unconditional QTL mapping with conditional QTL one of time-dependent measures, it is quite possible to reveal the dynamic gene expressions for the development of leaf staygreen.

The ability of a genotype to adapt its phenotype to different environments is referred as phenotypic plasticity (Ungerer et al., 2003). The phenotypic plasticity of quantitative traits arises in nature from QEIs at molecular levels (Campbell et al., 2003). Several examples of QEIs for developmental traits showed that the expression of particular chromosome regions differs across environments (Wu et al., 2010; Li S. et al., 2012). In this study, 53.0% of A-QTLs and 62.8% of AA-QTLs for the ontogeny of leaf greenness were significantly interacted with water environments, explaining from 0.82 to 5.42% and from 0.82 to 4.08% of the phenotypic variation, respectively. This indicated that the expressions of QTLs governing the ontogeny of leaf greenness were more susceptible to water environments, and to a certain extent, environmentally dependent. As for the attributes of QTL expressions influenced by water availability, there were significant differences among different developmental stages and between two mapping strategies. Firstly, although all of QEIs reacted to at least one water environment, the stages responsive to QEIs were widely various. Generally, most of them were highlighted in the periods after the first stage. For example, about 90% of unconditional A-QEIs and 85% of conditional A-QEIs flexibly occurred in the mid-anaphase, indicating that water environments highly affected the expressions of developmental QTLs in the later period of growth. Secondly, more QEIs occurred in unconditional QTLs than in conditional ones. For example, 27 A-QEIs and 24 E-QEIs were detected by the unconditional QTL mapping, where only 18 and 14 QEIs were unraveled by the conditional QTL mapping. Especially in the unconditional QTL mapping, four A-QEIs and seven E-QEIs had continually-expressed *AE* or *AAE* effects at two or three stages. However, the similar result was observed with only one E-QEI in the conditional QTL mapping. Of course, QEI effects were thus greatly distinct, except the QEIs involved in the first stage. This indicated that cumulative genetic effects were more prone to interact with water environments than net genetic effects. Thirdly, the behaviors of QEIs differed from responses to specific water environments. For example, nine unconditional A-QTLs expressed during the period from S1 to S3 got involved in negative *AE* effects with WW environments, but positive effects with DS environments. Thereafter, the case was opposite. This indicated that *AE* effects of these loci exposed during the early period could be more up-regulated by the DS. However, the other QEIs showed negative interaction effects with the DS, but positive effects with the WW. This clearly suggested that putative QTL expressions for the ontogeny of leaf greenness were environmentally-dependent and significantly up-regulated by WW environments. These results could provide detailed information on the variable performance of quantitative loci controlling the development of leaf greenness under different water environments.

Regarding the genetic components of leaf staygreen in wheat, several previous studies were elucidated that additive effects are predominant (Silva et al., 2000; Verma et al., 2004; Joshi

et al., 2007; Kumar et al., 2010). In some cases, epistatic effects (Zhang et al., 2009a,b, 2010; Kumar et al., 2012) and QEI effects (Yang et al., 2007; Peleg et al., 2009) were also considered to be important. However, current genetic gains for leaf staygreen trait are made only by the traditional mapping analysis depending on phenotypic data at one time point. Indeed, it is inadequate to deeply dissect genetic components and their dynamical behaviors for the ontogeny of leaf staygreen. In the present study, the genetic behaviors for respective components were obviously dynamical and time-dependent during the whole measuring stages (**Figure 4**). Of these, general effects were almost negative, but had a large alternation in the third duration, which might essentially illustrate why phenotypic values of leaf greenness always decreased in the duration of growth, along with significant reduction around the stage of S3 across environments (**Table 1**). In view of the respective effects of genetic components, *AA* and QEIs effects were superior to other genetic components by two sets of mapping analysis, while effect directions and magnitudes varied in response to specific periods. The similar result was also observed in the developmental behavior of rice tiller number (Liu et al., 2010). On the other hand, the general contribution rates showed progressive reduction over the measuring time (**Figure 4**), consistent with the findings from the developmental genetic attributes of plant height (Wang et al., 2010; Wu et al., 2010) and grain weight (Li S. et al., 2012) in wheat. Thus, genetic effects active in early stages might play a critical role in modulating the phenotypic variation of development traits, due to higher contribution rates. Similar to the variations of genetic effects, the respective contribution rates of genetic components were highly flexible, whereas their magnitudes were incompletely equal to those of genetic effects. It was considered that some specific effects might be counteracted or pyramided each other during the development of quantitative traits (Wu et al., 2010). Nevertheless, the contribution rates of *AA* and *AE* effects were more predominant than other effects. The finding is confirmed that the action a specific gene to one quantitative phenotype is the collective property of a network of polygenes and/or its tight interactions with environments, rather than the behavior of a single gene (Wade, 2002; Malmberg et al., 2005).

In this study, although putative A-QTLs were widely dispersed on almost all chromosomes except 1D–4D, 5A, and 6D, they were nearly concentrated in specific neighboring marker intervals in several chromosomes. Moreover, these important marker intervals harbored many reproducibly-expressed QTLs in two or more periods (**Figure 1**). Using a wheat microsatellite consensus map (Somers et al., 2004) as a reference map, some QTLs controlling leaf greenness in the present work have been previously mapped on similar chromosomal regions. For example, four A-QTLs were detected in the marker interval from Xgwm357 to Xwmc304 on chromosome 1A, which overlapped with the location of a staygreen QTL (*Qsg.bhu-1A*) reported by Kumar et al. (2010). The reproducibly-expressed QTL, *Qspad.acs-1B.1*, shared the similar interval (Xgwm11–Xwmc626) of putative loci controlling leaf chlorophyll content (Czyczyło-Mysza et al., 2013). Another reproducibly-expressed QTL, *Qspad.acs-2A.1*, was possibly equal to *Qchl a+b.igdb-2A* (Li

H. et al., 2012), as both were very close to the marker Xgwm339. Likewise, the other loci, such as *Qspad.acs-2B.1*, *Qspad.acs-3B.2*, *Qspad.acs-4A.3*, *Qspad.acs-5B.1*, *Qspad.acs-5D.2*, *Qspad.acs-7B.3*, *Qspad.acs-7D.1*, and *Qspad.acs-7D.2*, were identical or adjacent to the corresponding loci governing leaf chlorophyll content or its component content observed in different wheat populations (Zhang et al., 2009a; Kumar et al., 2010; Li H. et al., 2012; Czyczyło-Mysza et al., 2013). These common QTLs and their tightly-linked molecular markers would be of great importance for MAS. By contrast to current results reported by Czyczyło-Mysza et al. (2013), some typical marker intervals harboring QTLs clusters especially on chromosomes of 2B, 3A, 4B, 5B, 6B, and 7A (**Figure 1**) were also co-located considerable loci related to chlorophyll fluorescence parameters, carotenoid content, and even agronomic traits. For example, in the marker interval from Xwmc494 to Xgwm193 on chromosome 6B, five A-QTLs were mapped in this study, in which loci for related traits, such as amount of excitation energy trapped in PSII reaction centers (ET_0/CS_m), overall performance index of PSII photochemistry (PI), the maximum photochemical efficiency (Fv/Fm), light energy absorption (ABS/CS_m), and grain yield for the main stem (GWE), were also located. Another marker interval from Xwmc11 to Xbarc19 harboring five loci for leaf greenness was possibly equivalent to the QTL-rich interval from Xwmc11 to Xcfa2262, where Czyczyło-Mysza et al. (2013) identified several QTLs controlling leaf carotenoid content, ET_0/CS_m, number of active reaction centers (RC/CS_m), and GWE. This indicated that the hot-spot regions of QTLs could carry a wealth of genetic information on leaf greenness and its associated traits of wheat. Therefore, further studies on the possibility of a tight linkage or genetic pleiotropism on the QTL-rich regions will be very important, so as to elucidate the genetic nature of leaf greenness, and to use them in wheat improvement program.

CONCLUSION

Flag leaf greenness of wheat in reproductive phase was controlled by minor-effect polygenes, which were expressed selectively as a time- and environmentally-dependent pattern during ontogeny. No single QTL was continually active in measuring period. But more loci were identified in early development periods, showing the higher performance of genetic effects. QEIs mainly happened in the mid-anaphase of development, where drought stress was more conducted with negative regulation on the QTL expressions. By contrast, *AA* and *AE* effects could be predominant in regulating phenotypic variations during the ontogeny of leaf greenness. In this regards, cumulative genetic effects could be maintained longer and stronger activation, whereas net genetic effects were weaker and short lived. Several QTL cluster regions were suggestive of tight linkage or expression pleiotropy in the inheritance of these traits. Some reproducibly-expressed QTLs or common loci consistent with previously detected would be useful to the genetic improvement of staygreen types in wheat through MAS, especially in water-deficit environments.

AUTHOR CONTRIBUTIONS

DY designed the whole experiments and wrote the manuscript. ML and YL performed statistic analysis. LC, JC, and HC accomplished the phenotypic observation and measurement. SC performed the management of field experiments. All authors have read the manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

REFERENCES

Atchley, W. R., and Zhu, J. (1997). Developmental quantitative genetics, conditional epigenetic variability and growth in mice. *Genetics* 147, 765–776.

Barakat, M. N., Wahba, L. E., and Milad, S. I. (2013). Molecular mapping of QTLs for wheat flag leaf senescence under water-stress. *Biol. Plant.* 57, 79–84. doi: 10.1007/s10535-012-0138-7

Biswal, A. K., and Kohli, A. (2013). Cereal flag leaf adaptations for grain yield under drought: knowledge status and gaps. *Mol. Breed.* 31, 749–766. doi: 10.1007/s11032-013-9847-7

Borrell, A., Hammer, G., and Oosterom, E. V. (2001). Stay-green: a consequence of the balance between supply and demand for nitrogen during grain filling? *Ann. Appl. Biol.* 138, 91–95. doi: 10.1111/j.1744-7348.2001.tb00088.x

Campbell, B. T., Baenziger, P. S., Gill, K. S., Eskridge, K. M., Budak, H., Erayman, M., et al. (2003). Identification of QTLs and environmental interactions associated with agronomic traits on chromosome 3A of wheat. *Crop Sci.* 43, 1493–1505. doi: 10.2135/cropsci2003.1493

Cao, G., Zhu, J., He, C., Gao, Y., Yan, J., and Wu, P. (2001). Impact of epistasis and QTL×environment interaction on the developmental behavior of plant height in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 103, 153–160. doi: 10.1007/s001220100536

Chen, J., Liang, Y., Hu, X., Wang, X., Tan, F., Zhang, H., et al. (2010). Physiological characterization of 'stay green' wheat cultivars during the grain filling stage under field growing conditions. *Acta Physiol. Plant.* 32, 875–882. doi: 10.1007/s11738-010-0475-0

Christopher, J. T., Manschadi, A. M., Hammer, G. L., and Borrell, A. K. (2008). Developmental and physiological traits associated with high yield and stay-green phenotype in wheat. *Aus. J. Agricul. Res.* 59, 354–364. doi: 10.1071/AR07193

Czyczyło-Mysza, I., Tyrka, M., Marcińska, I., Skrzypek, E., Karbarz, M., Dziurka, M., et al. (2013). Quantitative trait loci for leaf chlorophyll fluorescence parameters, chlorophyll and carotenoid contents in relation to biomass and yield in bread wheat and their chromosome deletion bin assignments. *Mol. Breed.* 32, 189–210. doi: 10.1007/s11032-013-9862-8

Emebiri, L. C. (2013). QTL dissection of the loss of green colour during postanthesis grain maturation in two-rowed barley. *Theor. Appl. Genet.* 126, 1873–1884. doi: 10.1007/s00122-013-2102-0

Farooq, M., Hussain, M., and Siddique, K. H. M. (2014). Drought stress in wheat during flowering and grain-filling periods. *Crit. Rev. Plant Sci.* 33, 331–349. doi: 10.1080/07352689.2014.875291

Gong, Y., Zhang, J., Gao, J., Lu, J., and Wang, J. (2005). Slow export of photoassimilate from stay-green leaves during late grain-filling stage in hybrid winter wheat (*Triticum aestivum* L.). *J. Agron. Crop Sci.* 191, 292–299. doi: 10.1111/j.1439-037X.2005.00173.x

Gregersen, P. L., Culetic, A., Boschian, L., and Krupinska, K. (2013). Plant senescence and crop productivity. *Plant Mol. Biol.* 82, 603–622. doi: 10.1007/s11103-013-0013-8

Gregersen, P. L., Holm, P. B., and Krupinska, K. (2008). Leaf senescence and nutrient remobilisation in barley and wheat. *Plant Biol.* 10, 37–49. doi: 10.1111/j.1438-8677.2008.00114.x

Guóth, A., Tari, I., Gallé, A., Csiszár, J., Pécsváradi, A., Cseuz, L., et al. (2009). Comparison of the drought stress responses of tolerant and sensitive wheat cultivars during grain filling: changes in flag leaf photosynthetic activity, ABA levels, and grain yield. *J. Plant Growth Regul.* 28: 167–176. doi: 10.1007/s00344-009-9085-8

Han, Y., Xie, D., Teng, W., Sun, J., and Li, W. (2012). QTL underlying developmental behaviour of 100-seed weight of soybean. *Plant Breed.* 131, 600–606. doi: 10.1111/j.1439-0523.2012.01987.x

Han, Y., Xie, D., Teng, W., Zhang, S., Chang, W., and Li, W. (2011). Dynamic QTL analysis of linolenic acid content in different developmental stages of soybean seed. *Theor. Appl. Genet.* 122, 1481–1488. doi: 10.1007/s00122-011-1547-2

Harris, K., Subudhi, P. K., Borrell, A., Jordan, D., Rosenow, D., Nguyen, H., et al. (2007). Sorghum stay-green QTL individually reduce post-flowering drought-induced leaf senescence. *J. Exp. Bot.* 58, 327–338. doi: 10.1093/jxb/erl225

Hu, L., Ye, Y., Lv, T., Li, M., Liu, Y., Chang, L., et al. (2015). QTL mapping and genetic analysis for grain weight (*Triticum aestivum*) under different water environments. *Acta Pratacul. Sin.* 24, 118–129. doi: 10.11686/cyxb2015071

Jiang, G., He, Y., Xu, C., Li, X., and Zhang, Q. (2004). The genetic basis of stay-green in rice analyzed in a population of doubled haploid lines derived from an *indica* by *japonica* cross. *Theor. Appl. Genet.* 108, 688–698. doi: 10.1007/s00122-003-1465-z

Joshi, A. K., Kumari, M., Singh, V. P., Reddy, C. M., Kumar, S., Rane, J., et al. (2007). Stay green trait: variation, inheritance and its association with spot blotch resistance in spring wheat (*Triticum aestivum* L.). *Euphytica* 153, 59–71. doi: 10.1007/s10681-006-9235-z

Kassahun, B., Bidinger, F. R., Hash, C. T., and Kuruvinashetti, M. S. (2010). Stay-green expression in early generation sorghum [*Sorghum bicolor* (L.) Moench] QTL introgression lines. *Euphytica* 172, 351–362. doi: 10.1007/s10681-009-0108-0

Kumar, S., Sehgal, S. K., Kumar, U., Prasad, P. V. V., Joshi, A. K., and Gill, B. S. (2012). Genomic characterization of drought tolerance-related traits in spring wheat. *Euphytica* 186, 265–276. doi: 10.1007/s10681-012-0675-3

Kumar, U., Joshi, A. K., Kumari, M., Paliwal, R., Kumar, S., and Röder, M. S. (2010). Identification of QTLs for stay green trait in wheat (*Triticum aestivum* L.) in the 'Chirya 3' × 'Sonalika' population. *Euphytica* 174, 437–445. doi: 10.1007/s10681-010-0155-6

Li, H., Lin, F., Wang, G., Jing, R., Zheng, Q., Li, B., et al. (2012). Quantitative trait loci mapping of dark-induced senescence in winter wheat (*Triticum aestivum*). *J. Integ. Plant Biol.* 54, 33–44. doi: 10.1111/j.1744-7909.2011.01088.x

Li, L., Yang, D., Li, M., Chang, L., Cheng, H., Chai, S., et al. (2013). Effects of source-sink regulation on WSC in vegetative organs and thousand-grain mass of wheat under different water conditions. *Chin. J. Appl. Ecol.* 24, 1879–1888.

Li, S., Wang, C., Chang, X., and Jing, R. (2012). Genetic dissection of developmental behavior of grain weight in wheat under diverse temperature and water regimes. *Genetica* 140, 393–405. doi: 10.1007/s10709-012-9688-z

Liu, G., Zhu, H., Liu, S., Zeng, R., Zhang, Z., Li, W., et al. (2010). Unconditional and conditional QTL mapping for the developmental behavior of tiller number in rice (*Oryza sativa* L.). *Genetica* 138, 885–893. doi: 10.1007/s10709-010-9471-y

Ma, Z., Li, M., Yang, D., Chen, X., Chen, J., Liu, Y., et al. (2014). Relationship between grain filling and accumulation and remobilization of water soluble carbohydrates in leaf and stem of winter wheat during the

grain filling in different water conditions. *Acta Pratacul. Sin.* 23, 68–78. doi: 10.11686/cyxb20140408

Malmberg, R. L., Held, S., Waits, A., and Mauricio, R. (2005). Epistasis for fitness-related quantitative traits in *Arabidopsis thaliana* grown in the field and in the greenhouse. *Genetics* 171, 2013–2027. doi: 10.1534/genetics.105. 046078

Naruoka, Y., Sherman, J. D., Lanning, S. P., Blake, N. K., Martin, J. M., and Talbert, L. E. (2012).Genetic analysis of green leaf duration in spring wheat. *Crop Sci.* 52, 99–109. doi: 10.2135/cropsci2011.05.0269

Nawaz, A., Farooq, M., Cheema, S. A., Yasmeen, A., and Wahid, A. (2013). Stay green character at grain filling ensures resistance against terminal drought in wheat. *Int. J. Agricul. Biol.* 15, 1272–1276.

Peleg, Z., Fahima, T., Krugman, T., Abbo, S., Yakir, D., Korol, A. B., et al. (2009). Genomic dissection of drought resistance in durum wheat×wild emmer wheat recombinant inbreed line population. *Plant Cell Environ.* 32, 758–779. doi: 10.1111/j.1365-3040.2009.01956.x

Silva, S. A., Carvallo, F. I. F., Caetano, V. R., Oliveira, A. C., Coimbra, J. L. M., Vasconcellos, N. J. S., et al. (2000). Genetic basis of stay-green trait in bread wheat. *J. New Seeds* 2, 55–68. doi: 10.1300/J153v02n02_05

Somers, D. J., Isaac, P., and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 109, 1105–1114. doi: 10.1007/s00122-004-1740-7

Sun, D., Li, W., Zhang, Z., Chen, Q., Ning, H., Qiu, L., et al. (2006). Quantitative trait loci analysis for the developmental behavior of soybean (*Glycine max* L. Merr.). *Theor. Appl. Genet.* 112, 665–673. doi: 10.1007/s00122-005-0169-y

Takai, T., Fukuta, Y., Shiraiwa, T., and Horie, T. (2005). Time-related mapping of quantitative trait loci controlling grain-filling in rice (*Oryza sativa* L.). *J. Exp. Bot.* 56, 2107–2118. doi: 10.1093/jxb/eri209

Thomas, H., and Ougham, H. (2014).The stay-green trait. *J. Exp. Bot.* 65, 3889–3900. doi: 10.1093/jxb/eru037

Toker, C. (2004). Estimates of broad-sense heritability for seed yield and yield criteria in faba bean (*Vicia faba* L.). *Hereditas* 140, 222–225. doi: 10.1111/j.1601-5223.2004.01780.x

Trenberth, K. E. (2011). Changes in precipitation with climate change research. *Clim. Res.* 47, 123–138. doi: 10.3354/cr00953

Ungerer, M. C., Halldorsdottir, S. S., Purugganan, M. D., and Mackay, T. F. C.(2003). Genotype-environment interactions at quantitative trait loci affecting inflorescence development in *Arabidopsis thaliana*. *Genetics* 165, 353–365.

Verma, V., Foulkes, M. J., Worland, A. J., Sylvester-Bradley, R., Caligari, P. D. S., and Snape, J. W. (2004). Mapping quantitative trait loci for flag leaf senescence as a yield determinant in winter wheat under optimal and drought-stressed environments. *Euphytica* 135, 255–263. doi: 10.1023/B:EUPH.0000013255.31618.14

Wade, M. J. (2002). A gene's eye view of epistasis, selection and speciation. *J. Evol. Biol.* 15, 337–346. doi: 10.1046/j.1420-9101.2002.00413.x

Wang, A., Li, Y., and Zhang, C. (2012). QTL mapping for stay-green in maize (*Zea mays*).Can. *J. Plant Sci.* 92, 249–256. doi: 10.4141/cjps2011-108 doi: 10.4141/cjps2011-108

Wang, D., Zhu, J., Li, Z., and Paterson, A. H. (1999). Mapping QTLs with epistatic effects and QTL×environment interactions by mixed linear model approaches. *Theor. Appl. Genet.* 99, 1255–1264. doi: 10.1007/s001220051331

Wang, Z., Wu, X., Ren, Q., Chang, X., Li, R., and Jing, R. (2010). QTL mapping for developmental behavior of plant height in wheat (*Triticum aestivum* L.). *Euphytica* 174, 447–458. doi: 10.1007/s10681-010-0166-3

Wu, W., Li, W., Tang, D., Lu, H., and Worland, A. J. (1999). Time-related mapping of quantitative trait loci underlying tiller number in rice. *Genetics* 151, 297–303.

Wu, X., Wang, Z., Chang, X., and Jing, R. (2010). Genetic dissection of the developmental behaviours of plant height in wheat under diverse water regimes. *J. Exp. Bot.* 61, 2923–2937. doi: 10.1093/jxb/erq117

Yang, D., Jing, R., Chang, X., and Li, W. (2007). Quantitative trait loci mapping for chlorophyll fluorescence and associated traits in wheat (*Triticum aestivum*). *J. Integ. Plant Biol.* 49, 646–654. doi: 10.1111/j.1744-7909.2007.00443.x

Yang, D., Li, M., Cheng, H., Chai, S., and Li, W. (2014). Genetic analysis of accumulation and remobilization of water soluble carbohydrates regulated by drought in wheat RIL stem at grain-filling stage. *Chin. J. Appl. Ecol.* 25, 803–812.

Yang, D., Zhang, G., Li, X., Xing, H., Cheng, H., Ni, S., et al. (2012). Genetic characteristics associated with drought tolerance of plant height and thousand-grain mass of recombinant inbred lines of wheat. *Chin. J. Appl. Ecol.* 23, 1569–1576.

Yang, G., Xing, Y., Li, S., Ding, J., Yue, B., Deng, K., et al. (2006). Molecular dissection of developmental behavior of tiller number and plant height and their relationship in rice (*Oryza sativa* L.). *Hereditas* 143, 236–245. doi: 10.1111/j.2006.0018-0661.01959.x

Yang, J., Hu, C., Hu, H., Yu, R., Xia, Z., Ye, X., et al. (2008). QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. *Bioinform.* 24, 721–723. doi: 10.1093/bioinformatics/btm494

Ye, Y., Li, M., Liu, Y., Chen, J., Yang, D., Hu, L., et al. (2015). QTL mapping and QTL× environmental interactions for plant height in wheat (*Triticum aestivum* L.). *Acta Agricul. Boreali Sin.* 30, 83–91.

Yoo, S., Cho, S., Zhang, H., Paik, H., Lee, C., Li, J., et al. (2007).Quantitative trait loci associated with functional stay-green SNU-SG1 in rice. *Mol. Cells* 24, 83–94.

Zhang, K., Fang, Z., Liang, Y., and Tian, J. (2009a). Genetic dissection of chlorophyll content at different growth stages in common wheat. *J. Genet.* 88, 183–189. doi: 10.1007/s12041-009-0026-x

Zhang, K., Zhang, Y., Chen, G., and Tian, J. (2009b).Genetic analysis of grain yield and leaf chlorophyll content in common wheat. *Cereal Res. Commun.* 37, 499–511. doi: 10.1556/CRC.37.2009.4.3

Zhang, Z., Xu, P., Jia, J., and Zhou, R. (2010). Quantitative trait loci for leaf chlorophyll fluorescence traits in wheat. *Aus. J. Crop Sci.* 4, 571–579.

Zhu, J. (1995). Analysis of conditional genetic effects and variance components in developmental genetics. *Genetics* 141, 1633–1639.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

VaERD15, a Transcription Factor Gene Associated with Cold-Tolerance in Chinese Wild *Vitis amurensis*

Dongdong Yu[1,2,3†], Lihua Zhang[1,2,3†], Kai Zhao[1,2,3], Ruxuan Niu[1,2,3], Huan Zhai[1,2,3] and Jianxia Zhang[1,2,3]*

[1] College of Horticulture, Northwest A&F University, Yangling, China, [2] Key Laboratory of Horticultural Plant Biology and Germplasm Innovation in Northwest China, Ministry of Agriculture, Yangling, China, [3] State Key Laboratory of Crop Stress Biology in Arid Areas, Northwest A&F University, Yangling, China

Early responsive to dehydration (*ERD*) genes can be rapidly induced to counteract abiotic stresses, such as drought, low temperatures or high salinities. Here, we report on an *ERD* gene (*VaERD15*) related to cold tolerance from Chinese wild *Vitis amurensis* accession 'Heilongjiang seedling'. The full-length *VaERD15* cDNA is 685 bp, including a 66 bp 5′-untranslated region (UTR), a 196 bp 3′-UTR region and a 423 bp open reading frame encoding 140 amino acids. The VaERD15 protein shares a high amino acid sequence similarity with ERD15 of *Arabidopsis thaliana*. In our study, *VaERD15* was shown to have a nucleic localization function and a transcriptional activation function. Semi-quantitative PCR and Western blot analyses showed that *VaERD15* was constitutively expressed in young leaves, stems and roots of *V. amurensis* accession 'Heilongjiang seedling' plants, and expression levels increased after low-temperature treatment. We also generated a transgenic *Arabidopsis* Col-0 line that over-expressed *VaERD15* and carried out a cold-treatment assay. Real-time quantitative PCR (qRT-PCR) and Western blot analyses showed that as the duration of cold treatment increased, the expression of both gene and protein levels increased continuously in the transgenic plants, while almost no expression was detected in the wild type *Arabidopsis*. Moreover, the plants that over-expressed *VaERD15* showed higher cold tolerance and accumulation of proline, soluble sugars, proteins, malondialdehyde and three antioxidases (superoxide dismutase, peroxidase, and catalase). Lower levels of relative ion leakage also occurred under cold stress. Taken together, our results indicate that the transcription factor *VaERD15* was induced by cold stress and was able to enhance cold tolerance.

Keywords: grapevine, *Vitis amurensis*, *VaERD15*, cold tolerance, functional analysis

Abbreviations: *AtGAPDH, Arabidopsis thaliana* glyceraldehyde 3-phosphate dehydrogenase; CAT, catalase; *CBF*, C-repeat-binding factor; cDNA, complementary DNA; *DREB1*, dehydration responsive element-binding factor 1; *ERD,* early responsive to dehydration; EST, expressed sequence tag; GFP, green fluorescent protein; GUS, β-glucuronidase; MDA, malondialdehyde; MS, Murashige and Skoog; NCBI, National Center for Biotechnology Information; ORF, open reading frame; PCR, polymerase chain reaction; POD, peroxidase; qRT-PCR, real-time quantitative polymerase chain reaction; ROS, reactive oxygen species; SOD, superoxide dismutase; WT, wild type.

INTRODUCTION

Grapevine (*Vitis vinifera* L.) is one of the most important multi-use fruiting plants of the world, and its berries are used in wine-making, or eaten fresh or dried. In addition to their value as a food (providing sugars, roughage etc.), grapes and their processed products have also been shown to have nutraceutical benefits including those attributed to resveratrol, which has a role in preventing and treating cardiovascular disease and cancer (Jeandet et al., 1995; Yang et al., 2009). These beneficial properties have fostered the recent growth in the grape industry.

The majority of commercial grape cultivars belong to the European grape (*V. vinifera* L.). While these cultivars have excellent organoleptic qualities, they suffer relatively poor tolerance to the cold experienced during winter, resulting in significant damage to grapevines growing in the cooler regions of the world, including northern America and northern China.

China has abundant resources of wild grape germplasm. These include *V. amurensis*, a species that can tolerate very low winter temperatures approaching −32°C. Therefore, this species has great potential as a germplasm resource for cold-resistant breeding (He and Niu, 1989). Study of the cold-tolerance genes in *V. amurensis* has significantly contributed to understanding the mechanisms of cold-tolerance, as well as transgenic breeding (Xu et al., 2014a,b).

In *Arabidopsis*, *ERD* genes can be induced within 1 h by drought stress (Kiyosue et al., 1994). Kiyosue et al. (1994) divided a total of 26 *ERD* cDNA clones into 16 different gene families based on their expression in response to drought. Early research on *ERDs* was concentrated on *Arabidopsis*, and many studies have shown that *ERD* genes can be induced by a diversity of stresses including drought (Rai et al., 2012), low temperature (Kiyosue et al., 1998), salinity (Rai et al., 2015) and abscisic acid (ABA) (Aalto et al., 2012). In recent years, *ERD* genes have been isolated from corn (Liu et al., 2009), soybeans (Alves et al., 2011a), tobacco (Shao et al., 2014), and tomato (Ziaf et al., 2016). Importantly, their roles in response to a range of stresses have been identified. However, of these *ERD*s, only *ERD6*, *ERD10,* and *ERD15* are related to cold tolerance in *Arabidopsis*. The first of these, *ERD6*, encodes a sugar carrier protein and is induced by low temperatures and moisture stress (Kiyosue et al., 1998). The gene *AtERD10* is induced by cold and regulates *CBF* transcription factors, and the transfer DNA (T-DNA) insertion silence strain of *AtERD10* reduces stress tolerance of mutants compared with wild type *Arabidopsis* plants (Kim and Nam, 2010). An earlier report indicates that, at low temperatures, *SpERD15* can protect the cell membrane, improve photosynthetic efficiency and promote the accumulation of soluble substances (Ziaf et al., 2011). However, another report (Kariola et al., 2006) showed that over-expression of *AtERD15* decreased the sensitivity to ABA, tolerance to drought, and also to low temperatures. Conversely, the *AtERD15* mutant was more sensitive to ABA and enhanced the tolerance to salinity and drought (Kariola et al., 2006). Therefore, in different plants, *ERD15* may have different functions. There have been no previous reports on the *ERD* genes in Chinese wild *Vitis* spp., and for this reason we are keen to explore the role of *ERD15* in *V. amurensis*.

We have already confirmed that *V. amurensis* is the most cold-resistant species of the 18 wild grapes species native to China and the seven wild grapes species native to North America (Zhang et al., 2012). Subsequently, a cold-induced cDNA library was constructed using potted plants of *V. amurensis* accession 'Heilongjiang seedling', in which one EST sequence encoding the ERD protein was obtained, and the gene was named as *VaERD15*. Real-time quantitative PCR (qRT-PCR) analyses revealed that *VaERD15* was induced by cold stress (Zhang et al., 2013). In this study, we clone the full-length *VaERD15*, and confirm its function in response to cold stress.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

Plants of Chinese wild *V. amurensis* accession 'Heilongjiang Seedling', which originates in northeast China, were maintained in the grape germplasm repository at the Northwest A & F University, Yangling, Shaanxi, the People's Republic of China. This accession is highly resistant to cold (Zhang et al., 2012).

In early January 2010, 1-year-old shoots were taken from mature vines for sand-storage at under 4°C. At the end of March, the shoots were retrieved and used for cuttings, soaked for 2 h in Transplantone (500 mg/L), allowed to develop roots, and cultivated in a greenhouse (25°C, light 12000 lux). In July, well-grown and healthy potted plants were selected for further cold stress and total RNA and protein isolation.

Isolation and Sequence Analysis of the *VaERD15* Gene

'Heilongjiang seedling' plants growing in pots were placed in a pre-chilled growth chamber at 4°C. Young leaves, stems and roots were harvested at 0, 2, 6, 12, 24, and 48 h after exposure to 4°C. Samples were frozen in liquid nitrogen prior to extraction of RNA and protein.

Total RNA was extracted using the improved sodium dodecyl sulfate (SDS)/phenol method (Zhang et al., 2003). The first-strand cDNA was synthesized using the Easy Script First-Strand cDNA Synthesis Super Mix (Transgen, China), according to the manufacturer's protocol. The cDNA templates for 5′- and 3′-Rapid Amplification of cDNA Ends (RACE) were synthesized using the SMARTTM RACE cDNA Amplification Kit (Clontech, Palo Alto, CA, USA). Primers for 5′- and 3′-RACE are listed in **Table 1**. All amplified RACE fragments were sequenced three times for each sample. Based on the 5′-RACE and 3′-RACE results, a pair of full-length primers *VaERD15*-F/R (including initiation and termination codons) were designed (**Table 1**). Amplification and sequencing of the full-length cDNA of *VaERD15* were repeated using two replicates.

The cDNA sequence of *VaERD15* gene from *V. amurensis* was translated into amino acid sequences on the NCBI website[1]. The protein conserved domain analysis website[2] was used to predict conserved domains, theoretical molecular weights and isoelectric

---

[1]http://www.ncbi.nlm.nih.gov/gorf/gorf.html
[2]http://prosite.expasy.org/

TABLE 1 | Primers used in this study.

Sr. no.	Primer	Sequence	Amplicon (bp)
1	*VaERD15*-3′RACE	5′-GTTCGTTCCGTTGGCATATCGGACGGT-3′	
	VaERD15-5′RACE	5′-CTCAAACTACCAGGAAACCCTAACTCGC-3′	
2	*VaERD15*-F(rt *VaERD15* -F)	5′-ATGGCTATGGAGGTAATTTCACGTAC -3′	423
	VaERD15-R(rt *VaERD15* -R)	5′- TTACCGCGGCTGCTGAATCG -3′.	
3	*VaERD15*- XbaII-F	5′-GCTCTAGAATGGCTATGGAGGTAATTTCACG-3′	439
	VaERD15-Kpn I -R	5′-GGGGTACCCCGCGGCTGCTGAATC-3′	
4	VaERD15-NcoI -F(Southern)	5′-CTAGccatgg ATGGCTATGGAGGTAATTTCACG-3′	442
	VaERD15-*BgIII* -R	5′-GGAagatetCTACGGCGGCTGCTGAATCG-3′	
5	*Actin1*-F	5′-GATTCTGGTGATGGTGTGAGT-3′	168
	Actin1-R	5′-GACAATTTCCCGTTC AGC AGT-3′	
6	*AtGAPDH*-F	5′-TTGGTGAC AAC AGGTC AAGC A-3′	
	AtGAPDH-R	5′-AAACTTGTCGCTCAATGCAATC-3′	
7	rt *AtERD15*-F	5′-AACTTCGACTTGGTACCCTGAT-3′	116
	rt *AtERD15* -R	5′-GAAGATCAGCTACATCGATATGA-3′	

points. The homologies of *ERD15* in *V. amurensis* and in corn, pepper and *Arabidopsis* were analyzed by means of DNAMAN analysis software. Finally, phylogenetic analyses were generated using MEGA 5.0 software[3].

Expression Pattern Analysis of *VaERD15*

Total RNA was extracted from various grapevines after exposure to low-temperature stress (4°C) for 0, 2, 6, 12, 24, and 48 h, and first-strand cDNA was synthesized, as previously described. Semi-quantitative PCR was carried out using the primers rt*VaERD15*-F and rt*VaERD15*-R (**Table 1**). The volume of the semi-quantitative PCR amplification was 20 µl, and *Actin1* (accession no. AY680701) was used as the internal reference gene. All reactions were repeated for three biological replicates.

Western blotting was carried out to further analyze the expression pattern of *VaERD15*. Total protein was extracted from various samples according to the method of Méchin et al. (2006). The extracted protein concentration was determined using the method of Bradford (1976). Proteins samples (25 µg) were prepared for SDS-polyacrylamide gel electrophoresis (PAGE), and blotted onto a polyvinylidene fluoride membrane (Roche, product no. 03010040001). Immune antibodies were prepared in our laboratory, as previously reported (Zhang et al., 2014).

Subcellular Localization of *VaERD15*

The ORF of *VaERD15* without a termination codon was obtained by PCR amplification using specific primers *VaERD15*-XbaI-F and *VaERD15*-KpnI-R (**Table 1**). The PCR products digested by XbaII and KpnI were cloned into a pMD18-T vector (TaKaRa, Japan). The fragments were fused into the N-terminus of the GFP expression vector driven by the 35S promoter. The vector carrying 35S::GFP was used as a control. The plasmids of the 35S::*VaERD15*-GFP fusion construct and 35S::GFP were purified for subsequent experiments. Plasmids were then transformed into onion epidermal cells using the particle bombardment method, as described by Varagona et al. (1992). Transformed

onion epidermal cells were cultured on MS media under dark conditions for 24 h at 25°C. Expression of the genes transformed into the onion epidermal cells was observed using confocal laser scanning microscopy (LSM 510 META, ZEISS, Germany).

Transcriptional Activation Assay

For the transcriptional activation assay, the ORF of *VaERD15* was generated and fused into the frame to the NcoI and BamHI sites of the GAL4 DNA-binding domain in the pGBKT7 vector by recombination reactions (Invitrogen, USA). The expression vector pGBKT7 carrying the *GAL4* gene constructed by our laboratory was used as a positive control and the empty pGBKT7 vector was used as a negative control. These constructs were then transformed into the yeast strain AH109. The resulting transformants were streaked onto Synthetic Defined (SD)/-Trp medium. After incubation at 30°C for 3 days, transformed strains on the SD/-Trp plates were selected and streaked onto SD/-Trp/-His/-Ade plates containing x-α-gal; the level of transcriptional activation was evaluated by color reaction.

Generation and Detection of Transgenic *Arabidopsis* Seedlings

The full-length cDNA of *VaERD15* was amplified by PCR and cloned into the BglII/NcoI site of pCAMBIA3301, generating pCAMBIA3301-*VaERD15*. The specific primers *VaERD15*-NcoI-F/R are shown in **Table 1**. Constructs were verified by sequencing. The constructed plasmid was introduced into *Agrobacterium tumefaciens* GV3101 cells by electroporation. Transgenic *Arabidopsis* plants were obtained using the floral dipping method (Clough and Bent, 1998). Putative transgenic *Arabidopsis* plants harboring the pCAMBIA3301-*VaERD15* construct were selected on MS plates containing 10 mg/L Basta.

Homozygous T3 *Arabidopsis* strains were tested by PCR and genomic DNA was extracted from 3-week-old leaves of the putative transgenic *Arabidopsis* seedlings using the cetyltrimethyl ammonium bromide (CTAB) method with appropriate modification (Wu et al., 1998). The specific primers *VaERD15*-NcoI –F/R were used for amplifying the exogenous

---

[3]http://www.megasoftware.net

gene, with pCAMBIA3301-*VaERD15* plasmid DNA as a positive control, and DNA from WT *Arabidopsis* (Col-0) leaves as a negative control. At the same time, leaves of *VaERD15*-positive *Arabidopsis* plants were taken for GUS staining, with WT *Arabidopsis* plants serving as controls.

Southern blot analysis was carried out as previously described Southern (1975). After plant genomic DNA was extracted and purified, 50 μg DNA was digested by the restriction enzyme HindIII in 40 μl for 12 h. Electrophoresis and Southern blot analysis were then carried out according to the standard methods.

Cold Tolerance, qRT-PCR and Western Blot Assay

Two T3 *Arabidopsis* lines (L1, L2) were used for the cold tolerance assay. Similarly, robust transgenic and WT 3-week-old *Arabidopsis* plants were used for the cold treatment. Plants were placed in a pre-chilled chamber at 4°C to cold acclimate for 48 h, and then transferred to a pre-chilled chamber at −6°C for 72 h. Plants suffering from cold stress were transferred to room temperature (approximately 23°C) for 5 days to recover. The phenotypic changes of *Arabidopsis* plants were observed and photographed during this period. Leaf samples were collected at 0, 2, 4, 8, 12, 24, 48, and 72 h after exposure to cold stress and stored at −80°C for qRT-PCR and Western blot analyses. qRT-PCR was carried out to determine expression changes of *VaERD15* and *AtERD15* [primers: rt*VaERD15*-F/R and rt*AtERD15*- F/R (**Table 1**)], *in vivo*, using the Takara SYBR Premix Ex Taq^{TM} II (Perfect Real Time) on a Bio-Rad IQ5 Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA). The volume used for qRT-PCR amplification was 20 μl, and *AtGAPDH* (accession no. 101214) was used as the endogenous reference gene. Western blotting was carried out as previously described. All experiments were carried out for three biological replicates.

Biochemical Indicator Assays of Transgenic *Arabidopsis*

Biochemical indices relating to cold stress were determined in 3-week-old *Arabidopsis* seedlings in both transgenic and control plants. All plants had been subjected to the cold treatment described above. Determinations in this part of the study were carried out for three biological replicates.

Relative electrolyte leakage was assessed according to the previously described method (Weigel et al., 2001). Proline content was determined following the method of Shan et al. (2007). MDA content was measured according to the method of Puckette et al. (2007) with minor modifications. About 200 mg leaves frozen in liquid nitrogen was homogenized in 4 ml 10% trichloracetic acid (TCA), then centrifuged at 10,000 rpm for 10 min. The supernatant (2 ml) was mixed with 2 ml thiobarbituric acid (TBA) and heated at 95°C for 30 min, quickly cooled on ice and then centrifuged at 10,000 rpm for 10 min. The soluble sugar and soluble protein contents were measured according to the methods of Bradford (1976) and Machado et al. (2013), respectively.

The activity of SOD was measured using the nitroblue tetrazolium (NBT) method (Giannopolitis and Ries, 1977; Puyang et al., 2015). POD activity was measured using the method of Pagariya et al. (2012). CAT activity was determined as described by Tseng et al. (2007).

Statistical Analysis

All physiological data was analyzed using the IBM SPSS Statistics 18.0 software. The differences between the transgenic samples and the corresponding wild type samples were calculated by the independent sample *t*-test. Significant differences were represented by $*P < 0.05$; $**P < 0.01$.

RESULTS

Sequence Analysis of *VaERD15*

Specific primers were designed according to the acquired ORF of *VaERD15*, which had a total length of 423 bp with a 5′-untranslated region (UTR) of 66 bp, a 3′-UTR of 196 bp and an intron size of 88 bp. *VaERD15* encoded a predicted polypeptide of 140 amino acids with a molecular weight of 16.2 kD (Zhang et al., 2014), and pI of 4.80. Alignment analysis with the grapevine genome (Jaillon et al., 2007) showed that *VaERD15* was initially located on chromosome 13.

Multiple sequence alignment analysis (**Figure 1A**) showed the *VaERD15* shared 36% homology with the *ZmERD15* predicted protein (accession no. ACG25626.1), 35% homology with the *CaERD15* predicted protein (accession no. ABB89735.1) and 33% homology with the *AtERD15* predicted protein (accession no. AAM64638.1). Phylogenetic analyses (**Figure 1B**) indicated that the relationship of *VaERD15* with other *ERDs* from similar plant species can be divided into two families. The closest relationship with *VaERD15* was found in *Vitis vinifera* and *Cucumis sativus*.

Expression Pattern Analysis of *VaERD15*

Semi-quantitative PCR and Western blotting were carried out to determine the expression patterns of *VaERD15*. The relative expressions at transcription level are shown in **Figure 2A**; the transcript was detected in all tissues measured. At 0 h, high *VaERD15* expression was found in stems, low expression in leaves and zero expression in roots. After 2 h of cold stress, the expression of *VaERD15* in stems decreased significantly, but recovered slightly after 12 h. The expression of *VaERD15* in leaves and roots showed a rising trend that peaked at 24 h, but then decreased.

The Western blot analysis results are shown in **Figure 2B**. We observed 16.2 kD bands on the polyvinylidene difluoride, membrane, indicating that the tissue proteins in roots, stems and leaves of *V. amurensis* accession 'Heilongjiang seedling' specifically reacted with *VaERD15* antibodies. *VaERD15* expression in leaves increased gradually and peaked at 48 h under the cold stress treatment. *VaERD15* expression in stems showed a downward trend up to 12 h, which recovered at 24 h, but subsequently decreased. In the roots, *VaERD15* expression levels increased, with a maximum at 6 h and a second small peak after 24 h.
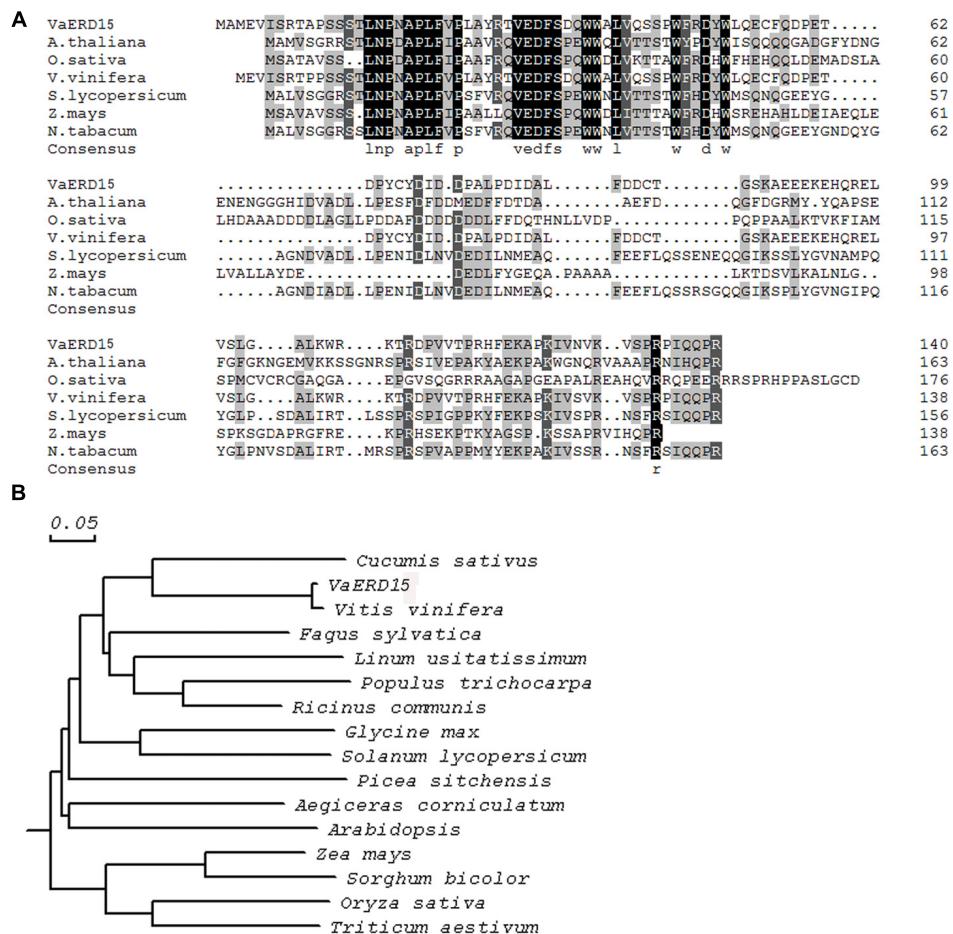
FIGURE 1 | (A) Amino acid sequence alignment of VaERD15 and other ERD15 from similar plant species. **(B)** Phylogenetic analysis of VaERD15 and other ERD15 from similar plant species. Sequences are from *Vitis amurensis* (*VaERD15*, JQ687321), *Arabidopsis thaliana* (*AtERD15*, AAM64638.1), *Brassica napus* (*BnERD15*, ADP37978.1), *Solanum lycopersicum* (*SlERD15*, NP_001234461), *Zea mays* (*ZmERD15*, ACG25626.1), and *Capsicum annuum* (*CaERD15*, ABB89735.1).

VaERD15 Functions as a Transcription Factor

The subcellular localization results showed that the *35S:VaERD15-GFP* fusion expression vector was transiently expressed in onion epidermal cells. Green fluorescence was only observed in the nuclei, while in the control green fluorescence was visible throughout the entire onion cell (**Figure 3**). This indicates that *VaERD15* is localized to the nuclei.

The AH109 strains with the recombinant plasmid of pGBKT7-*GAL4* (positive control) and pGBKT7-*VaERD15* or the empty pGBKT7 vector (negative control) were all able to grow well on the SD/-Trp medium (**Figure 4A**). This demonstrates that the pGBKT7-*VaERD15* recombinant plasmid, and the positive and negative controls were all transferred into the yeast. The strains transformed with *VaERD15* were able to grow well on the SD/-Trp/-His/-Ade + X-α-gal selective medium, and turned blue on the SD/-Trp/-His/-Ade + X-α-gal medium. Accordingly, the negative control did not grow on the SD/-Trp-His-Ade medium (**Figure 4B**), indicating that *VaERD15* could activate the expression of the reporter gene and synthesize the histidine and

adenine required for the normal growth of yeast AH109. Taken together, these results illustrate that *VaERD15* could function as a transcriptional activator in yeast.

Molecular Detection of Transgenic Plants

Arabidopsis plants over-expressing *VaERD15* were subjected to PCR. Five T3 *Arabidopsis* plants that survived on selection medium were used for the PCR experiments, and four expected bands were observed (**Figure 5A**). We also carried out a GUS staining assay. Results showed that transgenic *Arabidopsis* leaves were stained blue (**Figure 5B**), indicating that the *GUS* was expressed in the plant *in vivo*. This demonstrates that the plant genome had successfully integrated the target gene *VaERD15*.

Southern blot analysis was performed using four *VaERD15*-positive lines to further confirm that *VaERD15* had been integrated into the *Arabidopsis* genome. The Southern blot results showed that specific hybridization bands could be clearly observed in all positive lines (**Figure 5C**), but not in the WT plants. The four transgenic lines all had a specific hybridization signal, but the band sizes were not all the same. This dissimilarity
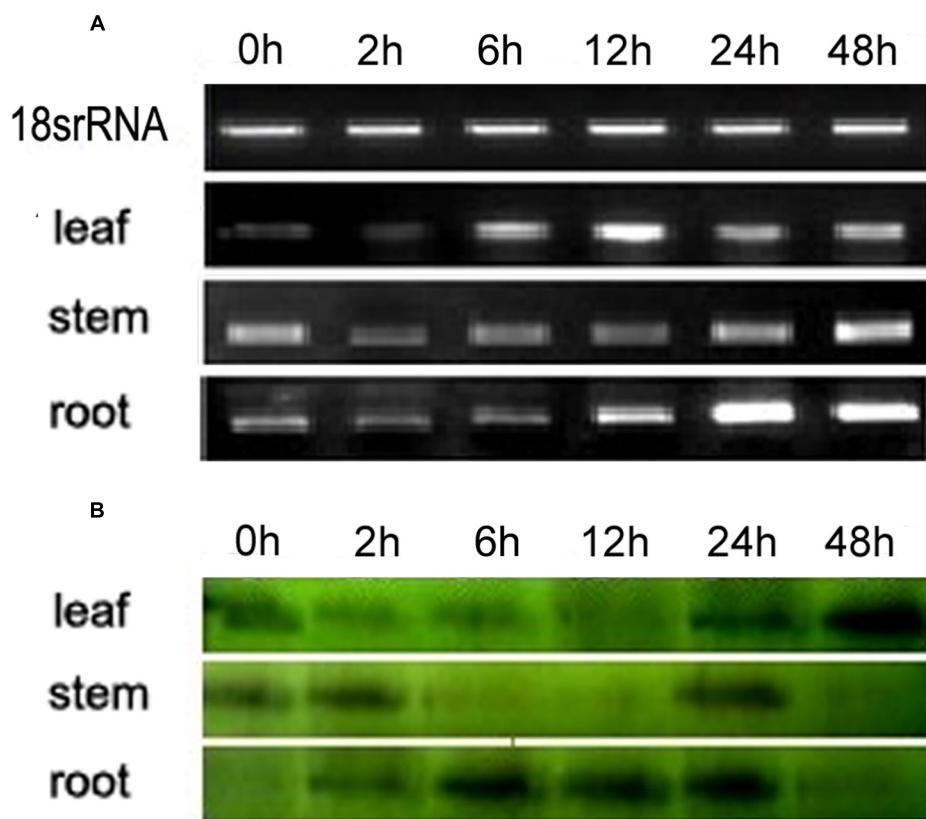
FIGURE 2 | Expression pattern analyses of *VaERD15* in different tissues of *Vitis amurensis* accession 'Heilongjiang seedling' after different periods of cold stress at 4°C. (A) Semi-quantitative PCR was used to analyze the transcriptional level expression of *VaERD15* with *Actin1* as a reference gene. **(B)** Western blotting was carried out to detect the expression levels of VaERD15 protein after cold stress. All experiments were carried out for three biological replicates.

FIGURE 3 | Subcellular localization of 35S::*VaERD15*-GFP fusion protein and control in onion epidermal cells. Cells were observed using fluorescent microscopy.
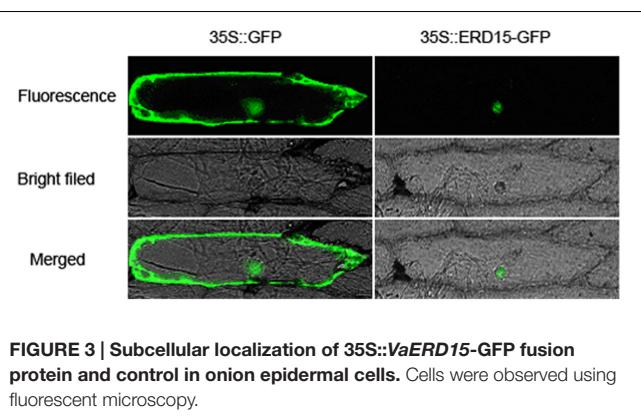
indicates that the transgenic *Arabidopsis* plants had different insert sites for *VaERD15* in their genomes.

Response of *VaERD15* to Cold Treatment

qRT-PCR and Western blotting were carried out to determine expression changes of *VaERD15 in vivo*. As shown in **Figure 6A**, *VaERD15* transcripts gradually increased in L1 and L2 lines with longer durations of cold stress, but the transcripts were

barely detectable in the WT plants. We also detected a change in endogenous *AtERD15* after transferring *VaERD15* into the *Arabidopsis* plant. The qRT-PCR results showed that there were no significant differences in the expression of endogenous *AtERD15* between wild type and transgenic *Arabidopsis*, except that the expression in transgenic plants was nearly four times higher than that in WT *Arabidopsis* at 12 h (**Supplementary Figure S1**). Western blot analysis showed that the *VaERD15* expression in two transgenic *Arabidopsis* lines both tended to increase with longer durations of cold stress, and the *VaERD15* expression level in L1 was higher than that in L2 (**Figure 6B**). However, only weak bands were observed in the WT plants, indicating that expression levels of *VaERD15* in WT and in transgenic *Arabidopsis* were different.

VaERD15 Enhanced Cold Tolerance in *Arabidopsis*

The cold tolerance assay showed transgenic and WT *Arabidopsis* were both subjected to freezing injury, but to different extents, with longer durations of cold stress. However, freezing injuries were more serious in WT *Arabidopsis*. Almost all the transgenic plants suffering from cold stress were able to resume normal growth after being transferred to room-temperature (∼23°C).

FIGURE 4 | Transcriptional activation assay of VaERD15 in yeast. (A) Fusion proteins of pGBKT7-*GAL4* (positive control), pGBKT7-*VaERD15* and pGBKT7 vector (negative control) were transformed into the yeast strain AH109. **(B)** Transformants were streaked onto plates of SD/Trp- and SD/Trp-/His-/Ade- + X-α-gal to detect their growth conditions and β-galactosidase activity.
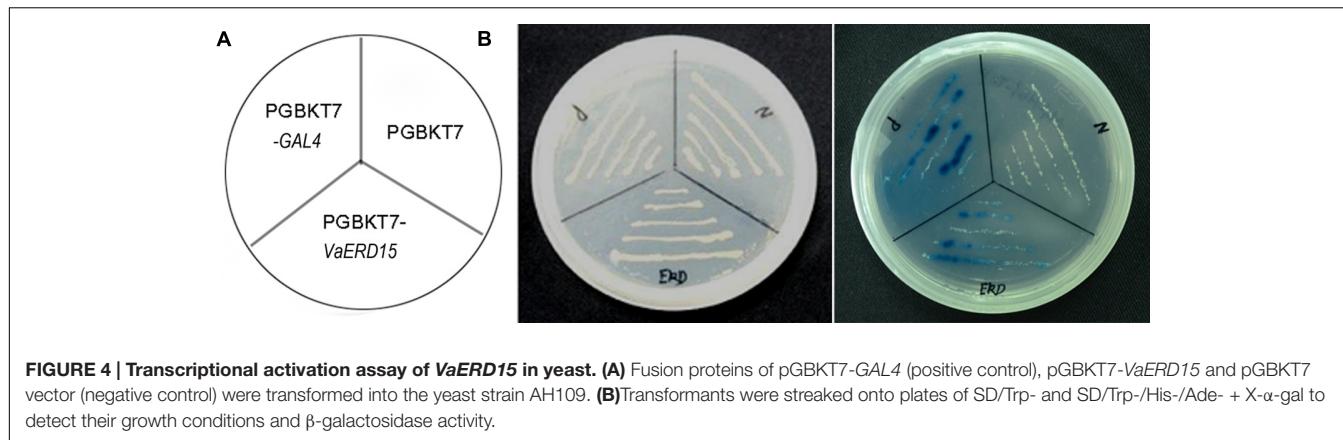
However, the same recovery did not occur with the WT plants (**Figure 7**).

Cold tolerance is strongly correlated with a number of physiological parameters in plants (Liu et al., 2010; Li et al., 2014; Xu et al., 2014a). To investigate whether the cold tolerance of *Arabidopsis* lines over-expressing *VaERD15* was improved, we measured several cold-related physiological indices including: relative electrolyte leakage; the contents of proline, MDA; soluble sugars and soluble proteins; and the activities of SOD, POD and CAT.

Results showed that relative electrolyte leakage in both transgenic and WT *Arabidopsis* plants showed an upward trend with longer durations of low temperature treatment (**Figure 8A**). Both transgenic and WT *Arabidopsis* plants showed similar conductivities in the early stages of cold stress but after 2 h, the conductivity in the WT plants was significantly higher than that in the transgenic plants. This indicates that WT plants suffered greater cell membrane damage than the transgenic plants.

We also investigated proline and MDA contents (**Figures 8B,C**). At the start of the stress period, the proline content of the transgenic plants was slightly higher than that of the WT plants but the difference was not significant. After 4 h of cold stress, proline in the transgenic plants accumulated rapidly, and proline abundance and growth rate were both significantly higher than in the WT plants. Changes in proline content were consistent with the phenotypic changes under cold stress. While the MDA content in the transgenic plants was lower than in the WT plants after 2 h of cold stress, it increased rapidly with longer durations of cold treatment and was consistently higher than in the WT plants.

Similar to conductivity, the soluble sugar contents of both transgenic and WT *Arabidopsis* plants increased with longer durations of cold treatment (**Figure 8D**). The soluble sugar content of the transgenic plants was slightly lower than in the WT plants for the first 12 h of cold stress, but then accumulated rapidly to become significantly higher than in the WT plants.

As shown in **Figure 8E**, the soluble protein content of the transgenic plants was consistently higher than that of the WT plants during cold stress. However, the pattern of change was very similar in both plant, falling to a minimum during the first 12 h, and then gradually increasing.

We also measured the activity of antioxidant enzymes. The results showed that SOD activity in transgenic *Arabidopsis* was consistently higher than in WT plants (**Figure 9A**), indicating that plants over-expressing *VaERD15* generally had elevated SOD activity. POD and CAT activities showed a late increase. The POD content curve of the transgenic plants was relatively flat, but that of the WT plants decreased over the first 8 h and then increased significantly after 12 h (**Figure 9B**). The CAT content in the transgenic plants was lower than in the WT plants at first, but rose slightly after 2 h (**Figure 9C**).

DISCUSSION

Grapes are of considerable economic importance and are grown over large areas of the world. However, close to the cooler limits of where this crop can be grown in the higher latitudes, both north and south of the equator, and at higher altitudes, chilling and frost damage can result in major economic loss. Therefore, the study of genes relating to cold tolerance is crucially important.

ERD genes were first isolated from *Arabidopsis* suffering from drought stress (Kiyosue et al., 1994). A large number of studies have shown that over-expressing *ERD* genes can improve the ability of plants to withstand biotic and abiotic stresses. AtERD10 and AtERD14 proteins have been reported to interact with phospholipid vesicles and protect membranes during conditions of high salinity, drought, and low temperature stress (Kovacs et al., 2008). *Brassica juncea ERD4* encoding a RNA-binding protein can respond to the induction of dehydration, ABA, salicylic acid, sodium chloride, cold and heat treatments, and overexpressing *BjERD4* can improve the tolerance of *Arabidopsis* plants to salt and dehydration stresses (Rai et al., 2015). Overexpressing *Arabidopsis ERD10* can activate *CBF/DREB1* genes and enhance the tolerance of plants to cold stress, while a T-DNA insertion mutant of *ERD10* is more sensitive than WT plants to cold stress (Kim and Nam, 2010). However, studies involving *ERD15* and responses to cold stress are few. In our study, a putative transcription factor *VaERD15*, was isolated from a cDNA library of *V. amurensis* induced by low temperatures, and its major function in cold tolerance was investigated.
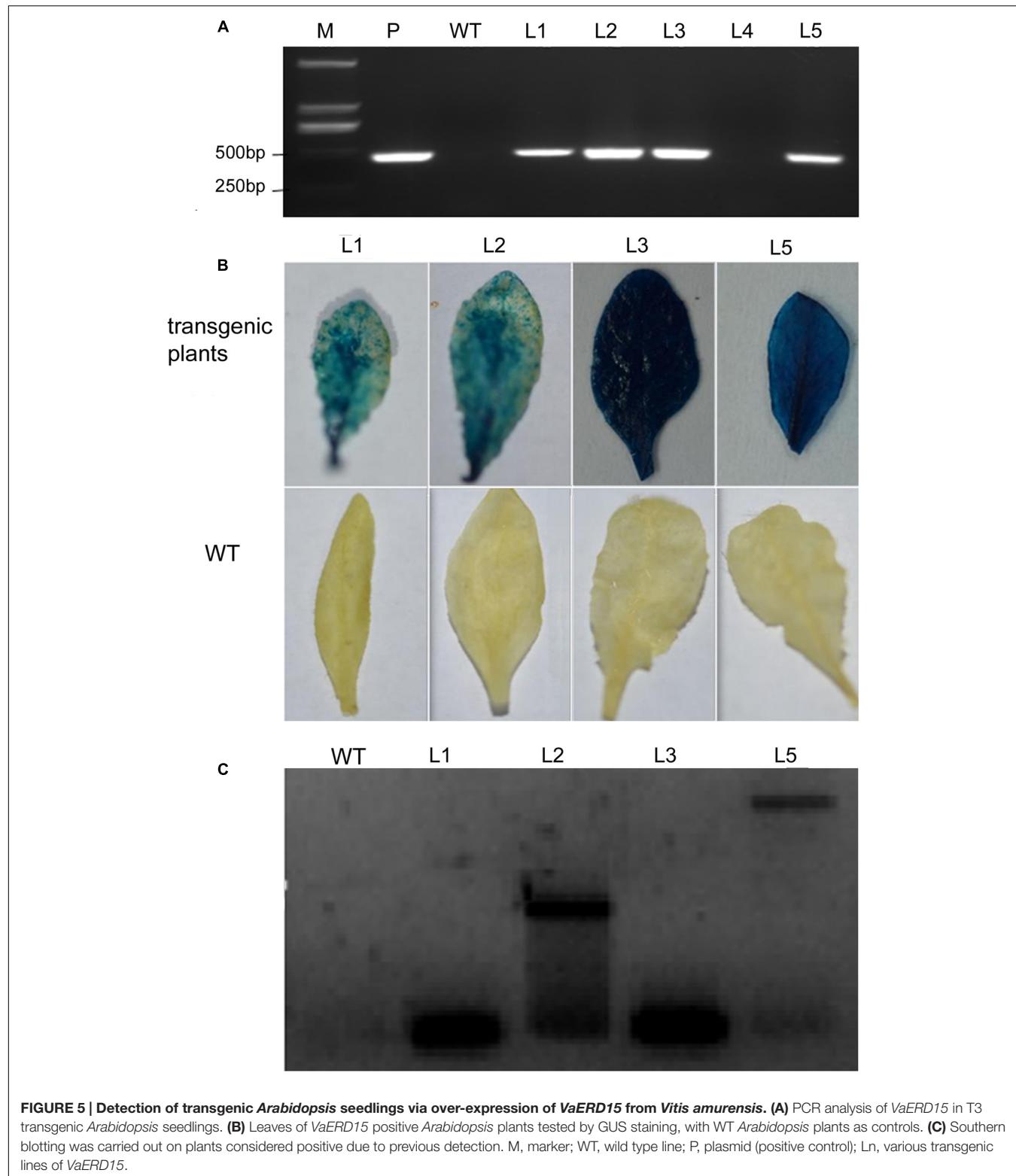
FIGURE 5 | Detection of transgenic *Arabidopsis* seedlings via over-expression of *VaERD15* from *Vitis amurensis*. (A) PCR analysis of *VaERD15* in T3 transgenic *Arabidopsis* seedlings. **(B)** Leaves of *VaERD15* positive *Arabidopsis* plants tested by GUS staining, with WT *Arabidopsis* plants as controls. **(C)** Southern blotting was carried out on plants considered positive due to previous detection. M, marker; WT, wild type line; P, plasmid (positive control); Ln, various transgenic lines of *VaERD15*.

Our results indicate that *VaERD15* is expressed in diverse plant tissues, which suggests that *VaERD15* is not specific to grapevine. Semi-quantitative PCR results show that accumulation of *VaERD15* transcripts in stems is significantly higher than in roots or leaves after 0 h of chilling. Similar research has shown that transcript accumulation of *SpERD15* under non-stress conditions is higher in roots and old leaves of tobacco (Ziaf et al., 2011). Increasing the duration of cold stress led to down-regulation in
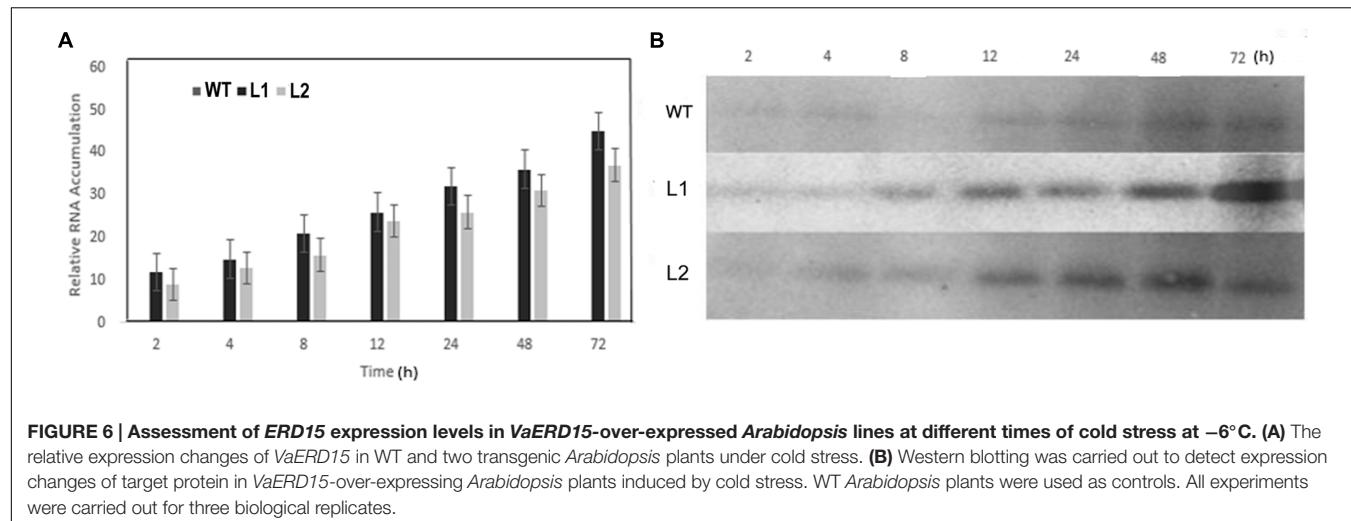
FIGURE 6 | Assessment of *ERD15* expression levels in *VaERD15*-over-expressed *Arabidopsis* lines at different times of cold stress at −6°C. (A) The relative expression changes of *VaERD15* in WT and two transgenic *Arabidopsis* plants under cold stress. **(B)** Western blotting was carried out to detect expression changes of target protein in *VaERD15*-over-expressing *Arabidopsis* plants induced by cold stress. WT *Arabidopsis* plants were used as controls. All experiments were carried out for three biological replicates.
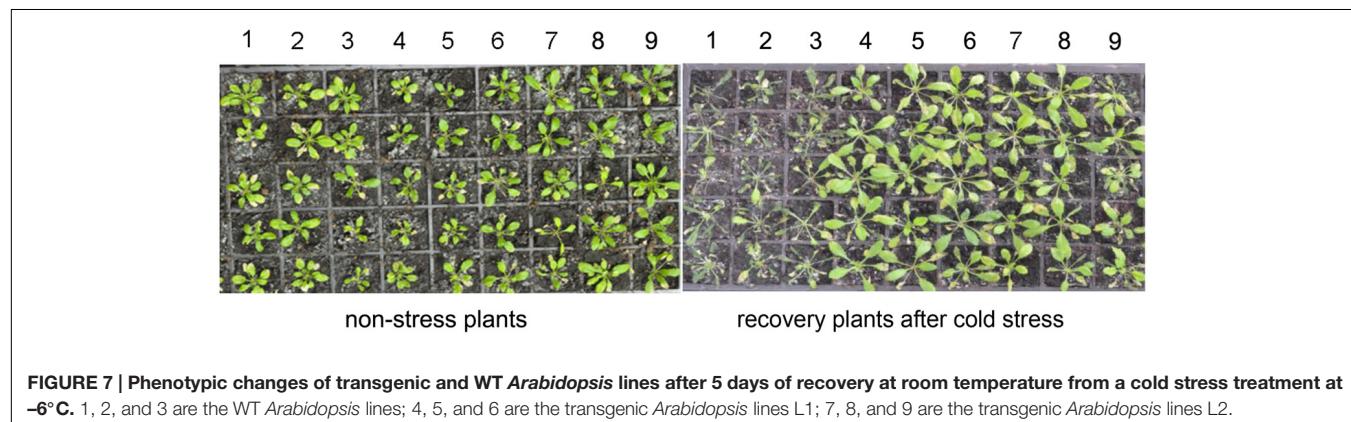
FIGURE 7 | Phenotypic changes of transgenic and WT *Arabidopsis* lines after 5 days of recovery at room temperature from a cold stress treatment at −6°C. 1, 2, and 3 are the WT *Arabidopsis* lines; 4, 5, and 6 are the transgenic *Arabidopsis* lines L1; 7, 8, and 9 are the transgenic *Arabidopsis* lines L2.
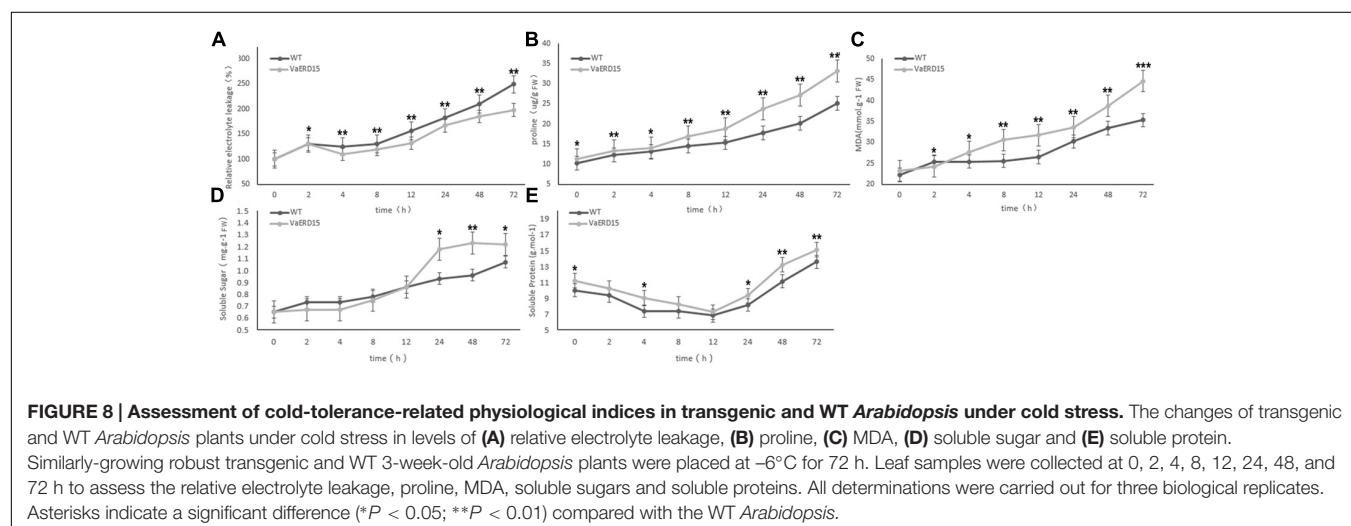
FIGURE 8 | Assessment of cold-tolerance-related physiological indices in transgenic and WT *Arabidopsis* under cold stress. The changes of transgenic and WT *Arabidopsis* plants under cold stress in levels of **(A)** relative electrolyte leakage, **(B)** proline, **(C)** MDA, **(D)** soluble sugar and **(E)** soluble protein. Similarly-growing robust transgenic and WT 3-week-old *Arabidopsis* plants were placed at −6°C for 72 h. Leaf samples were collected at 0, 2, 4, 8, 12, 24, 48, and 72 h to assess the relative electrolyte leakage, proline, MDA, soluble sugars and soluble proteins. All determinations were carried out for three biological replicates. Asterisks indicate a significant difference (*$P < 0.05$; **$P < 0.01$) compared with the WT *Arabidopsis*.

the expression of the desired gene in stems, with a concomitant increase in the expression of the gene in leaves and roots. As the expression patterns of *VaERD15* in leaves, roots and stems were different, we speculated that after low-temperature treatment of 'Heilongjiang seedling', *VaERD15* could be involved in different

regulatory pathways. The study by Yan et al. (2006) analyzed mRNA and protein level expression of 44 genes from *Oryza sativa*, and of the 27 up-regulated genes at the protein level, only five were up-regulated at a transcriptional level due to low-temperature treatment. Baginsky et al. (2005) also found
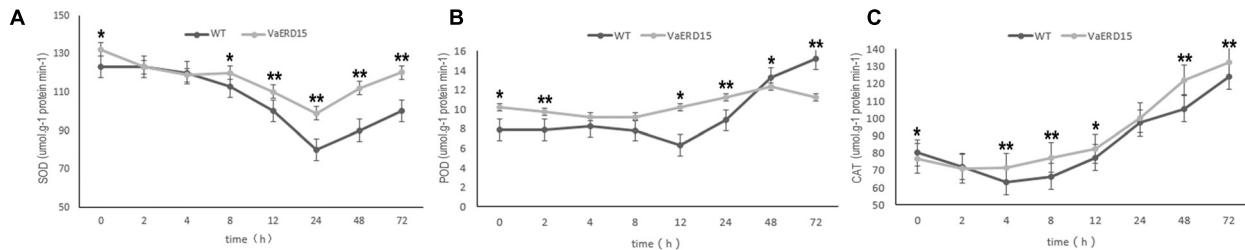
FIGURE 9 | Assessment of the activity of antioxidant enzymes in transgenic and WT *Arabidopsis* plants under cold stress. Changes in **(A)** SOD, **(B)** POD, **(C)** CAT activities in transgenic and WT *Arabidopsis* shoots after being subjected to cold stress. Similarly-growing robust transgenic and WT 3-week-old *Arabidopsis* plants were exposed to −6°C for 72 h. Leaf samples were collected at 0, 2, 4, 8, 12, 24, 48, and 72 h to determine the activities of SOD, POD and CAT. All determinations were carried out for three biological replicates. Asterisks indicate a significant difference (*$P < 0.05$; **$P < 0.01$) compared with the WT *Arabidopsis*.

that the correlation in expression at the mRNA and protein levels was lower when *Arabidopsis* chloroplasts and pollen were studied. This phenomenon, i.e., asynchrony in transcription and translation, can be caused by post-translational modifications in the protein-expression process, as well as by operator error (Baldi and Long, 2001). These previous reports support the idea that the expressions of *VaERD15* at the mRNA and protein levels are not identical.

Alves et al. (2011b) suggested that *GmERD15*, as a transcription factor localized to the nucleus and cytoplasm, can activate N-rich protein (*NPR*) gene expression under osmotic stress. *GmERD15* can specifically bind to a 187 bp fragment of *NPR-B* promoter in yeast, and activate the expression of downstream genes. Ziaf et al. (2011) reported that *SpERD15*, located mainly in the nucleus, could enhance the plant's ability to resist external stress by increasing the accumulation of intracellular solutes and by inhibiting lipid peroxidation. In this study, the analysis of subcellular localization in onion epidermis showed that *VaERD15* was mainly located in the nucleus. Transcription activation experiments confirmed that *VaERD15* was able to activate reporter genes in yeast. Therefore, these results show that *VaERD15* acts as a transcription factor.

To verify the function of *VaERD15* in mitigating cold stress, *Arabidopsis* plants over-expressing *VaERD15* were generated. A survival assay indicated that *Arabidopsis* plants over-expressing *VaERD15* survived better than WT plants under cold stress. We measured the expression levels of the *VaERD15* gene in transgenic *Arabidopsis* and WT plants and the results showed that the transcripts of *VaERD15* increased dramatically with longer durations of cold stress in transgenic *Arabidopsis* compared with WT plants. This observation is consistent with a number of other studies of cold-related genes in transformed plants (Park et al., 2010; Checker et al., 2012; Kidokoro et al., 2015). In addition, qRT-PCR results showed that the expression levels of endogenous *AtERD15* had a slight upward trend in WT plants with increasing duration of cold stress, while the expression of *AtERD15* showed a peak at 12 h in transgenic *Arabidopsis*, which indicated that the introduction of *VaERD15* could enhance the expression of endogenous *AtERD15*. These results illustrate that these transgenic plants are more tolerant to low temperatures. Further investigation indicated that transgenic

Arabidopsis over-expressing *VaERD15* also showed improved cold tolerance.

Physiological assessment showed cell concentrations of proline, MDA, soluble sugars and proteins, were higher in the over-expressing plants than in the WT plants, especially in the later stages of cold stress. These solutes act in different ways to mitigate stress, including protection of cellular structures, scavenging of ROS and detoxification of enzymes (Xiong et al., 2002; Verma and Dubey, 2003). It has been shown that increases in proline expression under drought, cold and salt stresses can help to protect plants from damage (Wanner and Junttila, 1999; Trovato et al., 2008; Xu et al., 2014b). In our study, rapid increases in proline occurred in transgenic plants after 4 h of cold stress, and after 72 h, proline was 1.5-times higher than in the WT plants. Many studies have reported similar results in plants over-expressing *VaICE1* or *VaICE2*, *OsCOIN*, *OsDREB1* or *DREB1*, *Osmyb4* under cold stress (Ito et al., 2006; Liu et al., 2007; Pasquali et al., 2008; Xu et al., 2014a). Proline accumulation in transgenic plants has been widely reported (Gao et al., 2011; Movahedi et al., 2015; Zhou et al., 2015). Moderate proline accumulations have been observed in transgenic tobacco plants over-expressing *SoMYB18* and a tendency for proline to decrease after cold stress has ceased has also been reported. These inconsistent results may be due to the observation that free proline is not a unique physiological index of osmotic potential reduction (Mao et al., 2011).

Electrolyte leakage is a key indicator of membrane injury caused by stress (Dexter et al., 1932; Jaglo-Ottosen et al., 1998). In the early stages, our *VaERD15*-over-expressing plants and control plants were both able to tolerate cold stress even though their conductivity levels increased. However, under more prolonged stress, the conductivity in the control plants was higher than that in the transgenic plants, indicating greater damage in the controls. The study by Parvanova et al. (2004) was consistent with our results and supported the idea that increased conductivity is indicative of membrane dysfunction.

Malondialdehyde is an end product of lipid peroxidation, and its level is therefore a key indicator of cold stress injury in plants (Jouve et al., 1993; Zhang and Kirkham, 1994). Previous studies on over-expression of *OsAPXa*, *SpERD15* in citrus, rice and tobacco plants reported that, compared with controls, lower

levels of MDA were found in transgenic plants after exposure to cold stress (Hara et al., 2003; Sato et al., 2011; Ziaf et al., 2011). In contrast with the above, our results showed that after 2 h of being subjected to cold temperatures, the MDA content of transgenic plants was consistently higher than that of WT plants. However, a similar finding to ours was reported for transgenic tobacco by Parvanova et al. (2004); their results suggested plants develop stress tolerance by producing large amounts of MDA.

Reactive oxygen species accumulation can lead to membrane peroxidation and thus destroy cell structure and function (Mittler et al., 2004). One way in which plants respond to stress is to accelerate free radical scavenging by increasing the activity of protective enzymes. The activities of three antioxidant-related enzymes were measured in our assay. In general, POD activity increased slowly, CAT activity increased rapidly while SOD activity initially increased, then decreased. The time courses of the activity trends for SOD and CAT in our transgenic plants were generally similar to those in WT plants, but the enzyme activities were significantly higher in the transgenic plants. Many reports on environmental stress have recorded the activities of antioxidant enzymes. For example, Yang et al. (2012) confirmed that the activities of POD, SOD and CAT increased in plants over-expressing *OsMYB2*, under salt stress. However, Yuan et al. (2015) reported that under cold stress, antioxidase activity was not significantly different in *VaPAT1*-over-expressing plants compared to WT plants. Shingote et al. (2015) found POD and CAT activities fell, while SOD activity increased in transgenic plants under cold stress. This indicates there may be synergy between the various antioxidant enzymes in the presence of active oxygen scavenging. Therefore, we hypothesize that within 24 h of the imposition of cold stress, the continuing decline in SOD activity that we observed was the result of over-accumulation of CAT in the cell.

We found that in *V. amurensis* the transcription factor *VaERD15* can significantly improve the tolerance of plants to low temperatures. Previous reports have shown that over-expression of individual genes can improve tolerance to cold stress (Saijo et al., 2000; Mukhopadhyay et al., 2004; Mishra et al., 2013). Other studies have also indicated that transformation of cold-resistant genes can activate and enhance the expression of related genes under cold stress (Ziaf et al., 2011; Xu et al., 2014a). Therefore, clarifying the interaction between *VaERD15* and other genes appears to be a promising direction for future research in this field. In summary, our findings confirm the significant value of continued investigation into the function and mechanisms of *ERD* genes in grapes, for the further development of cold-tolerant strains.

AUTHOR CONTRIBUTIONS

DY: Expression and analysis of *VaERD15* in *Arabidopsis*, determining the physiological and biochemical indices for transgenic plants. LZ: RT-PCR analysis of *AtERD15*, data collation and manuscript writing. KZ: Subcellular localization of VaERD15 and transcriptional activation analysis of *VaERD15*. RN: Cloning and sequence analysis of *VaERD15* gene, as well as RT-PCR analysis for *V. amurensis*. HZ: Semi-quantitative RT-PCR analysis for *V. amurensis*. JZ: Experimental design, plant material preparation and manuscript modification.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2017.00297/full#supplementary-material

FIGURE S1 | The relative expression changes of *VaERD15* and *AtERD15* in transgenic and WT *Arabidopsis* plants under cold stress. All experiments were carried out for three biological replicates.

REFERENCES

Aalto, M. K., Helenius, E., Kariola, T., Pennanen, V., Heino, P., Hõrak, H., et al. (2012). ERD15—an attenuator of plant ABA responses and stomatal aperture. *Plant Sci.* 182, 19–28. doi: 10.1016/j.plantsci.2011.08.009

Alves, M. S., Fontes, E. P., and Fietto, L. G. (2011a). Early responsive to dehydration 15, a new transcription factor that integrates stress signaling pathways. *Plant Signal. Behav.* 6, 1993–1996. doi: 10.4161/psb.6.12.18268

Alves, M. S., Reis, P. A., Dadalto, S. P., Faria, J. A., Fontes, E. P., and Fietto, L. G. (2011b). A novel transcription factor, ERD15 (Early Responsive to Dehydration 15), connects endoplasmic reticulum stress with an osmotic stress-induced cell death signal. *J. Biol. Chem.* 286, 20020–20030. doi: 10.1074/jbc.M111.233494

Baginsky, S., Kleffmann, T., Von Zychlinski, A., and Gruissem, W. (2005). Analysis of shotgun proteomics and RNA profiling data from *Arabidopsis thaliana* chloroplasts. *J. Proteome Res.* 4, 637–640. doi: 10.1021/pr049764u

Baldi, P., and Long, A. D. (2001). A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes. *Bioinformatics* 17, 509–519. doi: 10.1093/bioinformatics/17.6.509

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3

Checker, V. G., Chhibbar, A. K., and Khurana, P. (2012). Stress-inducible expression of barley Hva1 gene in transgenic mulberry displays enhanced tolerance against drought, salinity and cold stress. *Transgenic Res.* 21, 939–957. doi: 10.1007/s11248-011-9577-8

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x

Dexter, S. T., Tottingham, W., and Graber, L. (1932). Investigations of the hardiness of plants by measurement of electrical conductivity. *Plant Physiol.* 7, 63–78.

Gao, J.-J., Zhang, Z., Peng, R.-H., Xiong, A.-S., Xu, J., Zhu, B., et al. (2011). Forced expression of Mdmyb10, a myb transcription factor gene from apple, enhances tolerance to osmotic stress in transgenic *Arabidopsis*. *Mol. Biol. Rep.* 38, 205–211. doi: 10.1007/s11033-010-0096-0

Giannopolitis, C. N., and Ries, S. K. (1977). Superoxide dismutases I. Occurrence in higher plants. *Plant Physiol.* 59, 309–314. doi: 10.1104/pp.59.2.309

Hara, M., Terashima, S., Fukaya, T., and Kuboi, T. (2003). Enhancement of cold tolerance and inhibition of lipid peroxidation by citrus dehydrin in transgenic tobacco. *Planta* 217, 290–298. doi: 10.1007/s00425-003-0986-7

He, P.-C., and Niu, L.-X. (1989). Studies on cold resistance of wild *Vitis* species. *Acta Hortic. Sin.* 16, 81–88.

Ito, Y., Katsura, K., Maruyama, K., Taji, T., Kobayashi, M., Seki, M., et al. (2006). Functional analysis of rice DREB1/CBF-type transcription factors involved in cold-responsive gene expression in transgenic rice. *Plant Cell Physiol.* 47, 141–153. doi: 10.1093/pcp/pci230

Jaglo-Ottosen, K. R., Gilmour, S. J., Zarka, D. G., Schabenberger, O., and Thomashow, M. F. (1998). *Arabidopsis* CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* 280, 104–106. doi: 10.1126/science.280.5360.104

Jaillon, O., Aury, J.-M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–467. doi: 10.1038/nature06148

Jeandet, P., Bessis, R., Maume, B. F., Meunier, P., Peyron, D., and Trollat, P. (1995). Effect of enological practices on the resveratrol isomer content of wine. *J. Agric. Food Chem.* 2, 316–319. doi: 10.1021/jf00050a010

Jouve, L., Engelmann, F., Noirot, M., and Charrier, A. (1993). Evaluation of biochemical markers (sugar, proline, malonedialdehyde and ethylene) for cold sensitivity in microcuttings of two coffee species. *Plant Sci.* 91, 109–116. doi: 10.1016/0168-9452(93)90194-5

Kariola, T., Brader, G., Helenius, E., Li, J., Heino, P., and Palva, E. T. (2006). Early responsive to dehydration 15, a negative regulator of abscisic acid responses in *Arabidopsis*. *Plant Physiol.* 142, 1559–1573. doi: 10.1104/pp.106.086223

Kidokoro, S., Watanabe, K., Ohori, T., Moriwaki, T., Maruyama, K., Mizoi, J., et al. (2015). Soybean DREB1/CBF-type transcription factors function in heat and drought as well as cold stress-responsive gene expression. *Plant J.* 81, 505–518. doi: 10.1111/tpj.12746

Kim, S. Y., and Nam, K. H. (2010). Physiological roles of ERD10 in abiotic stresses and seed germination of *Arabidopsis*. *Plant Cell Rep.* 29, 203–209. doi: 10.1007/s00299-009-0813-0

Kiyosue, T., Abe, H., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1998). ERD6, a cDNA clone for an early dehydration-induced gene of *Arabidopsis*, encodes a putative sugar transporter. *Biochim. Biophys. Acta* 1370, 187–191. doi: 10.1016/S0005-2736(98)00007-8

Kiyosue, T., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1994). Cloning of cDNAs for genes that are early-responsive to dehydration stress (ERDs) in *Arabidopsis thaliana* L.: identification of three ERDs as HSP cognate genes. *Plant Mol. Biol.* 25, 791–798. doi: 10.1007/BF00028874

Kovacs, D., Kalmar, E., Torok, Z., and Tompa, P. (2008). Chaperone activity of ERD10 and ERD14, two disordered stress-related proteins. *Plant Physiol.* 147, 381–390. doi: 10.1104/pp.108.118208

Li, J., Wang, L., Zhu, W., Wang, N., Xin, H., and Li, S. (2014). Characterization of two VvICE1 genes isolated from 'Muscat Hamburg' grapevine and their effect on the tolerance to abiotic stresses. *Sci. Hortic.* 165, 266–273. doi: 10.1016/j.scienta.2013.11.002

Liu, K., Wang, L., Xu, Y., Chen, N., Ma, Q., Li, F., et al. (2007). Overexpression of OsCOIN, a putative cold inducible zinc finger protein, increased tolerance to chilling, salt and drought, and enhanced proline level in rice. *Planta* 226, 1007–1016. doi: 10.1007/s00425-007-0548-5

Liu, L., Duan, L., Zhang, J., Zhang, Z., Mi, G., and Ren, H. (2010). Cucumber (*Cucumis sativus* L.) over-expressing cold-induced transcriptome regulator ICE1 exhibits changed morphological characters and enhances chilling tolerance. *Sci. Hortic.* 124, 29–33. doi: 10.1016/j.scienta.2009.11.018

Liu, Y., Li, H., Shi, Y., Song, Y., Wang, T., and Li, Y. (2009). A maize early responsive to dehydration gene, ZmERD4, provides enhanced drought and salt tolerance in *Arabidopsis*. *Plant Mol. Biol. Rep.* 27, 542–548. doi: 10.1007/s11105-009-0119-y

Machado, R. A., Ferrieri, A. P., Robert, C. A., Glauser, G., Kallenbach, M., Baldwin, I. T., et al. (2013). Leaf-herbivore attack reduces carbon reserves and regrowth from the roots via jasmonate and auxin signaling. *New Phytol.* 200, 1234–1236. doi: 10.1111/nph.12438

Mao, X., Jia, D., Li, A., Zhang, H., Tian, S., Zhang, X., et al. (2011). Transgenic expression of TaMYB2A confers enhanced tolerance to multiple abiotic stresses in *Arabidopsis*. *Funct. Integr. Genomics* 11, 445–465. doi: 10.1007/s10142-011-0218-3

Méchin, V., Damerval, C., and Zivy, M. (2006). Total protein extraction with TCA-acetone. *Methods Mol. Biol.* 355, 1–8. doi: 10.1385/1-59745-227-0:1

Mishra, M. K., Chaturvedi, P., Singh, R., Singh, G., Sharma, L. K., Pandey, V., et al. (2013). Overexpression of *WsSGTL1* gene of *Withania somnifera* enhances salt tolerance, heat tolerance and cold acclimation ability in transgenic *Arabidopsis* plants. *PLoS ONE* 8:e63064. doi: 10.1371/journal.pone.0063064

Mittler, R., Vanderauwera, S., Gollery, M., and Van Breusegem, F. (2004). Reactive oxygen gene network of plants. *Trends Plant Sci.* 9, 490–498. doi: 10.1016/j.tplants.2004.08.009

Movahedi, A., Zhang, J., Yin, T., and Zhuge, Q. (2015). Functional analysis of two orthologous NAC Genes, CarNAC3, and CarNAC6 from *Cicer arietinum*, involved in abiotic stresses in poplar. *Plant Mol. Biol. Rep.* 33, 1539–1551. doi: 10.1007/s11105-015-0855-0

Mukhopadhyay, A., Vij, S., and Tyagi, A. K. (2004). Overexpression of a zinc-finger protein gene from rice confers tolerance to cold, dehydration, and salt stress in transgenic tobacco. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6309–6314. doi: 10.1073/pnas.0401572101

Pagariya, M. C., Devarumath, R. M., and Kawar, P. G. (2012). Biochemical characterization and identification of differentially expressed candidate genes in salt stressed sugarcane. *Plant Sci.* 184, 1–13. doi: 10.1016/j.plantsci.2011.12.002

Park, M. R., Yun, K. Y., Mohanty, B., Herath, V., Xu, F., Wijaya, E., et al. (2010). Supra-optimal expression of the cold-regulated OsMyb4 transcription factor in transgenic rice changes the complexity of transcriptional network with major effects on stress tolerance and panicle development. *Plant Cell Environ.* 33, 2209–2230. doi: 10.1111/j.1365-3040.2010.02221.x

Parvanova, D., Ivanov, S., Konstantinova, T., Karanov, E., Atanassov, A., Tsvetkov, T., et al. (2004). Transgenic tobacco plants accumulating osmolytes show reduced oxidative damage under freezing stress. *Plant Physiol. Biochem.* 42, 57–63. doi: 10.1016/j.plaphy.2003.10.007

Pasquali, G., Biricolti, S., Locatelli, F., Baldoni, E., and Mattana, M. (2008). Osmyb4 expression improves adaptive responses to drought and cold stress in transgenic apples. *Plant Cell Rep.* 27, 1677–1686. doi: 10.1007/s00299-008-0587-9

Puckette, M. C., Weng, H., and Mahalingam, R. (2007). Physiological and biochemical responses to acute ozone-induced oxidative stress in *Medicago truncatula*. *Plant Physiol. Biochem.* 45, 70–79. doi: 10.1016/j.plaphy.2006.12.004

Puyang, X., An, M., Han, L., and Zhang, X. (2015). Protective effect of spermidine on salt stress induced oxidative damage in two Kentucky bluegrass (*Poa pratensis* L.) cultivars. *Ecotoxicol. Environ. Saf.* 117, 96–106. doi: 10.1016/j.ecoenv.2015.03.023

Rai, A., Suprasanna, P., D'souza, S. F., and Kumar, V. (2012). Membrane topology and predicted RNA-binding function of the 'early responsive to dehydration (ERD4)' plant protein. *PLoS ONE* 7:e32658. doi: 10.1371/journal.pone.0032658

Rai, A. N., Tamirisa, S., Rao, K., Kumar, V., and Suprasanna, P. (2015). *Brassica* RNA binding protein ERD4 is involved in conferring salt, drought tolerance and enhancing plant growth in *Arabidopsis*. *Plant Mol. Biol.* 90, 375–387. doi: 10.1007/s11103-015-0423-x

Saijo, Y., Hata, S., Kyozuka, J., Shimamoto, K., and Izui, K. (2000). Over-expression of a single Ca^{2+}-dependent protein kinase confers both cold and salt/drought tolerance on rice plants. *Plant J.* 23, 319–327. doi: 10.1046/j.1365-313x.2000.00787.x

Sato, Y., Masuta, Y., Saito, K., Murayama, S., and Ozawa, K. (2011). Enhanced chilling tolerance at the booting stage in rice by transgenic overexpression of the ascorbate peroxidase gene, OsAPXa. *Plant Cell Rep.* 30, 399–406. doi: 10.1007/s00299-010-0985-7

Shan, D. P., Huang, J. G., Yang, Y. T., Guo, Y. H., Wu, C. A., Yang, G. D., et al. (2007). Cotton GhDREB1 increases plant tolerance to low temperature and is negatively regatively regulated by gibberellic acid. *New Phytol.* 176, 70–81. doi: 10.1111/j.1469-8137.2007.02160.x

Shao, H., Chen, S., Zhang, K., Cao, Q., Zhou, H., Ma, Q., et al. (2014). Isolation and expression studies of the ERD15 gene involved in drought-stressed responses. *Genet. Mol. Res.* 13, 10852–10862. doi: 10.4238/2014.December.19.6

Shingote, P. R., Kawar, P. G., Pagariya, M. C., Kuhikar, R. S., Thorat, A. S., and Babu, K. (2015). SoMYB18, a sugarcane MYB transcription factor improves salt and dehydration tolerance in tobacco. *Acta Physiol. Plant.* 37, 1–12. doi: 10.1007/s11738-015-1961-1

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98, 503–517. doi: 10.1016/S0022-2836(75)80083-0

Trovato, M., Mattioli, R., and Costantino, P. (2008). Multiple roles of proline in plant stress tolerance and development. *Rend. Lincei* 19, 325–346. doi: 10.1007/s12210-008-0022-8

Tseng, M. J., Liu, C.-W., and Yiu, J.-C. (2007). Enhanced tolerance to sulfur dioxide and salt stress of transgenic Chinese cabbage plants expressing both superoxide dismutase and catalase in chloroplasts. *Plant Physiol. Biochem.* 45, 822–833. doi: 10.1016/j.plaphy.2007.07.011

Varagona, M. J., Schmidt, R. J., and Raikhel, N. V. (1992). Nuclear localization signal (s) required for nuclear targeting of the maize regulatory protein Opaque-2. *Plant Cell* 4, 1213–1227. doi: 10.1105/tpc.4.10.1213

Verma, S., and Dubey, R. (2003). Lead toxicity induces lipid peroxidation and alters the activities of antioxidant enzymes in growing rice plants. *Plant Sci.* 164, 645–655. doi: 10.1016/S0168-9452(03)00022-0

Wanner, L. A., and Junttila, O. (1999). Cold-induced freezing tolerance in *Arabidopsis. Plant Physiol.* 120, 391–400. doi: 10.1104/pp.120.2.391

Weigel, R. R., Bäuscher, C., Pfitzner, A. J., and Pfitzner, U. M. (2001). NIMIN-1, NIMIN-2 and NIMIN-3, members of a novel family of proteins from *Arabidopsis* that interact with NPR1/NIM1, a key regulator of systemic acquired resistance in plants. *Plant Mol. Biol.* 46, 143–160. doi: 10.1023/A:1010652620115

Wu, D., Lin, Q., and Li, G. (1998). Cetyltrimethyl ammonium bromide (CTAB) method for extracting *Rhizophora* apiculata's DNA and its random amplified polymorphic DNA (RAPD) reaction. *Chin. J. Biochem. Mol. Biol.* 15, 67–70.

Xiong, L., Schumaker, K. S., and Zhu, J.-K. (2002). Cell signaling during cold, drought, and salt stress. *Plant Cell* 14, S165–S183. doi: 10.1105/tpc.000596

Xu, W., Jiao, Y., Li, R., Zhang, N., Xiao, D., Ding, X., et al. (2014a). Chinese wild-growing *Vitis amurensis* ICE1 and ICE2 encode MYC-type bHLH transcription activators that regulate cold tolerance in *Arabidopsis. PLoS ONE* 9:e102303. doi: 10.1371/journal.pone.0102303

Xu, W., Li, R., Zhang, N., Ma, F., Jiao, Y., and Wang, Z. (2014b). Transcriptome profiling of *Vitis amurensis*, an extremely cold-tolerant Chinese wild *Vitis* species, reveals candidate genes and events that potentially connected to cold stress. *Plant Mol. Biol.* 86, 527–541. doi: 10.1007/s11103-014-0245-2

Yan, S.-P., Zhang, Q.-Y., Tang, Z.-C., Su, W.-A., and Sun, W.-N. (2006). Comparative proteomic analysis provides new insights into chilling stress responses in rice. *Mol. Cell. Proteom.* 5, 484–496. doi: 10.1074/mcp.M500251-MCP200

Yang, A., Dai, X., and Zhang, W.-H. (2012). A R2R3-type MYB gene, OsMYB2, is involved in salt, cold, and dehydration tolerance in rice. *J. Exp. Bot.* 63, 2541–2556. doi: 10.1093/jxb/err431

Yang, J., Timothy, E. M., and Liu, R. H. (2009). Phytochemical profiles and antioxidant activities of wine grapes. *Food Chem.* 116, 332–339. doi: 10.1016/j.foodchem.2009.02.021

Yuan, Y., Fang, L., Karungo, S. K., Zhang, L., Gao, Y., Li, S., et al. (2015). Overexpression of VaPAT1, a GRAS transcription factor from *Vitis amurensis*, confers abiotic stress tolerance in *Arabidopsis. Plant Cell Rep.* 35, 655–666. doi: 10.1007/s00299-015-1910-x

Zhang, J., and Kirkham, M. (1994). Drought-stress-induced changes in activities of superoxide dismutase, catalase, and peroxidase in wheat species. *Plant Cell Physiol.* 35, 785–791. doi: 10.1093/oxfordjournals.pcp.a078658

Zhang, J.-J., Wang, Y.-J., Wang, X.-P., Yang, K.-Q., and Yang, J.-X. (2003). An improved method for rapidly extracting total RNA from *Vitis. J. Fruit Sci.* 3:126.

Zhang, J., Wu, X., Niu, R., Liu, Y., Liu, N., Xu, W., et al. (2012). Cold-resistance evaluation in 25 wild grape species. *Vitis* 51, 153–160.

Zhang, J., Liu, N., Niu, R., Liu, Y., Zhai, H., Xu, W., et al. (2013). Construction of a cDNA library of the Chinese wild Vitis amurensis under cold stress and analysis of potential hardiness-related expressed sequence tags. *Genet. Mol. Res.* 12, 1182–1193. doi: 10.4238/2013.April.12.5

Zhang, J.-X., Zhai, H., Niu, R.-X., and Li, R.-M. (2014). Prokaryotic expression and polyclonal antibody preparation of China wild *Vitis amurensis. J. Northwest For. Univ.* 29, 100–105.

Zhou, M., Wang, C., Qi, L., Yang, X., Sun, Z., Tang, Y., et al. (2015). Ectopic expression of *Fagopyrum tataricum* FtMYB12 improves cold tolerance in *Arabidopsis thaliana. J. Plant Growth Regul.* 34, 362–371. doi: 10.1007/s00344-014-9472-7

Ziaf, K., Loukehaich, R., Gong, P., Liu, H., Han, Q., Wang, T., et al. (2011). A multiple stress-responsive gene ERD15 from *Solanum pennellii* confers stress tolerance in tobacco. *Plant Cell Physiol.* 52, 1055–1067. doi: 10.1093/pcp/pcr057

Ziaf, K., Munis, M. F. H., Samin, G., Zhang, X., Li, J., Zhang, J., et al. (2016). Characterization of ERD15 gene from cultivated tomato (*Solanum lycopersicum*). *Pak. J. Agric. Sci.* 53, 27–33. doi: 10.21162/PAKJAS/16.3695

Check for updates

Methodology for High-Throughput Field Phenotyping of Canopy Temperature Using Airborne Thermography

David M. Deery[1]*, Greg J. Rebetzke[1], Jose A. Jimenez-Berni[2], Richard A. James[1], Anthony G. Condon[1], William D. Bovill[1], Paul Hutchinson[2†], Jamie Scarrow[2], Robert Davy[3] and Robert T. Furbank[1,4]

[1] CSIRO Agriculture and Food, Canberra, ACT, Australia, [2] High Resolution Plant Phenomics Centre, Australian Plant Phenomics Facility, CSIRO Agriculture and Food, Canberra, ACT, Australia, [3] CSIRO Information Management and Technology, Canberra, ACT, Australia, [4] ARC Centre of Excellence for Translational Photosynthesis, Australian National University, Canberra, ACT, Australia

Lower canopy temperature (CT), resulting from increased stomatal conductance, has been associated with increased yield in wheat. Historically, CT has been measured with hand-held infrared thermometers. Using the hand-held CT method on large field trials is problematic, mostly because measurements are confounded by temporal weather changes during the time required to measure all plots. The hand-held CT method is laborious and yet the resulting heritability low, thereby reducing confidence in selection in large scale breeding endeavors. We have developed a reliable and scalable crop phenotyping method for assessing CT in large field experiments. The method involves airborne thermography from a manned helicopter using a radiometrically-calibrated thermal camera. Thermal image data is acquired from large experiments in the order of seconds, thereby enabling simultaneous measurement of CT on potentially 1000s of plots. Effects of temporal weather variation when phenotyping large experiments using hand-held infrared thermometers are therefore reduced. The method is designed for cost-effective and large-scale use by the non-technical user and includes custom-developed software for data processing to obtain CT data on a single-plot basis for analysis. Broad-sense heritability was routinely >0.50, and as high as 0.79, for airborne thermography CT measured near anthesis on a wheat experiment comprising 768 plots of size 2 × 6 m. Image analysis based on the frequency distribution of temperature pixels to remove the possible influence of background soil did not improve broad-sense heritability. Total image acquisition and processing time was *ca.* 25 min and required only one person (excluding the helicopter pilot). The results indicate the potential to phenotype CT on large populations in genetics studies or for selection within a plant breeding program.

Keywords: field experiments, wheat, thermal imaging, image analysis, data processing, pixel histogram analysis

1. INTRODUCTION

The gaseous exchange of water for carbon occurs at the stomata. From this exchange, plant surfaces, particularly leaves, are cooled by evaporation, and their temperatures typically decrease with increased evaporation. Stomatal closure and reduced transpiration manifest as a warmer canopy temperature (CT), while cooler CT is related to more open stomata and higher transpiration. Cooler CT has been associated with genetic gains in wheat yield, higher stomatal conductance, and maximum photosynthetic rates under non-water-limited conditions (Fischer et al., 1998). Similarly, cooler CT has been associated with increased grain yield in warm, irrigated conditions in Mexico (Reynolds et al., 1994; Amani et al., 1996; Ayeneh et al., 2002), and in a study comparing a selection of spring wheat cultivars from Australia and the International Maize and Wheat Improvement Center (CIMMYT) (Rattey et al., 2011). Similar findings were reported in water-limited environments, with cooler CT in wheat associated with increased yield (Blum et al., 1989; Rashid et al., 1999; Olivares-Villegas et al., 2007). When measured during grain-filling, cooler CT has been associated with increased rooting depth (Reynolds et al., 2007a), water use, and grain yield (Lopes and Reynolds, 2010). Conversely, warmer CT has been associated with conservative water use in different crops. In wheat, Pinter et al. (1990) reported that varieties with warmer CT in well-watered conditions had reduced stomatal conductance, used less water and were higher yielding when grown under water limitation.

Researchers have investigated the genetic basis underpinning CT in wheat using different populations. For example, Saint Pierre et al. (2010) studied five populations grown in three environments (water-limited, well-watered, and heat stress) and reported that gene effects were mostly additive with some dominance. Genetic mapping has revealed multiple quantitative trait loci for CT that are often pleiotropic with other important agronomic traits including yield and biomass (Pinto et al., 2010; Bennett et al., 2012; Mason et al., 2013; Rebetzke et al., 2013b). These studies generally report a strong association between cooler CT and yield, particularly when CT is measured during grain-filling. However, the polygenic control, together with the environmental sensitivity of stomatal conductance and CT (Rebetzke et al., 2013b), may reduce the heritability of the trait and hence the utility of CT for selection within a breeding program. Mason and Singh (2014) investigated CT as an indirect selection criterion for wheat under water limitation and heat stress environments. They concluded that the most useful application of CT within a breeding program would occur in the early generations, where yield testing is not performed and therefore indirect selection would be beneficial.

In the aforementioned studies, CT was measured with hand-held infrared thermometers. Use of hand-held instruments in large experiments is laborious, time-consuming and sensitive to weather fluctuations over short periods of time. Moreover, difficulties associated with maintaining a constant view angle and avoiding "contamination" from soil further complicate hand-held CT measurements. To address these issues, infrared thermography has been proposed as a method for CT

phenotyping, owing to the advent of relatively affordable thermal cameras and user-friendly software for image processing (Jones et al., 2009; Takai et al., 2010; Prashar et al., 2013; Prashar and Jones, 2014). Recent studies have used unmanned aerial vehicles (UAVs) for the acquisition of thermal images for quantifying water stress in various field crops including cotton (Sullivan et al., 2007) and perennials including olives, mandarins, oranges, and apples (Berni et al., 2009a,b; Zarco-Tejada et al., 2012; Gómez-Candón et al., 2016). Chapman et al. (2014) demonstrated the use of UAV for various phenotyping applications including CT in sugarcane using thermal imaging.

For successful deployment of CT phenotyping within breeding programs, a scalable, and reliable methodology must first be developed and validated. Such a methodology must enable acquisition of CT from a large number of plots in a short time period (in the order of seconds), to reduce variance associated with weather fluctuations. The method must be accurate and precise to enable reliable and confident discrimination between genotypes. Moreover, the method must enable fast data acquisition and timely data processing. It must also be routine in delivery and readily accessible.

In this paper, we evaluate such a method developed for assessing CT on large field experiments. The method involves (i) airborne thermography from manned helicopter using a radiometrically-calibrated thermal camera to acquire CT data for large experiments in the order of seconds, and then (ii) data processing within minutes. The aim of this paper is to demonstrate the repeatability, scalability, and operative nature of the airborne thermography method for potential use in plot-scale phenotyping within a genetics study or within a plant breeding program.

2. MATERIALS AND METHODS

2.1. Field Experiments

A field experiment containing contrasting wheat genotypes was grown in two successive years at the Managed Environment Facility (MEF) (Rebetzke et al., 2013a), located at Yanco (34.62°S, 146.43°E, elevation 164 m) in SE Australia. The soil at the Yanco MEF has been classified as chromosol and has a clay-loam texture (Isbell, 1996). The experiment was sown on 28th May in 2013 and 11th June in 2014 following canola or field pea break-crops and then managed with adequate nutrition and chemical controls as required for pest, weed, and leaf diseases. The experiments comprised 768 experimental plots, of size 2 × 6 m with 18 cm row spacing (orientated North-South), and included a range of germplasm that conformed to the following criteria described in Rebetzke et al. (2013a): contemporary high-yielding, elite germplasm with agronomically-acceptable flowering time and plant height, to minimize confounding variation in CT with canopy architecture.

Genotypes were sown into a partial-replicate design trial (average number of replicates was 1.4) at a sowing density of 200 seeds per m². As described in Rebetzke et al. (2013a), two irrigation treatments (384 experimental plots per treatment) were used to simulate appropriate target environments, namely: *Treatment 1*, where irrigation was supplied to achieve a water

limitation pattern close to the long-term climate median for the site and; *Treatment 2*, where irrigation was supplied to achieve the equivalent of a decile eight rainfall (wettest 20% of years) for the site. The mean grain yields for Treatment 1 and Treatment 2 were 2.2 and 2.0 t/ha in 2013 and 1.7 and 2.1 t/ha in 2014, respectively. For the majority of entries and for both treatments, the anthesis growth stage occurred between the 19th and 24th of September in 2013 and between the 24th of September and 2nd of October in 2014.

2.2. Hand-Held Thermography

Hand-held CT measurements were made by a single operator walking through the plots with an infrared thermometer (Mikron 1600, Mikron Infrared Instrument Co., Inc., Oakland, NJ, USA). To minimize capturing soil in the instrument's field-of-view, the infrared thermometer was held obliquely to each plot and scanned across the canopy at an angle of *ca.* 20° (above the horizontal) for *ca.* 4 s to derive an average CT-value for each plot (after Rebetzke et al., 2013b). Measurements were taken on the morning of 18th October 2013, between 11:00 and 11:30, (Treatment 2 only) and on the afternoon of 25th October 2013, between 13:50 and 14:20, (Treatment 2 only). The majority of entries in the experiment were in the grain-filling growth stage. Weather conditions on both days were sunny and clear, and winds light (≤20 km/h). Air temperature, relative humidity, and wind speed were recorded at a weather station located *ca.* 400 m from the experiment site. Weather conditions were only recorded prior to the commencement of hand-held CT measurements on the 18th October 2013, while on the 25th October 2013, weather conditions were recorded prior to the start of measurements and at the completion of measurements.

2.3. Airborne Thermography

Airborne thermal images were acquired using the system described below on the 24th October 2013 at 10:00, 11:30, 12:30, and 13:30 and on the 2nd October 2014 at 09:00, 10:00, 11:00, 12:00, 13:00, and 14:00. On the 24th October 2013, the majority of entries in the experiment were in the grain-filling growth stage. On the 2nd October 2014, the majority of entries in the experiment were at anthesis. Weather conditions were sunny and clear, and winds light (≤20 km/h) on all days. The image acquisition and processing pipeline is depicted in **Figures 1**, **2**, and the major steps are described below.

2.3.1. Image Acquisition

Thermal images were acquired using a thermal infrared camera (FLIR® SC645, FLIR Systems, Oregon, USA, for which the technical specifications are: ±2°C or ±2% of reading; < 0.05°C pixel sensitivity; 640 × 480 pixels; 0.7 kg without lens; 13.1 mm lens). The camera was mounted in a commercially-available helicopter cargo pod (R44 Helipod II Slim Line Top Loader, Simplex Aerospace, Oregon, USA) and fitted to a Robinson R44 Raven helicopter (**Figure 1**). Highly visible infra-red (IR) targets, made of black fabric and of size *ca.* 1 m², were systematically positioned throughout the field to identify the experiment from adjacent collocated experiments. The IR targets were initially used for flight navigation and later for spatial referencing in

post-processing of the thermal images. In contrast to other studies (e.g., Gómez-Candón et al., 2016), the IR targets were not used for temperature correction of the thermal images. Prior to acquiring thermal images, a GPS tracking line (an "AB line") for subsequent flights was recorded by flying *ca.* 10 m above ground level (AGL) directly along the middle of the intended flight line.

In order to capture the experiment in a single flight pass whilst maximizing image resolution and avoiding motion blur, images were typically acquired at heights of 60 to 90 m AGL and at a flight velocity of 25–35 knots (45–65 km/h). Using the camera described above, at 60 m AGL, an image swath 43.6 by 32.1 m was obtained with a pixel size 7 × 7 cm, which equated to 204 temperature pixels per m². At 90 m AGL, an image swath 65.3 by 48.1 m was obtained with a pixel size 10 × 10 cm, which equated to 100 temperature pixels per m².

Thermal images were recorded on a laptop computer with FLIR® ResearcherIR™ software which was also used to control the camera. This proprietary software is provided for camera control and comprises basic image analysis features. The laptop and camera were manually operated by the helicopter passenger. Immediately prior to acquiring data for a particular experiment, the passenger would manually apply the shutter-based non-uniformity correction (NUC) and focus the camera, thereby ensuring image sharpness and that the NUC was not automatically applied during the run. Whilst acquiring thermal images, the passenger checked the images for complete coverage of the experiment using the IR targets and, in this way, provided real time assessment of the images and feedback on the helicopter flight path to the pilot (**Figure 2A**).

This method enabled capture of multiple high quality single images with at least 30% frame overlap in the direction of travel. Image acquisition with this system took < 10 s for the experiment described above comprising 768 plots.

2.3.2. Image Processing

The thermal images were pre-processed with FLIR® ResearcherIR™ software using the basic image analysis and processing features provided. Pre-processing included trimming of the image stack, to exclude extraneous images, and conversion from the RAW file format to Matlab (MAT) file format. This processing took *ca.* 2 min and was independent of experiment size.

Experimental plots were segmented from each thermal image using custom software developed with Python 2.7 (Python Software Foundation, https://www.python.org/); alias "ChopIt". The ChopIt software works on a frame-by-frame basis extracting data from the raw imagery, whereby the user navigates through the image stack to ensure that each plot in the experiment has been sampled. A screenshot of the ChopIt graphical user interface is shown in **Figure 2B**. The ChopIt software is designed for semi-automated plot segmentation whereby the user controls the area sampled within plots by placement of bounding corners. The software also assigns a unique identifying number to each plot. The core geometric algorithm in ChopIt divides a four-sided region into a predefined number of rows and columns based on the placement of the bounding corners. The algorithm

FIGURE 1 | Airborne thermography image acquisition system comprising a helicopter cargo pod with thermal camera and acquisition kit mounted on the skid of a Robinson R44 Raven helicopter. Photo insert shows the inside of the helicopter cargo pod with arrow denoting FLIR® SC645 thermal camera: $\pm 2°$C or $\pm 2\%$ of reading; $< 0.05°$C pixel sensitivity; $640x480$ pixels; 0.7 kg without lens.
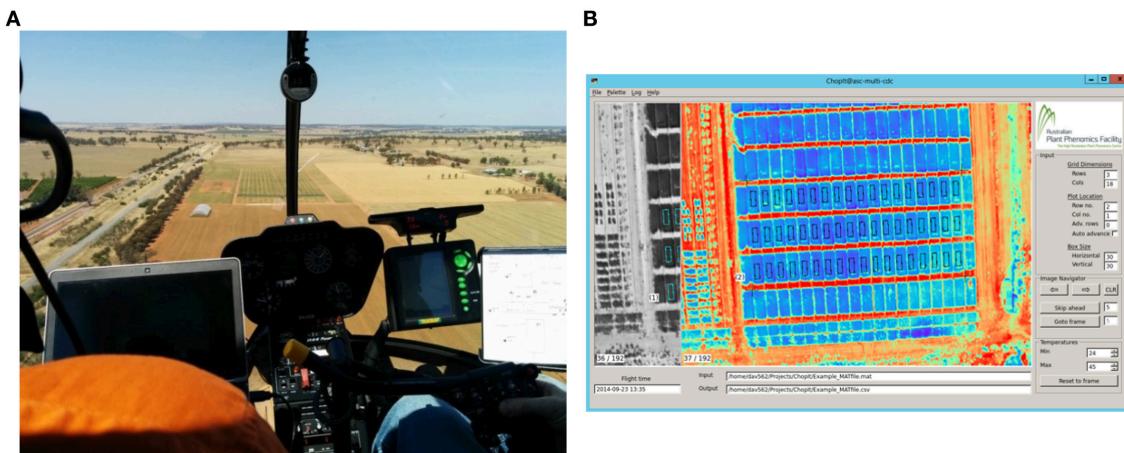
FIGURE 2 | Airborne thermography image acquisition and processing pipeline. Total time to acquire and process images for an experiment comprising 1000 plots of size 2 × 6 m is *ca.* 25 min. **(A)** Image acquisition with helicopter. The images are recorded on a laptop and the passenger, left, provides real time assessment of the images and feedback to the pilot. This step takes < 10 s for an experiment comprising 1000 plots of size 2 × 6 m. **(B)** Screenshot of custom developed software called ChopIt. ChopIt is used for plot segmentation and extraction of CT from each individual plot for statistical analysis. This step takes *ca.* 20 min for an experiment comprising 1000 plots of size 2 × 6 m.

uses the concept of vanishing points and thus can accommodate situations where the image plane is not parallel to the ground. For a given row and column value, a plot rectangle is defined with a surrounding buffer, and the CT data are extracted from within the plot rectangle. The ChopIt software produces two output files comprising the CT data for each plot rectangle assigned by the user: (1) SQLite database file comprising all the CT pixel values for each experimental plot rectangle; and (2) an Excel file comprising a descriptive statistical summary for each experimental plot rectangle.

The process of plot segmentation and extraction of CT for each individual plot for statistical analysis took *ca.* 20 min for the experiment described above comprising 768 plots. Total image acquisition and processing time was *ca.* 25 min.

2.3.3. Image Quality Control

The custom-developed ChopIt software provides a high level of quality control for the user to manually exclude poor quality sections of the plot or removed sections (e.g., where biomass cuts have been earlier sampled). This user-enabled flexibility in the image analysis protocol is demonstrated in **Figure 3**, where a section comprising a previous biomass sampling has been excluded on a plot with approximate dimensions of 2 × 6 m. In this fashion, sections of plots comprising biomass samples were excluded in the study reported herein. In addition to this feature of manually excluding poor quality sections of the plot during plot segmentation, post-processing of the temperature pixels is also possible, as all the pixel data for each plot are stored in a SQLite database file.
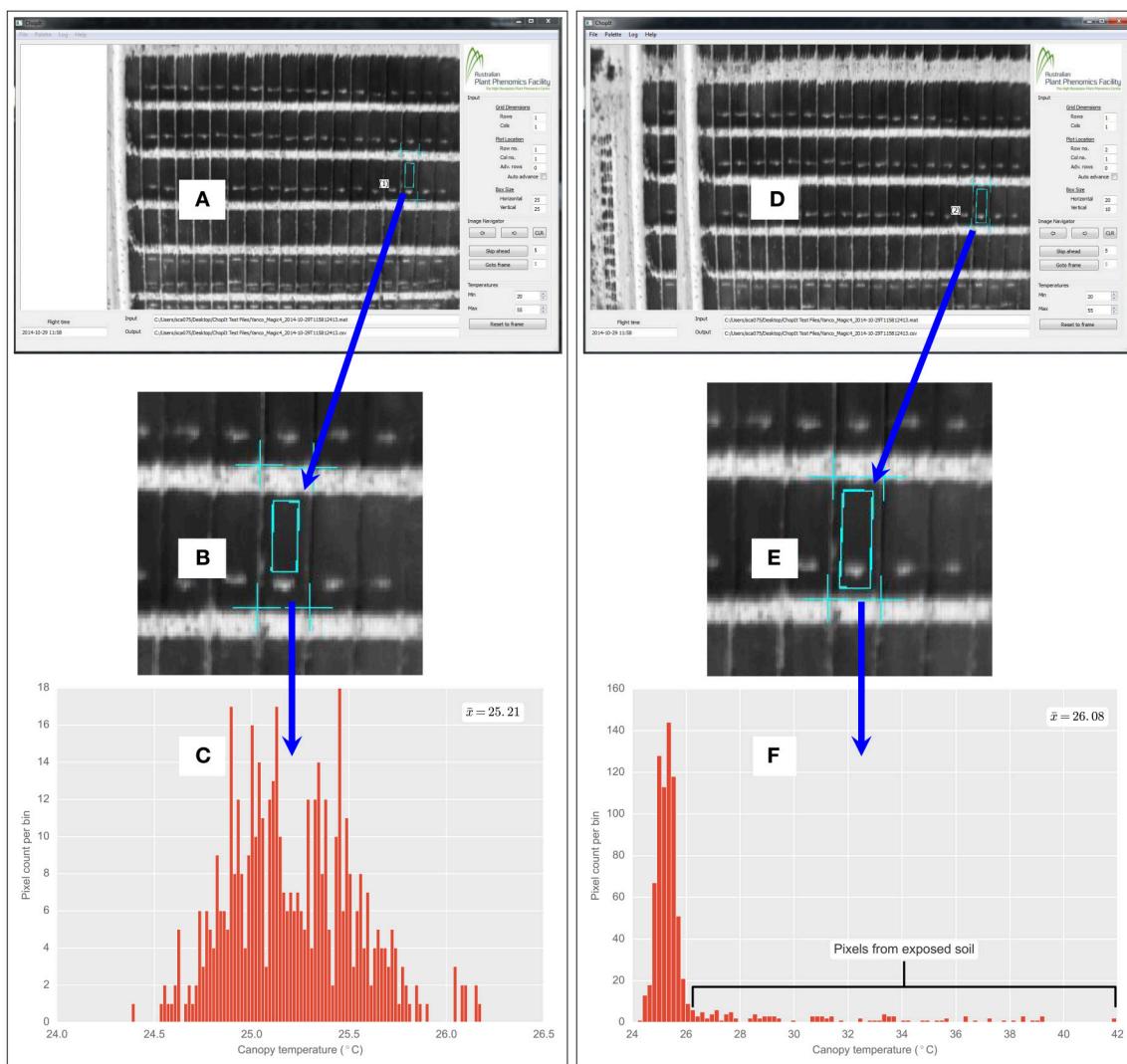
FIGURE 3 | Custom developed ChopIt airborne thermography image processing software. Demonstration of quality control by user and flexibility of image analysis protocol, whereby the user can manually exclude exposed soil patches within a field plot. In this example, exposed soil patch is from biomass sample taken earlier on a 2 × 6 m field plot of wheat. **(A)** ChopIt user interface, where user has avoided exposed soil within the plot. **(B)** Magnified view showing exposed soil patch. **(C)** CT histogram is therefore void of pixels from the exposed soil patch. Compare with **(D)**, where for the same plot as **(A)**, user has included the exposed soil, evident in magnified view **(E)** and pixels from the soil patch are evident in the CT histogram **(F)**. Where $\bar{x}$ denotes the respective mean for **(C,F)**. The $\bar{x}$ from **(C)** is 0.87°C cooler than **(F)**.

2.4. Analysis of the Pixel Frequency Distribution

2.4.1. Rationale

The water limitation imposed on the crop in the MEF can often result in incomplete crop ground cover. The incomplete ground cover may have implications for the airborne thermography measurements through the potential aggregation of crop canopy and the background soil temperatures, which in the case of dry soil is often warmer than the crop canopy. The potential for the background soil temperature to bias estimates of CT is exacerbated when the size of the image pixels is the same as, or greater than, the individual plant organs that comprise the crop canopy. In such cases, a pixel is likely to comprise both soil and plant canopy temperatures, thereby resulting in "mixed pixels". The presence of mixed pixels is likely to bias the observed temperature toward the soil background temperature (Jones and Sirault, 2014).

In the airborne thermography system described above (Section 2.3), at an above-ground altitude of *ca.* 60 m, the pixel size is *ca.* 7 × 7 cm. This pixel resolution is several times greater than the leaf width of a typical wheat plant (*ca.* 1 cm) and, together with variation in plant establishment and canopy architecture, can result in mixed pixels and the need for image analysis to remove temperature pixels arising from the background soil that can bias the intrinsic measures of plant-based CT. Methods for handling thermal

images containing mixed pixels were reviewed by Jones and Sirault (2014). These methods include automated thresholding algorithms such as the Otsu method (Otsu, 1979) and work best when discrete peaks are present in the histogram, representing multiple frequency distributions of temperature pixels.

A bimodal distribution with discrete peaks representing soil and plant canopy was not evident in the data acquired. Rather, pixel frequency distributions were unimodal with a long tail of warm temperature pixels, as shown in **Figures 3**, **4**. The unimodal distribution may have resulted from the *ca.* 7×7 cm pixel resolution, whereby no clear difference between the mean temperature of the background soil and the mean temperature of the plant canopy was evident (Jones and Sirault, 2014). To account for the unimodal pixel distribution, the below-described methods of pixel frequency analysis were used.

2.4.2. Methods for Analysis of the Pixel Frequency Distribution

The frequency distribution of the temperature pixels from a given plot rectangle produced from the ChopIt software was analyzed to determine if the observed temperature was biased by the background soil and whether this influenced the measurement repeatability. Three methods for analysis of pixel temperature bias were evaluated, depicted in **Figure 4**, and hereafter referred to as M1, M2, and M3:

1. M1: The mean of all pixels (no pixels discarded).
2. M2: The mean of the coolest 25th percentile remaining after discarding the coolest 5th percentile. This method was designed to extract the plant CT only. The percentile of pixels discarded and sampled were selected arbitrarily with the intention of maximizing the sampling of pixels arising from plant material and discarding mixed pixels and pixels comprising background soil. The coolest 5th percentile was discarded to avoid sampling background soil that might be cooler than the plant material (Jones, 2002; Jones and Sirault, 2014).
3. M3: A method designed to discard warm pixels arising from soil patches and calculated as follows:

   a. For a given set of plot pixel temperatures, x, the mode of the distribution was estimated.
   b. Then, a filter cut-off, c, was calculated according to: $c = min(x) + 2(mode(x) - min(x))$
   c. The set, x, was then filtered by retaining only values where $x < c$.
   d. The mean of this filtered set was then calculated.

From the representative example of pixel frequency distribution shown in **Figure 4**, the difference between M1 and M2 was $2.4°C$, the difference between M1 and M3 was $0.6°C$ and the difference between M3 and M2 was $1.7°C$.

2.5. Statistical Analysis

CT data were analyzed after first checking for normality and error variance homogeneity at each date by time sampling event.

Each event was analyzed separately with the best spatial models being determined after first fitting the experimental design and then modeling the residual variation with autoregressive row and column terms in the Genstat® statistical program (https://www.vsni.co.uk/software/genstat/). Significant spatial effects were identified and residuals assessed before determinations made to the need for fitting of other (e.g., linear) effects (Gilmour et al., 1997). Generalized heritabilities were then estimated after Holland et al. (2003).

3. RESULTS

3.1. Hand-Held Thermography

The results from the hand-held thermography are summarized in **Table 1**, and box-plots for each sample time are shown in **Figure 5A**. The range in plot CT was large in each sampling event. Broad-sense heritabilities for CT using the hand-held thermography method were 0.17 and 0.13 for the morning (18 October 2013) and afternoon measurements (25 October 2013), respectively (Treatment 2 only). The time taken to measure CT using hand-held thermography on 384 plots was *ca.* 30 min on both days. On the 25th October 2013, the air temperature changed from $17.8°C$ prior to the start of measurements to $19.0°C$ at the completion of measurements. At the same time, the relative humidity remained constant at 28%, and the wind speed remained constant at 17 km/h. Weather conditions were only recorded prior to the commencement of measurements on the 18th October 2013: air temperature was $11.9°C$, relative humidity was 48% and wind speed was 11 km/h.

3.2. Airborne Thermography

Box-plots summarizing the airborne thermography CT data for each flight time and irrigation treatment are shown for 2013 and 2014 in **Figures 5C,D**, respectively. Each box-plot represents CT data from 384 experimental plots (384 experimental plots per treatment and 768 experimental plots in total). The CT for box-plots shown in **Figures 5C,D** were derived from each experimental plot using M1, the mean of all pixels from a given plot rectangle produced from the ChopIt software with no pixels removed. For 2013 (**Figure 5C**) and 2014 (**Figure 5D**), Treatment 2 is consistently cooler than Treatment 1, owing to the greater water limitation in Treatment 1. Pearson correlations between the hand-held CT, obtained on the 18th October 2013 11:00 and the 25th October 2013 14:00, and the airborne thermography CT, obtained on the 24th October 2013, are shown in **Figure 5B**. The correlations between hand-held CT and airborne CT were <0.25.

Broad-sense heritabilities for airborne thermography CT for each flight time and irrigation treatment are shown for 2013 and 2014 in **Figure 5**. For the 2013 data, broad-sense heritability was calculated using CT derived from each experimental plot using M1 (**Figure 6A**). For the 2014 data, to test the influence of soil temperature bias on measurement repeatability, broad-sense heritability was calculated for CT estimated from M1, M2, and M3 (**Figure 6B**). With the exception of two early morning measurements on Treatment 1 in 2014, broad-sense heritability was high and ranged from 0.34 to 0.79. **Figure 6B** shows that for
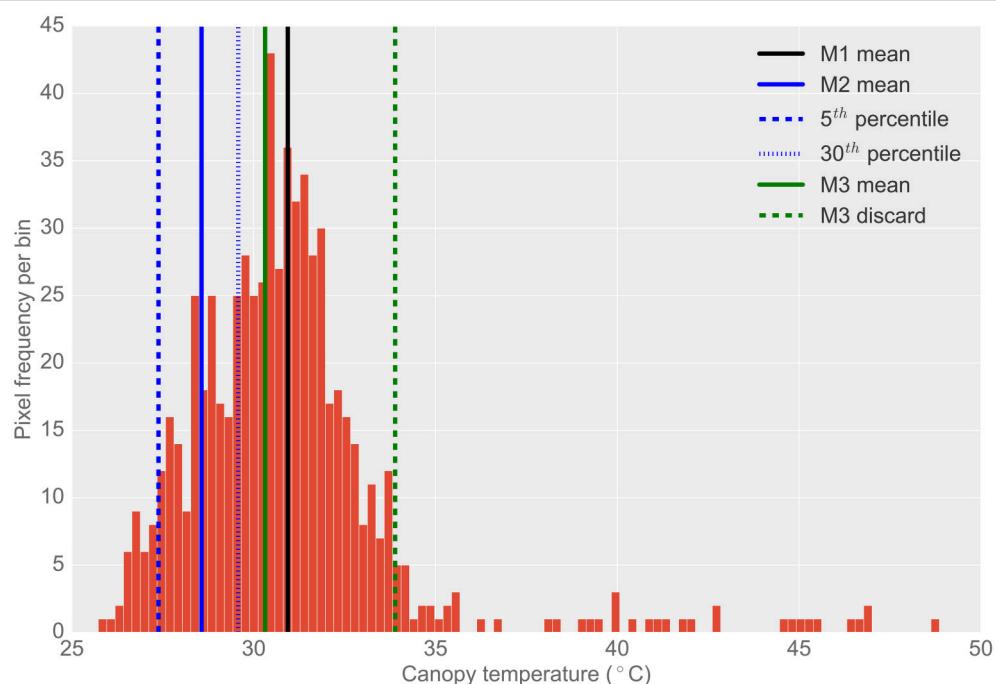
FIGURE 4 | Typical histogram of the frequency distribution of CT pixels from airborne thermography for a single 2 × 6 m experimental plot. Different histogram trimming methods used to minimize bias from background soil temperature are shown. M1, The mean of all pixels (no pixels discarded). M2, The mean of coolest 25th percentile remaining after discarding coolest 5th percentile, designed to extract leaf temperature only. M3, Designed to discard warm pixels from soil patches resulting from poor establishment or biomass sampling. Total pixels in histogram (i.e., M1): 671. Total pixels using M2: 167. Total pixels using M3: 623. Difference between M1 and M2 was 2.4°C. Difference between M1 and M3 was 0.6°C. Difference between M3 and M2 was 1.7°C. Refer Section 2.4.2 for details on M1, M2, and M3.

TABLE 1 | Summary of hand-held CT sampling events, weather conditions and resulting broad-sense heritabilities.

Date and time	Broad-sense heritability	Air temp. (°C)	Rel. humidity (%)	Wind speed (km/h)
18 October 2013 11:00 to 11:30	0.17	11.9 (09:00)	48.0 (09:00)	11.0 (09:00)
25 October 2013 13:50 to 14:20	0.13	17.8 (13:30) 19.0 (14:00)	28.0 (13:30) 28.0 (14:00)	17.0 (13:30) 17.0 (14:00)

Hand-held CT measurements were made on two occasions in 2013 on Treatment 2 only. In Treatment 2, irrigation was supplied to achieve the equivalent of a decile eight rainfall (wettest 20% of years) for the site. Treatment 2 comprised 348 plots. Values in parenthesis denote the time of day when air temperature, relative humidity, and wind speed were recorded. Note, that weather conditions were only recorded prior to the commencement of measurements on the 18th October 2013.

a given flight time, there was very little difference in broad-sense heritability for the three pixel handling methods. This result is in accordance with Figure S1, which shows the Pearson correlation calculated for all methods at each flight time. At any given flight time, all three pixel handling methods were highly correlated with Pearson correlations exceeding 0.86 and averaging 0.93. The background soil temperature did influence the observed CT but in this example, did not influence measurement repeatability (i.e., broad-sense heritability).

3.3. Analysis of the Pixel Frequency Distribution

The aerial CT data incorporates influences from the background soil and the plant canopy. To investigate the significance of the effect of background soil, pairwise difference plots between M1, M2, and M3 (i.e., M1 and M2, M1 and M3, M3 and M2) were generated for airborne thermography data captured from Yanco MEF, 2nd October 2014, using the method of Bland and Altman (1986):

- Where i is a given set of temperature pixels derived at a particular time from a given plot rectangle produced from the ChopIt software described in Section 2.3.2.
- For a given pair, the mean of the pair as the abscissa (x-axis) value e.g., $\frac{M1_i + M2_i}{2}$.
- For a given pair, the difference between the two methods as the ordinate (y-axis) value e.g., $M1_i - M2_i$.

The difference against mean plots are shown in Figure S2. For M1 and M2 (Figure S2A), and M3 and M2 (Figure S2C), the differences increased with time of day until 11:00 h, then from 12:00 to 14:00 h the differences decreased (M1 and M2 mean decreased 0.19°C). From the mean difference calculated across all sample times, M1 and M3 were on average 1.13°C and 0.94°C warmer, respectively, than M2. Further, M1 and M3 were as much as *ca.* 3.0°C warmer than M2 at sample times close to solar
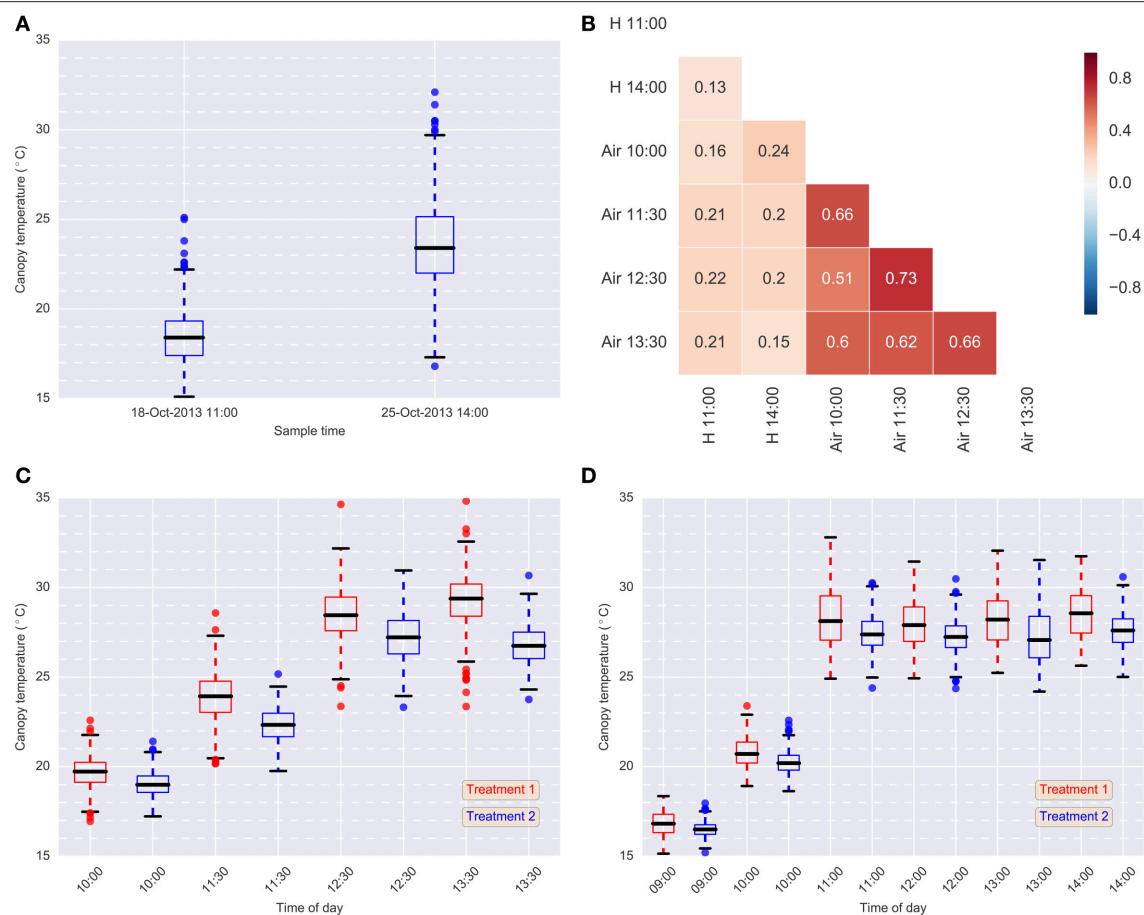
FIGURE 5 | Data summary. (A) Box-plots of hand-held CT data for each sample event in 2013 (Treatment 2 only); **(B)** Pearson correlations between hand-held CT, H, (18-Oct-2013 11:00 and 25-Oct-2013 14:00) and airborne CT, Air, (24-Oct-2013) (Treatment 2); **(C)** box-plots of airborne thermography CT data for each flight time and treatment on 24-Oct-2013; **(D)** box-plots of airborne thermography CT data for each flight time and treatment on 2-Oct-2014. Each box-plot represents CT data from 384 experimental plots (384 experimental plots per treatment and 768 experimental plots in total). Airborne CT for each experimental plot was derived using M1: mean of all pixels (no pixels discarded). In **(C,D)**, Treatment 2 was consistently cooler than Treatment 1, owing to the greater water limitation applied to Treatment 1. In Treatment 2, irrigation was supplied to achieve the equivalent of a decile eight rainfall (wettest 20% of years) for the site, while in Treatment 1, irrigation was supplied to achieve a water limitation close to the long-term climate median for the site.

noon (11:00, 12:00, and 13:00 h). The majority of differences between M1 and M3 (Figure S2B) were close to zero and the mean difference across all sample times was 0.19°C.

4. DISCUSSION

4.1. High Broad-Sense Heritability Obtained from Airborne Thermography Methodology

The main finding reported herein is the large broad-sense heritability obtained for CT from the airborne thermography method, which contrasts with the low heritabilities reported with hand-held thermography sampling methods. Further, this was demonstrated in a large experiment comprising diverse wheat germplasm typical of a commercial wheat breeding program. Across both years, the broad-sense heritability for the airborne

thermography ranged from 0.34 to 0.79, while for the hand-held infra-red thermometer, broad-sense heritability ranged from 0.13 to 0.17. Further, aside from two early morning measurements (09:00 and 10:00 h) on Treatment 1 in 2014, which ranged from 0.34 to 0.46, broad-sense heritability for the airborne thermography ranged from 0.52 to 0.79. The larger broad-sense heritabilities obtained from the airborne thermography can be attributed to the acquisition of thermal images of the entire experiment at effectively a single point in time, thereby overcoming confounding changes in local weather conditions during sampling to provide reliable assessment of CT for large experiments comprising hundreds of 10 m^2 sized plots. Moreover, by measuring CT at effectively a single point in time, statistical analysis need only account for the spatial variation in CT, likely due to the below ground effects of soil structure and water availability, which can be accommodated by the experiment design and spatial analysis (Gilmour et al.,
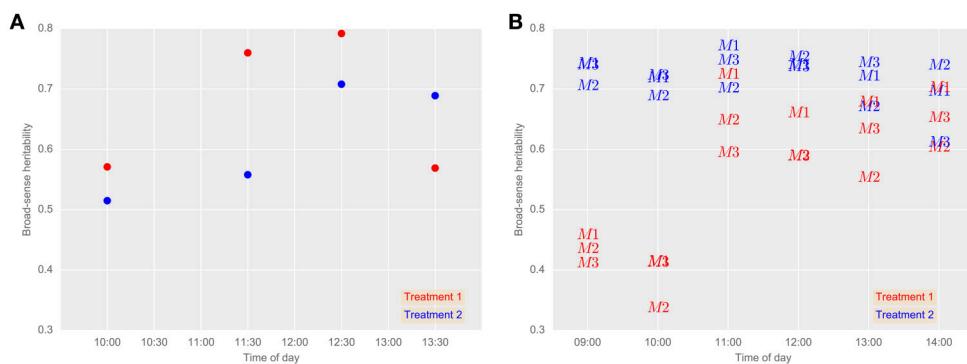
FIGURE 6 | Broad-sense heritabilities for airborne thermography CT for each flight event and treatment in 2013 (A) and 2014 (B). (A) CT for each experimental plot was derived using M1: mean of all pixels (no pixels discarded). 24th October 2013 (grain-filling). **(B)** Comparison between airborne thermography pixel handling methods. 2nd October 2014 (anthesis). M1, mean of all pixels (no pixels discarded). M2, The mean of coolest 25th percentile remaining after discarding coolest 5th percentile, in order to extract plant CT only. M3, Designed to discard warm pixels from soil patches resulting from poor establishment or biomass sampling. Refer Section 2.4.2 for details on M1, M2, and M3. In Treatment 2, irrigation was supplied to achieve the equivalent of a decile eight rainfall (wettest 20% of years) for the site, while in Treatment 1, irrigation was supplied to achieve a water limitation close to the long-term climate median for the site. Generally, broad-sense heritability was high regardless of the pixel handling method used, except for early morning measurements on Treatment 1 in 2014.

1997). In contrast, for the hand-held thermography method, the spatial analysis is confounded by temporal variation in weather conditions, which are more difficult to account for in the statistical analysis.

Broad-sense heritabilities obtained from the hand-held infrared thermometer were small, ranging from 0.13 to 0.17, and typical of our experience with experiments of similar size previously undertaken at the Yanco site (data not shown). Further, Pearson correlations between the hand-held (18-Oct-2013 11:00 and 25-Oct-2013 14:00) and airborne thermography (24-Oct-2013) measures were <0.25, and the correlation between the two hand-held measurement events was low (0.13) (**Figure 5B**). It is likely that during the time required to measure all plots with the hand-held thermography method (*ca.* 30 min), the seemingly small changes in local weather conditions confounded the CT measurements, thereby resulting in low broad-sense heritabilities (**Table 1**). Another contributing factor might also be the range in area and canopy structure sampled by the user as they moved through the experiment. By contrast, an airborne thermography measurement of each plot in the entire experiment took approximately 3 s – a measurement of CT for 768 experimental plots at effectively a single point in time and at a common height above the ground. The heritability of hand-held CT can potentially be improved by using the time of sampling in the statistical analysis. For example, Rebetzke et al. (2013b) improved the heritability of hand-held CT by fitting "time of sampling" as a fixed linear effect in a mixed linear model.

To the best of our knowledge, heritability of CT is typically small on a single-plot basis and seldom reported in the literature. More commonly reported is heritability of CT estimated on a line-mean basis where multiple environments are included in the calculation. Heritabilities of CT calculated on a line-mean basis are often small to moderate in size for both diverse germplasm and related families such as recombinant inbred lines (RILs) and doubled-haploid (DH) lines. For

example, Rebetzke et al. (2013b), using hand-held CT in three wheat populations containing 144–178 DH lines assessed in four irrigated environments, reported small narrow-sense heritabilities (0.12–0.32) on a single-plot basis and moderate to high line-mean heritability ranging from 0.38 to 0.91. Pinto et al. (2010) reported broad-sense heritability of 0.49 for CT measured during grain-filling on a RIL wheat population comprising 167 lines grown in six field experiments under drought and heat environments. In a separate study under similar environmental conditions, Lopes and Reynolds (2012) reported moderate broad-sense heritability on both a wheat population comprising 169 RILs (0.34) and 294 elite wheat lines from CIMMYT (0.38). Others have reported moderate line-mean heritability for diverse wheat germplasm calculated from studies comprising RILs in multiple environments (e.g., Reynolds et al., 2007b; Rattey et al., 2011; Lopes et al., 2012). In the above-mentioned studies, CT was measured using hand-held infrared thermometers. In the study reported herein, **Figure 5** shows that the broad-sense heritability for the airborne thermography method, calculated on a single-plot basis, was typically >0.50 and as high as 0.79, which is considerably greater than literature reported calculations of CT heritability on both a single-plot and line-mean basis.

4.2. Analysis of the Temperature Pixel Frequency Distribution Did Not Improve Broad-Sense Heritability

In this study, methods based on filtering the frequency distribution of the temperature pixels to remove the influence of background soil did not improve broad-sense heritability (**Figure 5**). However, it is likely that the accuracy of the CT data was improved with the CT derived from M2. The difference against mean plots (Figure S2) show that for M1 and M2 (Figure S2A), and M3 and M2 (Figure S2C), the differences increased with the time of day until 11:00 h. This is possibly

because the soil temperature increased more than the plant temperatures, thereby biasing the CT derived from M1 and M3. For M1 and M2, and M3 and M2, the decrease in differences from 12:00 to 14:00 h (for M1 and M2, the mean decreased 0.19°C) may have been due to the lower sun angle in the afternoon increasing the shaded portion of soil and thereby cooling it. That many of the differences between M1 and M3 (**Figure S2B**) were close to zero, indicates that M2 was more effective at deriving plant-based CT than M1 and M3. In contrast to M1 and M3, M2 was derived after discarding the warmest 70th percentile and it is therefore unlikely to be biased by the soil temperature, which, for dry soil, is likely to be warmer than the plant canopy. This approach of sampling cooler pixels, is likely to result in M2 more accurately approximating the actual plant CT than M1 and M3. The potential for improved CT accuracy may be beneficial in applications using energy balance equations to calculate stomatal conductance or transpiration, where soil-biased CT can lead to significant errors (Leinonen et al., 2006; Guilioni et al., 2008).

There may be phenotyping applications where the background soil could significantly reduce the accuracy and precision of CT measurements. For example: when multiple biomass samples have been taken from a plot leaving large areas of exposed soil; in plots with poor plant establishment; in early generation breeding trials or situations where seed number is limited and phenotyping is required on single plants or spaced rows; where row-spacing is too wide to completely cover the soil and in experiments that use raised beds with wide row spacings. To remove the influence of background soil, the custom developed ChopIt image processing software provides a high level of quality control to manually exclude poor quality sections of the plot or sections of the plot where biomass samples have been removed (**Figure 3**). In addition to this feature, post-processing based on the pixel frequency distribution (e.g., M2 and M3) is possible as all the pixels for a particular plot rectangle are stored in a SQLite database file.

4.3. Frame by Frame Image Processing

The ChopIt software enables processing of the images on a frame-by-frame basis and was custom built for the application of field phenotyping of CT. Our image processing method contrasts with the widely used mosaicking method, where a large number of single frame images containing many plots are used to create a mosaic from which plot level information is extracted (e.g., Berni et al., 2009b; Chapman et al., 2014; Gómez-Candón et al., 2016). Mosaics were attempted with thermal images obtained from our image acquisition system. However, the use of mosaics presents a number of issues, namely: the fact that mosaicking software tends to modify the pixel's value in favor of the visual result; the mosaicking process is computationally intensive; and mosaicking requires accurate measurements of the external orientation of the images via the integration of the camera with a GPS and inertial measurement unit (IMU). For our application, processing thermal images on a single frame basis confers a number of advantages over mosaicking including: a reduction in image processing time; higher CT accuracy from working with original temperature values from the raw images without the application of any pixel interpolation or blending; and no mosaicking

"black box", which introduces another layer of measurement uncertainty to the process.

Conversely, the requirement to process the images on a frame by frame basis introduced a trade-off between encompassing the entire experiment in a single helicopter pass, whilst maximizing the pixel resolution by flying no higher than necessary. However, the requirement to encompass an entire experiment in a single pass conferred many advantages including reduced helicopter flight time and cost, faster image processing, reduced image processing errors, and the influence of changing weather conditions on the observed CT were minimized.

4.4. Unmanned Aerial Vehicles

Unmanned aerial vehicles (UAVs) and tethered balloons have also been used for the acquisition of thermal images in field phenotyping applications (e.g., Sullivan et al., 2007; Berni et al., 2009a,b; Jones et al., 2009; Zarco-Tejada et al., 2012; Chapman et al., 2014; Gómez-Candón et al., 2016). The smaller form and, in some jurisdictions, non-requirement for a licensed operator may enable opportunistic sampling on small experiments, whereas the manned helicopter system used in this study might otherwise be considered too expensive to hire or may not be locally available. However, UAVs are often limited to a small camera payload (e.g., 1.5–1.1 kg in Chapman et al., 2014 and 3.0 kg in Gómez-Candón et al., 2016); have limited endurance (e.g., 30–60 min in Chapman et al., 2014); are highly susceptible to wind; are often required to be operated within line of sight and sometimes require a license to operate. Moreover, the image mosaicking process often reported in the literature with UAVs necessitates multiple passes of the experiment to achieve sufficient image overlap. For example, Chapman et al. (2014) used a transect width of 10 m, while Gómez-Candón et al. (2016) used track and cross-track overlaps of 80 and 60%, respectively. As discussed above, such requirement for multiple passes increases the required flight time for a given experiment and increases the likelihood that changes in local weather conditions will confound to compromise measurements of CT. However, the ChopIt frame-by-frame image processing software could potentially be used with images acquired from a UAV platform, provided the image acquisition considerations described in Section 2.3.1 are adhered to.

In contrast to many UAVs described in the literature, the thermal image acquisition system used in this study, comprising a manned helicopter fitted with a helicopter cargo pod (**Figure 1**), has a payload limit of 45 kg. The large payload limit permits the use of a radiometrically-calibrated thermal camera with high accuracy and pixel to pixel sensitivity that negates the need for ground infra-red calibration targets and temperature correction during post-processing (Gómez-Candón et al., 2016). Together, these simplify the image processing. Moreover, the large payload provides the option to add more cameras and sensors if required for additional tasks and enables carriage of a high-capacity battery sufficient for several hours operating time. Further, the use of manned helicopter enables acquisition of CT measurements from multiple large field trials in a short time, which would otherwise require a UAV to fly beyond visual line-of-site, which is not permitted in some jurisdictions.

4.5. Potential for Deployment of CT within Commercial Breeding Programs

In breeder's trials, plot size is often smaller than those sown in this study (e.g., Rebetzke et al., 2014). A single 10 s pass of the helicopter-mounted thermal camera can capture up to 5000 individual 4 m^2 plots in a breeder's yield trial. This application is ideally suited to the airborne thermography method, which we have shown readily scales up to experiments comprising 1000 individual 10 m^2 plots. For 1000 plots of 2 × 6 m, acquisition of CT on a per plot basis using the airborne thermography and data handling described here takes *ca.* 25 min and aside from the helicopter pilot, requires only one person. The method could be used within a breeding program to assess spatial uniformity of yield experiments and provide guidance to appropriate statistical spatial models. The demonstrated link between CT and grain yield (Reynolds et al., 1994; Amani et al., 1996; Fischer et al., 1998; Ayeneh et al., 2002; Rattey et al., 2011; Rebetzke et al., 2013b) should provide opportunity to select for CT in early generation screening. For example, selection of CT in early generations was demonstrated in studies reporting reasonable genetic correlation between small plot CT, leaf porosity and full plot yield (Condon et al., 2004, 2007). In addition, augmenting breeder's visual selection with early generation measurements of CT can potentially identify a greater number of equally high yielding lines compared to breeder's visual selection alone (van Ginkel et al., 2004). Importantly, economic analysis indicates that the incorporation of CT measurements within a wheat breeding program is likely to provide an economic benefit (Brennan et al., 2007).

To assist uptake by breeders, several improvements in the airborne thermography method described here are possible, including: remote automation of the image acquisition process; use of a smaller manned helicopter to reduce the operating cost (e.g., Robinson R22 Raven helicopter); and use of GPS geo-referencing to improve image processing. Differences in canopy architecture that may influence CT could be accounted for by making use of measurements of fractional ground cover, from digital camera (e.g., Li et al., 2010), and canopy height that can now be routinely measured by ground-based LiDAR (e.g., Deery et al., 2014) but possibly aerial LiDAR in the future. Together, these potential improvements could reduce the cost per plot of the airborne thermography method.

The high helicopter operating cost, AU$1000/h, may prohibit the use of the airborne thermography method within some breeding (and research) programs. However, the cost per plot of the airborne method, on 3000 plots of size 10 m^2, equates to AU$0.39 (*ca.* US$0.30) (Table S1), which is within 30% of the hand-held cost per plot reported by Brennan et al. (2007) (US$0.19 in 2007, which equates to US$0.22 in 2016 after adjusting for inflation). Given the similar cost per plot of the two methods, together with the greater repeatability of the airborne CT method compared with the hand-held CT method, the airborne CT method could be a cost-effective CT phenotyping method for use within breeding (and research) programs.

5. CONCLUSION

CT, as a surrogate measure for stomatal conductance and potentially photosynthesis, has been associated with genotypic variation in grain yield in numerous studies and therefore mooted as a possible phenotypic selection tool for use in genetics studies or in breeder's trials. For this to be realized, an inexpensive, scalable, and reliable CT methodology is required. The airborne thermography methodology described herein is such a method. The method is highly repeatable, as evidenced by the high broad-sense heritabilities obtained. The method is scalable: for an experiment comprising 768 plots of size 2 × 6 m, it takes *ca.* 25 min to obtain a CT measurement for each individual plot for statistical analysis. Moreover, the method requires only one person (not including the helicopter pilot) and utilizes purpose built image processing software for use by a non-technical user.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01808/full#supplementary-material

REFERENCES

Amani, I., Fischer, R. A., and Reynolds, M. F. P. (1996). Canopy temperature depression association with yield of irrigated spring wheat cultivars in a hot climate. *J. Agron. Crop Sci.* 176, 119–129.

Ayeneh, A., van Ginkel, M., Reynolds, M., and Ammar, K. (2002). Comparison of leaf, spike, peduncle and canopy temperature depression in wheat under heat stress. *Field Crops Res.* 79, 173–184. doi: 10.1016/S0378-4290(02)00138-7

Bennett, D., Reynolds, M., Mullan, D., Izanloo, A., Kuchel, H., Langridge, P., et al.(2012). Detection of two major grain yield QTL in bread wheat (*Triticum aestivum* L.) under heat, drought and high yield potential environments. *Theor. Appl. Genet.* 125, 1473–1485. doi: 10.1007/s00122-012-1927-2

Berni, J., Zarco-Tejada, P., Sepulcre-Cantó, G., Fereres, E., and Villalobos, F. (2009a). Mapping canopy conductance and CWSI in olive orchards using high resolution thermal remote sensing imagery. *Remote Sensing Environ.* 113, 2380–2388. doi: 10.1016/j.rse.2009.06.018

Berni, J. A. J., Zarco-Tejada, P. J., Suarez, L., and Fereres, E. (2009b). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Trans. Geosci. Remote Sensing* 47, 722–738. doi: 10.1109/TGRS.2008.2010457

Bland, J. M., and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327, 307–310.

Blum, A., Shpiler, L., Golan, G., and Mayer, J. (1989). Yield stability and canopy temperature of wheat genotypes under drought-stress. *Field Crops Res.* 22, 289–296.

Brennan, J. P., Condon, A. G., van Ginkel, M., and Reynolds, M. P. (2007). An economic assessment of the use of physiological selection for stomatal aperture-related traits in the CIMMYT wheat breeding programme. *J. Agri. Sci.* 145, 187–194. doi: 10.1017/S0021859607007009

Chapman, S. C., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, M. F., et al. (2014). Pheno-copter: a low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping. *Agronomy* 4, 279–301. doi: 10.3390/agronomy4020279

Condon, A., Reynolds, M., Rebetzke, G., van Ginkel, M., Trethowan, R., Bonnett, D., et al. (2004). "Physiological traits as indirect selection criteria for yield potential in bread wheat," in *Cereals 2004, Proceedings 54th Australian Cereal Chemistry Conference and 11th Wheat Breeders Assembly, Canberra, 21–24 September 2004*, eds C. Black, J. Panozzo, and G. Rebetzke (Canberra, ACT), 112–115.

Condon, A. G., Reynolds, M. P., Rebetzke, G. J., van Ginkel, M., Richards, R. A., and Farquhar, G. D. (2007). "Using stomatal aperture-related traits to select for high yield potential in bread wheat, chapter 12," in *Wheat Production in Stressed Environments: Proceedings of the 7th International Wheat Conference, 27 November–2 December 2005, Mar del Plata, Argentina*, Vol. 12, eds H. Buck, J. Nisi, and N. Salomón (Dordrecht: Springer Netherlands), 617–624.

Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., and Furbank, R. (2014). Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy* 4, 349–379. doi: 10.3390/agronomy4030349

Fischer, R. A., Rees, D., Sayre, K. D., Lu, Z.-M., Condon, A. G., and Saavedra, A. L. (1998). Wheat yield progress associated with higher stomatal conductance and photosynthetic rate, and cooler canopies. *Crop Sci.* 38, 1467–1475.

Gilmour, A. R., Cullis, B. R., and Verbyla, A. P. (1997). Accounting for natural and extraneous variation in the analysis of field experiments. *J. Agri. Biol. Environ. Stat.* 2, 269–293.

Gómez-Candón, D., Virlet, N., Labbé, S., Jolivot, A., and Regnard, J.-L. (2016). Field phenotyping of water stress at tree scale by UAV-sensed imagery: new insights for thermal acquisition and calibration. *Precision Agric.* 17, 786–800. doi: 10.1007/s11119-016-9449-6

Guilioni, L., Jones, H. G., Leinonen, I., and Lhomme, J. P. (2008). On the relationships between stomatal resistance and leaf temperatures in thermography. *Agric. For. Meteorol.* 148, 1908–1912. doi: 10.1016/j.agrformet.2008.07.009

Holland, J. B., Nyquist, W. E., and Cervantes-Martínez, C. T. (2003). Estimating and interpreting heritability for plant breeding: an update. *Plant Breeding Rev.* 22, 9–112. doi: 10.1002/9780470650202.ch2

Isbell, R. F. (1996). *The Australian Soil Classification*, Vol. 4. Collingwood, VIC: CSIRO Australia.

Jones, H., and Sirault, X. (2014). Scaling of thermal images at different spatial resolution: the mixed pixel problem. *Agronomy* 4, 380–396. doi: 10.3390/agronomy4030380

Jones, H. G. (2002). Use of infrared thermography for monitoring stomatal closure in the field: application to grapevine. *J. Exp. Bot.* 53, 2249–2260. doi: 10.1093/jxb/erf083

Jones, H. G., Serraj, R., Loveys, B. R., Xiong, L., Wheaton, A., and Price, A. H. (2009). Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. *Funct. Plant Biol.* 36, 978–989. doi: 10.1071/FP09123

Leinonen, I., Grant, O. M., Tagliavia, C. P. P., Chaves, M. M., and Jones, H. G. (2006). Estimating stomatal conductance with thermal imagery. *Plant Cell Environ.* 29, 1508–1518. doi: 10.1111/j.1365-3040.2006.01528.x

Li, Y., Chen, D., Walker, C., and Angus, J. (2010). Estimating the nitrogen status of crops using a digital camera. *Field Crops Res.* 118, 221–227. doi: 10.1016/j.fcr.2010.05.011

Lopes, M., Reynolds, M., Jalal-Kamali, M., Moussa, M., Feltaous, Y., Tahir, I., et al. (2012). The yield correlations of selectable physiological traits in a population of advanced spring wheat lines grown in warm and drought environments. *Field Crops Res.* 128, 129–136. doi: 10.1016/j.fcr.2011.12.017

Lopes, M. S., and Reynolds, M. P. (2010). Partitioning of assimilates to deeper roots is associated with cooler canopies and increased yield under drought in wheat. *Funct. Plant Biol.* 37, 147–156. doi: 10.1071/FP09101

Lopes, M. S., and Reynolds, M. P. (2012). Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. *J. Exp. Bot.* 63, 3789–3798. doi: 10.1093/jxb/ers071

Mason, R. E., Hays, D. B., Mondal, S., Ibrahim, A. M. H., and Basnet, B. R. (2013). QTL for yield, yield components and canopy temperature depression in wheat under late sown field conditions. *Euphytica* 194, 243–259. doi: 10.1007/s10681-013-0951-x

Mason, R. E., and Singh, R. P. (2014). Considerations when deploying canopy temperature to select high yielding wheat breeding lines under drought and heat stress. *Agronomy* 4, 191–201. doi: 10.3390/agronomy4020191

Olivares-Villegas, J. J., Reynolds, M. P., and McDonald, G. K. (2007). Drought-adaptive attributes in the Seri/Babax hexaploid wheat population. *Funct. Plant Biol.* 34, 189–203. doi: 10.1071/FP06148

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybernet.* 9, 62–66. doi: 10.1109/tsmc.1979.4310076

Pinter, P., Zipoli, G., Reginato, R., Jackson, R., Idso, S., and Hohman, J. (1990). Canopy temperature as an indicator of differential water use and yield performance among wheat cultivars. *Agric. Water Manage.* 18, 35–48.

Pinto, R. S., Reynolds, M. P., Mathews, K. L., McIntyre, C. L., Olivares-Villegas, J.-J., and Chapman, S. C. (2010). Heat and drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. *Theo. Appl. Genet.* 121, 1001–1021. doi: 10.1007/s00122-010-1351-4

Prashar, A., and Jones, H. (2014). Infra-red thermography as a high-throughput tool for field phenotyping. *Agronomy* 4, 397–417. doi: 10.3390/agronomy4030397

Prashar, A., Yildiz, J., McNicol, J. W., Bryan, G. J., and Jones, H. G. (2013). Infra-red thermography for high throughput field phenotyping in *Solanum tuberosum*. *PLoS ONE* 8:e65816. doi: 10.1371/journal.pone.0065816

Rashid, A., Stark, J. C., Tanveer, A., and Mustafa, T. (1999). Use of canopy temperature measurements as a screening tool for drought tolerance in spring wheat. *J. Agron. Crop Sci.* 182, 231–238.

Rattey, A., Shorter, R., and Chapman, S. (2011). Evaluation of CIMMYT conventional and synthetic spring wheat germplasm in rainfed sub-tropical environments. II. Grain yield components and physiological traits. *Field Crops Res.* 124, 195–204. doi: 10.1016/j.fcr.2011.02.006

Rebetzke, G. J., Chenu, K., Biddulph, B., Moeller, C., Deery, D. M., Rattey, A. R., et al. (2013a). A multisite managed environment facility for targeted trait and germplasm phenotyping. *Funct. Plant Biol.* 40, 1–13. doi: 10.1071/FP12180

Rebetzke, G. J., Fischer, R. T. A., van Herwaarden, A. F., Bonnett, D. G., Chenu, K., Rattey, A. R., et al. (2014). Plot size matters: interference from intergenotypic competition in plant phenotyping studies. *Funct. Plant Biol.* 41, 107–118. doi: 10.1071/FP13177

Rebetzke, G. J., Rattey, A. R., Farquhar, G. D., Richards, R. A., and Condon, A. G. (2013b). Genomic regions for canopy temperature and their genetic association

with stomatal conductance and grain yield in wheat. *Funct. Plant Biol.* 40, 14–33. doi: 10.1071/FP12184

Reynolds, M., Balota, M., Delgado, M., Amani, I., and Fischer, R. (1994). Physiological and morphological traits associated with spring wheat yield under hot, irrigated conditions. *Aust. J. Plant Physiol.* 21, 717–730. doi: 10.1071/PP9940717

Reynolds, M., Dreccer, F., and Trethowan, R. (2007a). Drought-adaptive traits derived from wheat wild relatives and landraces. *J. Exp. Bot.* 58, 177–186. doi: 10.1093/jxb/erl250

Reynolds, M., Saint-Pierre, C., Saad, A., Vargas, M., and Condon, A. (2007b). Evaluating potential genetic gains in wheat associated with stress-adaptive trait expression in elite genetic resources under drought and heat stress. *Crop Sci.* 47, S172–S189.

Saint Pierre, C., Crossa, J., Manes, Y., and Reynolds, M. P. (2010). Gene action of canopy temperature in bread wheat under diverse environments. *Theor. Appl. Genet.* 120, 1107–1117. doi: 10.1007/s00122-009-1238-4

Sullivan, D., Fulton, J., Shaw, J., and Bland, G. (2007). Evaluating the sensitivity of an unmanned thermal infrared aerial system to detect water stress in a cotton canopy. *Trans. ASABE* 50, 1955–1962.

Takai, T., Yano, M., and Yamamoto, T. (2010). Canopy temperature on clear and cloudy days can be used to estimate varietal differences in stomatal conductance in rice. *Field Crops Res.* 115, 165–170. doi: 10.1016/j.fcr.2009.10.019

van Ginkel, M., Reynolds, M., Trethowan, R., and Hernandez, E. (2004). "Can canopy temperature depression measurements help breeders in selecting for yield in wheat under irrigated production conditions?," in *New Directions for a Diverse Planet: Proceedings for the 4th International Crop Science Congress, Brisbane, Australia, 26 September - 1 October 2004*, eds R. A. Fischer, N. Turner, J. Angus, L. McIntyre, M. Robertson, A. Borrell, and D. Lloydpage (Brisbane, QLD: The Regional Institute Ltd.), 3.4.6.

Zarco-Tejada, P., González-Dugo, V., and Berni, J. (2012). Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing Environ.* 117, 322–337. doi: 10.1016/j.rse.2011.10.007

Assessing Wheat Traits by Spectral Reflectance: Do We Really Need to Focus on Predicted Trait-Values or Directly Identify the Elite Genotypes Group?

Miguel Garriga[1], Sebastián Romero-Bravo[1], Félix Estrada[1], Alejandro Escobar[1], Iván A. Matus[2], Alejandro del Pozo[1], Cesar A. Astudillo[3] and Gustavo A. Lobos[1]*

[1] Facultad de Ciencias Agrarias, Plant Breeding and Phenomic Center, PIEI Adaptación de la Agricultura al Cambio Climático, Universidad de Talca, Talca, Chile, [2] CRI-Quilamapu, Instituto de Investigaciones Agropecuarias, Chillán, Chile, [3] Department of Computer Science, Faculty of Engineering, Universidad de Talca, Curicó, Chile

Phenotyping, via remote and proximal sensing techniques, of the agronomic and physiological traits associated with yield potential and drought adaptation could contribute to improvements in breeding programs. In the present study, 384 genotypes of wheat (*Triticum aestivum* L.) were tested under fully irrigated (FI) and water stress (WS) conditions. The following traits were evaluated and assessed via spectral reflectance: Grain yield (GY), spikes per square meter (SM2), kernels per spike (KPS), thousand-kernel weight (TKW), chlorophyll content (SPAD), stem water soluble carbohydrate concentration and content (WSC and WSCC, respectively), carbon isotope discrimination (Δ^{13}C), and leaf area index (LAI). The performances of spectral reflectance indices (SRIs), four regression algorithms (PCR, PLSR, ridge regression RR, and SVR), and three classification methods (PCA-LDA, PLS-DA, and *k*NN) were evaluated for the prediction of each trait. For the classification approaches, two classes were established for each trait: The lower 80% of the trait variability range (Class 1) and the remaining 20% (Class 2 or elite genotypes). Both the SRIs and regression methods performed better when data from FI and WS were combined. The traits that were best estimated by SRIs and regression methods were GY and Δ^{13}C. For most traits and conditions, the estimations provided by RR and SVR were the same, or better than, those provided by the SRIs. PLS-DA showed the best performance among the categorical methods and, unlike the SRI and regression models, most traits were relatively well-classified within a specific hydric condition (FI or WS), proving that classification approach is an effective tool to be explored in future studies related to genotype selection.

Keywords: phenomic, high-throughput phenotyping, phenotyping, carbon isotope discrimination, reflectance

INTRODUCTION

Wheat is one of the most important cereals in the human diet worldwide. This cereal is consumed in different types of processed foods, providing around 20% of total daily calories (Shewry, 2009). Due to world population growth, it is expected that current wheat production will need to be doubled by the middle of the century (Tilman et al., 2011; FAO, 2015). To accomplish this level of production, wheat yield must increase by 1.60% per year (Dixon et al., 2009), which is far from the 1.26% increase that was reached during the last decade (FAOSTAT, 2013). Additionally, the current effects of climate change on weather patterns, and unexpected events, are threatening maximum thresholds in many areas (Ayeneh et al., 2002; Azimi et al., 2010; Rebetzke et al., 2012; Hernández-Barrera et al., 2016).

This challenging scenario should encourage wheat breeders to accelerate the development and release of new high-yield cultivars that are adapted to more complex environmental conditions (Velu and Prakash, 2013). One strategy for improving and expediting the selection of these elite genotypes is the acquisition of high-dimensional phenotypic data (high-throughput phenotyping) (Bowman et al., 2015; Camargo and Lobos, 2016; Crain et al., 2016).

Remote sensing techniques, such as spectrometry, are increasingly used for plant phenotyping (Cabrera-Bosquet et al., 2012; Araus and Cairns, 2014). Spectral reflectance or the spectrum of energy reflected by the plant is closely associated with absorption at certain wavelengths that are linked to specific characteristics or plant conditions (Lobos and Hancock, 2015). Spectrometers can acquire detailed information regarding the electromagnetic spectrum in a short time, making this technology ideal for assessing hundreds or thousands of genotypes within a few hours (Cabrera-Bosquet et al., 2012). This would enable researchers and breeders to estimate multiple morpho-physiological and physico-chemical traits, which would otherwise be impossible to evaluate due to the time and cost involved (Lobos and Hancock, 2015). This would be reflected in reduced breeding program costs and, by allowing for the early selection of genetic material of interest, increase the chances of releasing improved cultivars in less time (Lobos and Hancock, 2015; Camargo and Lobos, 2016).

For the estimation of wheat traits, such as grain yield, biomass, leaf area index, plant height, and isotopic carbon discrimination, the majority of previous studies have resorted primarily to the use of "Spectral Reflectance Indices" (SRIs) (Aparicio et al., 2002; Babar et al., 2006a,b; Marti et al., 2007; Prasad et al., 2007; Gutierrez et al., 2010; Lobos et al., 2014; Hernández et al., 2015), whereas there has been less attention paid to the development of multivariate regression models, using part or all of the spectral reflectance (Pimstein et al., 2011; Dreccer et al., 2014; Li F. et al., 2014; Li X. et al., 2014; Hernández et al., 2015; Siegmann and Jarmer, 2015; Yao et al., 2015).

Current research into plant phenotyping and phenomics for plant breeding has focused on using the spectral signature to estimate predicted trait-values rather than exploring other tools that could directly identify elite genotypes for the desired trait. The use of reflectance data and categorical methods for breeding purposes has been scarcely addressed by the scientific community. Nonetheless, some successful experiments have been carried out: To classify lines for waxy alleles in durum wheat (Delwiche et al., 2006; Lavine et al., 2014) and bread wheat (Delwiche et al., 2011); to identify wheat lines possessing wheat-rye translocations (Delwiche et al., 1999); to classify barley varieties (Porker et al., 2017); and to select haploid kernels from hybrid kernels in maize (Jones et al., 2012).

The aim of the present study was, based on plant reflectance data, to assess the feasibility of using a categorical approximation to select featured genotypes, by comparing the performance of a large set of SRIs, multivariate regression models (PCR, PLSR, ridge regression, and SVR), and categorical models (PCA-LDA, PLS-DA, and kNN) in the prediction of grain yield, agronomic yield components, and physiological traits.

MATERIALS AND METHODS
Plant Material and Experimental Conditions

A set of 384 cultivars and advanced lines of spring bread wheat (*Triticum aestivum* L.) with good agronomic characteristics and disease tolerance were evaluated (list at del Pozo et al., 2016). These genotypes were sourced from three wheat-breeding programs (INIA-Chile, INIA-Uruguay, and CIMMYT-Mexico) and are currently being used to breed for adaptation to drought stress and to develop suitable genotypes for wheat production in the drylands of Chile and other countries.

Experiments were conducted in 2012 in two contrasting Mediterranean environments of Chile: Cauquenes (35°58′S, 72°17′W), with typical rain-fed conditions such that the plants were grown under water stress (WS); and Santa Rosa (36°32′S, 71°55′W), under fully irrigated (FI) conditions. The precipitation in these locations during the experiments was 183 and 700 mm, respectively (**Table 1**). Under FI conditions, four furrow irrigations of approximately 50 mm each were applied at the end of tillering (Zadocks Stage 21; Z21), the flag leaf stage (Z37), heading (Z50), and middle grain filling (Z70) (Zadoks et al., 1974). There was a difference of 28 days (77–105 days) between the earlier and later genotypes in reaching the heading stage; 89% of the genotypes 81–94 days.

The experiment was conducted in an incomplete block design (α-lattice), with two replicates per genotype ($n = 384 \times 2$). Plots had five rows, each 2 m in length, the distance between the rows was 0.2 m, and the distance between plots was 0.4 m. Similar agronomic practices were performed at the two locations. Sowing ($20 \, g \, m^{-2}$) dates were 23 May at Cauquenes and 7 August at Santa Rosa. Before sowing, the plots were fertilized with $260 \, kg \, ha^{-1}$ of ammonium phosphate (46% P_2O_5 and 18% N), $90 \, kg \, ha^{-1}$ of potassium chloride (60% K_2O), $200 \, kg \, ha^{-1}$ of sul-po-mag (22% K_2O, 18% MgO and 22% S), $10 \, kg \, ha^{-1}$ of boronatrocalcite (11% B), and $3 \, kg \, ha^{-1}$ of zinc sulfate (35% Zn). During tillering, an extra $153 \, kg \, ha^{-1}$ of N was applied. Flufenacet + Flurtamone + Diflufenican (96 g a.i.) was applied for pre-emergence weed control and a further application of

TABLE 1 | Monthly maximum, minimum, and mean temperature, and monthly rainfall at the two experimental sites in central Chile during the trial (May 2012–January 2013).

Cauquenes		May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.
Temp. (°C)	Max.	18.1	14.1	15.8	14.7	19.6	19.9	25.7	24.6	30.8
	Min.	5.3	5.8	4.5	5.5	6.6	5.9	8.5	9.3	12.5
	Mean	11.7	10.0	11.6	11.3	15.0	12.8	16.8	17.0	21.2
Rainfall (mm)		10.0	9.0	14.0	22.8	3.0	24.5	69.6	29.7	0.0
Santa Rosa		**May**	**Jun.**	**Jul.**	**Aug.**	**Sept.**	**Oct.**	**Nov.**	**Dec.**	**Jan.**
Temp. (°C)	Max.	16.2	13.1	12.6	13.2	18.2	18.9	24.2	23.1	30.0
	Min.	4.4	5.1	0.8	3.5	4.1	5.3	7.5	8.1	11.1
	Mean	9.8	8.6	6.1	7.6	10.6	11.8	15.3	15.6	19.9
Rainfall (mm)		90.0	186.9	52.5	139.3	12.7	48.0	59.2	109.6	1.2

MCPA (525 g a.i.) + Metsulfuron-metil (5 g a.i.) was used for post-emergence weed control.

Trait Measurements

Grain Yield and Agronomic Yield Components
The number of spikes per m^2 (SM2) was determined for a 1 m length of an inside row. The number of kernels per spike (KPS) and thousand-kernel weight (TKW) were determined in 25 spikes selected at random from the central row. Grain yield (GY) was assessed by harvesting the whole plot.

Leaf Chlorophyll and Water-Soluble Carbohydrate
Chlorophyll (Chl) content was determined using the SPAD index for five flag leaves per plot, at anthesis (*an*) and at grain filling (*gf*), with a portable leaf chlorophyll-meter (SPAD 502, Minolta Spectrum Technologies Inc., Plainfield, IL, USA). Water-soluble carbohydrate was determined using the anthrone reactive method (Yemm and Willis, 1954), for five main stems (excluding leaf laminas and sheaths) per plot, at *an* and at maturity (*m*), and expressed as concentration (WSC, mg g^{-1} DW) and content (WSCC, mg stem^{-1}).

Carbon Isotope Discrimination
For kernels sampled randomly at *m*, the stable carbon (^{13}C/^{12}C) isotope ratio was measured using an elemental analyzer (ANCA-SL, PDZ Europa, UK) coupled with an isotope ratio mass spectrometer, at the Laboratory of Applied Physical Chemistry at Ghent University (Belgium). The carbon isotope discrimination (Δ^{13}C) was calculated as follows: Δ^{13}C (‰) = (δ^{13}C$_a$− δ^{13}C$_p$)/[1+ (δ^{13}C$_p$)/1000], where *a* and *p* refer to air and plant, respectively (Farquhar et al., 1989). δ^{13}C$_a$ was taken as 8.0‰.

Leaf Area Index
The leaf area index (LAI) at *an* (under FI conditions only) was determined by measuring the incident light falling on the crop and the amount of light in each plot at ground level, using a BF5 Sunshine Sensor and SunScan Canopy Analyser (Delta-T, Cambridge, UK). The radiation transmitted and dispersed by the canopy was recorded, and the LAI then calculated.

Spectral Reflectance Measurements
Canopy spectral reflectance (350–2,500 nm) was measured using a portable spectroradiometer (FieldSpec 3 JR, ASD, Boulder, CO, USA) at two developmental stages: Anthesis (*AN*; denoted with capital letters to avoid confusion with trait measurements stages) and grain filling (*GF*). The optical fiber (2.3 mm diameter with 25° full conical angles) was placed 80 cm above the canopy, at a 45° angle. From 11:00 to 17:00 h on clear sunny days, measurements were taken by moving (sweeping) the sensor over the plot, covering the three central rows. The equipment was set up to take three scans per plot and the average for each wavelength was considered in further analyses. To limit variations in reflectance induced by changes in the angle of the sun, radiometric calibration was performed every 15 min, using a white barium sulfate panel as the reference (Spectralon, ASD, Boulder, CO, USA). Prior to modeling, exploratory analysis and spectral noise deletion were performed using the software Spectral Knowledge (SK-UTALCA) (Lobos and Poblete-Echeverría, 2017).

Modeling Analysis
Spectral reflectance assessed at *AN* and *GF* stages was used to estimate the traits that were evaluated at anthesis (*an*) (Chl content, WSC, WSCC, and LAI), grain filling (*gf*) (Chl content), and maturity (*m*) (SM2, KPS, TKW, GY, WSC, WSCC, and Δ^{13}C). The following analyses were considered.

Spectral Reflectance Indices
Spectral reflectance was used to assess the predictive performance of a set of 255 SRIs loaded on SK-UTALCA (Lobos and Poblete-Echeverría, 2017). Using the linear regression analysis option, the relationships between the each of the measured traits and each of the SRIs at *AN* and *GF* were examined using the coefficient of determination (r^2) and the root mean square error (RMSE). WS and FI conditions were analyzed independently, but also as one environment (WS+FI).

Multivariate Regression Methods
Four different regression methods were considered: Principal Components Regression (PCR), Partial Least Square Regression (PLSR), Ridge Regression (RR), and Support Vector Machine

Regression (SVR). Prior to modeling with R 3.1.2 software (R Development Core Team, 2011), any samples with missing values were excluded, outliers were identified by Local Outlier Detection (LOF), and the data were normalized.

PCR is a combination of Principal Component Analysis (PCA) and Multiple Linear Regression (MLR), which first reduces the dimensionality of the spectral data, concentrating the information into so-called principal components. The transformed data is then used to train a MLR model that fits a linear equation (Jolliffe, 2002).

PLSR reduces the dimensionality of the data by constructing so-called latent factors. Unlike PCA, PLSR produces a set of factors that take into consideration the values of the independent and dependent variables simultaneously. In this sense, PLSR finds vectors that not only represent the variance of the data but that are also related to the response (Wold et al., 2001; Hastie et al., 2005).

RR works in a similar way to least square fitting, but adds a term that penalizes the values of the coefficients. The role of the penalization term is to "shrink" the estimates of the coefficients toward zero (Hastie et al., 2005). This penalization can be controlled using a tuning parameter λ, which has to be estimated independently. The optimization of λ was performed using a grid of 100 possible values of λ in a range of $[10^{-2}, 10^{10}]$ with 10-fold cross-validation. The best λ identified was used to build the model.

SVR is a method derived from the Support Vector Machine (SVM). The SVM transforms data into a new high-dimensional space using a kernel function. In this newly created space, a predictive model is built using a subset of representative instances called support vectors. SVR estimates a linear dependency by fitting an optimal approximating hyperplane to the training samples in the multidimensional feature space. In the present study, several kernels (linear, polynomial, radial basis function, and sigmoidal) were automatically tested and selected based on a minimization criterion. The parameters C (regularization parameter) and ε (loss function parameter) were fixed to 1 and 0.1, respectively.

All models were validated by 10-fold cross-validation (10xCV) and their performances evaluated by the coefficient of determination (R^2), the root mean square error (RMSE), and the Index of Agreement (IA) in calibration and validation. The IA (Willmott, 1981) is a standardized measure for estimating the prediction error of the model and ranges from 0 (faulty model) to 1 (perfect fit).

Multivariate Classification Methods

Spectral reflectance data were also modeled by three different supervised classification methods: Principal Components—Linear Discriminant Analysis (PCA-LDA), Partial Least Square Discriminant Analysis (PLS-DA), and the k-Nearest Neighbor (kNN) algorithm. Two different categories were established by taking into account the total variation of each trait, as measured at each of the developmental stages and in each of the environments (FI, WS, or WS+FI). The first category, labeled as "Class 1," corresponds to instances with values in the lower 80% of the trait range. The remaining 20% of

instances were considered as belonging to the elite group and were labeled as "Class 2" (Supplementary Table 1). The goal of this dichotomization was to develop predictive models that were able to identify those elite genotypes that had the highest trait performance (the upper 20% of the trait range). Model calculations were done using the Classification Tool Box (Version 4.2) developed by Milano Chemometrics and the QSAR Research Group (Ballabio and Consonni, 2013) and implemented in Matlab 8.2.0 (The Math Works Inc., MA, USA).

PCA-LDA is a classification technique based on the linear discriminant functions. PCA is used to reduce the dimensionality of the spectral matrix and LDA acts as the classifier. Classes are separated by maximizing the variance between the groups, and minimizing the variance within the groups, to determine the best fit of parameters for the classification (Lehmann et al., 2015). Before calculating the PCA-LDA models, the input data were mean-centered and the optimal number of principal components was searched in the interval 1–20, with 10xCV, on the basis of minimizing the error rate of validation. The discrimination of classes was established as linear.

PLS-DA is a pattern recognition method that combines the properties of PLSR, discriminating between the categories using the Discriminant Analysis technique (Ballabio and Consonni, 2013). PLS-DA works by finding the latent variables that describe the variance in both the independent X variables (spectra) and the dependent Y variables (classes) and are able to separate the data into two or more classes (Barker and Rayens, 2003). In PLS-DA, a model is developed for each class and the probability that a sample belongs to a specific class is calculated based on the estimated class values (Ballabio and Consonni, 2013). In the present study, the PLS-DA models were calculated using mean-centered data and the optimal number of latent variables was searched in the range [1, 20], with 10xCV.

Finally, the k-nearest neighbors (kNN) method is based on the determination of the distances between an instance whose identity is assumed to be unknown and each instance belonging to the training set. Once the distances are computed, the elements are ranked according to their proximity to the query instance, selecting the k elements that are closest to this. Finally, the category of the query item is estimated using a majority voting scheme among the labels of the k selected items (Cunningham and Delany, 2007). In general, the distance function can be any mathematical function that expresses dissimilarity but, for simplicity, a common choice is the Euclidian distance. In this study, the data was mean-centered and the best value for the parameter k was obtained from values in [1,10], with 10xCV.

The predictive powers of the categorical PCA-LDA, PLS-DA, and kNN models were evaluated by calculations of accuracy, error rate, and prediction rates (determined by the sensibility of each class) for both classes in calibration and validation.

RESULTS

The range of values, and their means, for each of the traits evaluated in the 384 wheat genotypes grown under FI and WS conditions, are presented in **Table 2**.

Trait[y]	FI		WS	
	Range	Mean ± SD	Range	Mean ± SD
SM2m[z]	320.0–1125.0	627.7 ± 133.7	75.0–625.0	321.1 ± 69.3
KPSm	15.0–68.9	38.6 ± 6.5	9.4–55.6	32.3 ± 6.3
TKWm (g)	30.8–88.6	49.3 ± 6.6	29.4–89.4	44.4 ± 6.8
GYm (ton ha^{-1})	5.1–12.9	9.7 ± 1.2	0.9–6.9	3.1 ± 0.9
Chlan (SPAD index)	35.9–58.2	49.3 ± 3.3	31.3–52.0	41.4 ± 3.5
Chlgf (SPAD index)	30.1–56.0	47.8 ± 3.7	0.8–48.8	34.0 ± 10.3
WSCan (mg g^{-1} DW)	16.1–610.9	141.2 ± 49.2	20.6–708.4	226.8 ± 51.7
WSCm (mg g^{-1} DW)	5.6–686.3	43.4 ± 33.3	5.0–218.3	48.0 ± 23.5
WSCCan (mg stem^{-1})	19.3–926.0	172.1 ± 86.3	32.7–1271.6	410.8 ± 140.9
WSCCm (mg stem^{-1})	6.1–926.6	47.4 ± 45.1	4.9–262.9	52.1 ± 34.4
Δ^{13}Cm (‰)	17.1–20.2	18.8 ± 0.5	12.3–16.5	14.9 ± 0.5
LAIan	2.4–8.4	5.2 ± 1.0	–	–

[y] SM2: spikes m^{-2}; KPS, kernels spike^{-1}; TKW, thousand kernels weight; GY, grain yield; Chl, SPAD index; water soluble carbohydrates concentration (WSC) and content (WSCC); Δ^{13}C, isotopic discrimination of ^{13}C; LAI, leaf area index.

[z] Trait measurement at anthesis (an), grain filling (gf), or maturity (m).

Spectral Reflectance Indices

Coefficients of determination greater than 0.8 were only reached when the hydric conditions were combined (WS+FI) for GYm (AN: 0.82 and GF: 0.92) and Δ^{13}Cm (AN: 0.82 and GF: 0.92) (**Table 3**; Supplementary Table 2). Among the 255 SRIs tested, NWI-3 [(R$_{970}$−R$_{920}$)/(R$_{970}$+R$_{920}$)] worked at AN and WI2 (R$_{970}$/R$_{900}$) worked at GF. When the hydric conditions were kept separate, predictions with r^2 values greater than 0.25 were achieved only for WS conditions and spectral measurements performed at GF (NWI-3; GYm: 0.51 and Δ^{13}Cm: 0.26).

Coefficients of determination between 0.40 and 0.79 were found with combined hydric conditions for the following traits: SM2m [AN MTCI ((R$_{800}$−R$_{750}$)/(R$_{750}$−R$_{670}$)): 0.63; GF SAVI2 (1.5∗(R$_{807}$−R$_{736}$)/(R$_{807}$+R$_{736}$+0.5)): 0.66]; Chlan [AN TCARI2 (3∗((R$_{700}$−R$_{600}$)−((0.2∗(R$_{700}$−R$_{550}$))∗(R$_{700}$/(R$_{850}$+R$_{670}$)))): 0.59; GF MTCI: 0.60]; Chlgf [AN NWI-3: 0.42; GF NWI-3: 0.44]; WSCan [AN MTCI: 0.41; GF NDSI4 ((R$_{933}$−R$_{948}$)/(R$_{933}$+R$_{948}$)): 0.44]; and WSCCan (AN MTCI: 0.47; GF WI2: 0.50). When the hydric conditions were kept separate, r^2 values between 0.4 and 0.79 were found only for LAIan under FI conditions [AN DATT ((R$_{850}$−R$_{710}$)/(R$_{850}$−R$_{680}$): 0.44] (**Table 3**).

Multivariate Regression Methods

As observed for the SRIs, the four multivariate regression models (PCR, PLSR, RR, and SVR) showed the highest predictive power for most traits when data from both hydric conditions were combined (WS+FI), with the exception of TKWm at AN under FI conditions (**Table 3**, Supplementary Figure 1, and Supplementary Table 2). The R_{cv}^2 values were similar between RR and SVR, and greater than those for the other two multivariate models. Using RR or SVR, R_{cv}^2 values greater than 0.8 were found for GYm (AN: 0.90 and GF: 0.93) and Δ^{13}Cm (AN: 0.92 and GF:

0.94). In addition, R_{cv}^2 values between 0.40 and 0.79 were found for the following traits when using SVR with combined hydric conditions: SM2m (AN: 0.73 and GF: 0.74), Chlan (AN: 0.59 and GF: 0.66), Chlgf (AN: 0.44 and GF: 0.51), WSCan (AN: 0.48 and GF: 0.49), and WSCCan (AN: 0.53 and GF: 0.56). When the hydric conditions were kept separate, r^2 values in this same range were achieved only for GYm (GF: 0.56) under WS conditions, and for LAI (AN: 0.45) under FI conditions (**Table 3**).

Multivariate Classification Methods

The general performances of the categorical models were very similar in terms of model accuracy with 10xCV (**Figure 1A**); PCA-LDA, kNN, and PLS-DA showed average accuracies of 0.81, 0.76, and 0.71, respectively. PLS-DA, however, showed the lowest error rate of validation (**Figure 1B**); the average errors were approximately 0.42, 0.42, and 0.30 for PCA-LDA, kNN, and PLS-DA, respectively. Moreover, the error rate of validation showed differences within each model, being similar for the two hydric conditions when these were kept separate, but higher when the WS and FI conditions were combined.

The general performance of the three categorical models was evaluated based on the prediction rate of cross-validation for both classes (**Figure 2**), which proved to be different between models. PCA-LDA and kNN showed greater prediction levels for samples included in Class 1 (0.96 and 0.88, respectively), but very low prediction levels for samples from Class 2 (0.21 and 0.27, respectively) (**Figures 2A, B**). Meanwhile, PLS-DA showed lower prediction levels for Class 1, however, the prediction rates were similar for both Class 1 (∼0.70) and Class 2 (∼0.71) (**Figure 2C**).

Considering the prediction rates for both classes, the best genotype discriminations were obtained for most traits by PLS-DA when both hydric conditions were combined, with the exception of WSCm and WSCCm, at WS (AN and GF). Prediction rates from cross-validation by PLS-DA for Class 1 under WS+FI conditions ranged from 0.64 (KPSm at AN and GF, and WSCCm at AN) to 0.77 (SM2m at AN and GF), while prediction rates for Class 2 were between 0.52 (WSCm at AN) and 0.98 (GY at GF). However, the prediction rates for several traits were greater when the hydric conditions were kept separate, when compared to those rates achieved with combined WS+FI conditions. This was observed mainly for Class 1 (e.g., KPSm AN-FI, GYm GF-WS, and Δ^{13}Cm GF-WS) but also for Class 2 (e.g., TKWm AN-FI, WSCm GF-WS, and WSCCm AN-WS).

Importantly, unlike the SRI and multivariate regression methods, the PLS-DA model generated high prediction levels for the individual hydric conditions, in both classes, for most of the traits evaluated. Under FI conditions, the prediction rates for Class 1 ranged from 0.56 (Chlan at GF) to 0.85 (LAIan at AN), while the prediction rates for Class 2 were between 0.47 (KPSm at AN) and 0.82 (TKWm and LAIan at AN). Under WS conditions, the prediction rates for Class 1 were between 0.56 (WSCCan at GF) and 0.85 (GYm at GF), while those for Class 2 ranged from 0.43 (KPSm at AN) to 0.78 (GYm at GF). Prediction levels between 0.40 and 0.79 for both classes were found for most traits when the hydric conditions were combined, although some of these showed prediction levels greater than 0.80 in one of the two classes: Class 1 (SM2m AN-FI and GF-FI, GYm GF-WS, Δ^{13}Cm

TABLE 3 | Best spectral reflectance indices (SRI), regression and classification models calculated by trait, reflectance assessment and hydric condition.

Trait[u]	Reflectance assessment[w]	Hydric condition[x]	Spectral reflectance indices		Regression models					Classification models		
			SRI[y]			PCR	PLSR	RR	SVR	PCA-LDA	PLS-DA	kNN
				r²		R^2	R^2	R^2	R^2	Pred. Rate (Class 2)	Pred. rate (Class 2)	Pred. rate (Class 2)
SM2m[v]	AN	WS	SR (550;670)	0.15[z]		0.11	0.12	0.12	0.11	0.23	0.55	0.17
		FI	NRI (1510;660)	0.24		0.23	0.25	0.34	0.26	0.41	0.71	0.30
		WS+FI	MTCI (800;750;670)	0.63		0.39	0.59	0.73	0.73	0.57	0.87	0.50
	GF	WS	CI (415;695)	0.11		0.03	0.10	0.20	0.15	0.18	0.64	0.29
		FI	NDSI (403;830)	0.20		0.18	0.18	0.24	0.20	0.23	0.68	0.21
		WS+FI	SAVI (807;736)	0.66		0.59	0.68	0.75	0.74	0.51	0.90	0.46
KPSm	AN	WS	NDSI (543;548)	0.03		0.01	0.01	0.06	0.05	0.04	0.43	0.09
		FI	RBI (695;445)	0.10		0.03	0.07	0.13	0.16	0.06	0.47	0.12
		WS+FI	NDSI (543;548)	0.19		0.15	0.16	0.20	0.22	0.07	0.71	0.27
	GF	WS	NDSI (933;948)	0.13		0.05	0.07	0.13	0.14	0.07	0.59	0.22
		FI	SR (690;655)	0.10		0.03	0.07	0.07	0.05	0.03	0.59	0.26
		WS+FI	SR (960;950)	0.24		0.21	0.21	0.24	0.26	0.07	0.73	0.28
TKWm	AN	WS	SR (690;655)	0.13		0.06	0.08	0.08	0.07	0.05	0.69	0.16
		FI	NRI (1510;660)	0.29		0.30	0.30	0.30	0.30	0.21	0.82	0.35
		WS+FI	NRI (1510;660)	0.27		0.21	0.21	0.26	0.28	0.25	0.73	0.39
	GF	WS	WI (950;900)	0.10		0.02	0.06	0.21	0.27	0.24	0.69	0.30
		FI	NRI (1510;660)	0.15		0.16	0.17	0.20	0.19	0.26	0.74	0.30
		WS+FI	WI (950;900)	0.16		0.17	0.18	0.33	0.38	0.24	0.80	0.33
GYm	AN	WS	PRI * CI (570;530;760;700)	0.09		0.07	0.11	0.14	0.19	0.18	0.68	0.22
		FI	SR (440;685)	0.12		0.02	0.10	0.16	0.23	0.10	0.61	0.24
		WS+FI	NDWI (970;920)	0.82		0.63	0.68	0.89	0.90	0.54	0.93	0.44
	GF	WS	NDWI (970;920)	0.51		0.33	0.36	0.48	0.56	0.49	0.78	0.32
		FI	WI (900;970)	0.14		0.09	0.11	0.14	0.15	0.12	0.69	0.18
		WS+FI	WI (970;900)	0.92		0.83	0.85	0.92	0.93	0.53	0.98	0.51
Chlan	AN	WS	PRI (550;531)	0.09		0.00	0.09	0.09	0.10	0.06	0.65	0.25
		FI	MCARI (700;670;550)	0.18		0.07	0.17	0.21	0.20	0.03	0.63	0.16
		WS+FI	TCARI (700;600;550;850;670)	0.59		0.48	0.53	0.58	0.59	0.41	0.96	0.45
	GF	WS	AI (740;887;691;698)	0.14		0.01	0.12	0.26	0.25	0.20	0.63	0.17
		FI	PRI (512;531)	0.11		0.02	0.06	0.07	0.07	0.06	0.75	0.22
		WS+FI	MTCI (800;750;670)	0.60		0.51	0.55	0.61	0.66	0.41	0.97	0.41
Chlgf	AN	WS	NDSI (410;550)	0.04		0.00	0.04	0.07	0.05	0.10	0.71	0.23
		FI	MCARI (700;670;550)	0.21		0.07	0.14	0.14	0.11	0.14	0.80	0.27
		WS+FI	NDWI (970;920)	0.42		0.34	0.38	0.42	0.44	0.57	0.95	0.50

(Continued)

TABLE 3 | Continued

Trait[u]	Reflectance assessment[w]	Hydric condition[x]	Spectral reflectance indices SRI[y]	r²	Regression models PCR R²	PLSR R²	RR R²	SVR R²	Classification models PCA-LDA Pred. Rate (Class 2)	PLS-DA Pred. rate (Class 2)	kNN Pred. rate (Class 2)
	GF	WS	SR (690;655)	0.05	0.04	0.05	0.11	0.11	0.18	0.66	0.18
		FI	PRI (512;531)	0.14	0.04	0.10	0.13	0.12	0.16	0.71	0.17
		WS+FI	NDWI (970;920)	0.44	0.42	0.43	0.44	0.51	0.52	0.96	0.48
WSCan	AN	WS	NDSI (1060;1118)	0.05	0.02	0.05	0.09	0.10	0.03	0.53	0.26
		FI	SAVI (807;736)	0.12	0.10	0.13	0.19	0.11	0.20	0.63	0.16
		WS+FI	MTCI (800;750;670)	0.41	0.23	0.30	0.49	0.48	0.29	0.85	0.36
	GF	WS	RE (670;780)	0.04	0.02	0.03	0.07	0.06	0.03	0.59	0.19
		FI	SAVI 2 (800;670)	0.09	0.08	0.08	0.09	0.10	0.10	0.63	0.29
		WS+FI	NDSI (933;948)	0.44	0.37	0.41	0.46	0.49	0.22	0.87	0.31
WSCm	AN	WS	NDTI (1650;2215)	0.06	0.03	0.04	0.06	0.03	0.11	0.55	0.22
		FI	WDVI (830;660)	0.04	0.00	0.01	0.01	0.01	0.01	0.51	0.27
		WS+FI	NDTI (1650;2215)	0.03	0.01	0.02	0.03	0.03	0.04	0.52	0.26
	GF	WS	NDSI (940;1122)	0.07	0.03	0.03	0.07	0.06	0.15	0.59	0.17
		FI	NDSI (1060;1118)	0.03	0.00	0.02	0.02	0.03	0.06	0.50	0.22
		WS+FI	SR (960;950)	0.03	0.00	0.01	0.04	0.05	0.02	0.54	0.17
WSCCan	AN	WS	NRI (1510;660)	0.05	0.03	0.04	0.07	0.06	0.05	0.56	0.03
		FI	NRI (1510;660)	0.20	0.21	0.23	0.30	0.22	0.24	0.78	0.22
		WS+FI	MTCI (800;750;670)	0.47	0.29	0.43	0.55	0.53	0.28	0.92	0.42
	GF	WS	NDSI (442;438)	0.04	0.03	0.03	0.10	0.07	0.03	0.66	0.23
		FI	SAVI 2 (800;670)	0.15	0.15	0.15	0.18	0.16	0.18	0.70	0.31
		WS+FI	WI (970;900)	0.50	0.40	0.46	0.56	0.56	0.31	0.90	0.31
WSCCm	AN	WS	NDTI (1650;2215)	0.06	0.03	0.04	0.06	0.05	0.09	0.66	0.20
		FI	WDVI (830;660)	0.04	0.00	0.00	0.02	0.02	0.01	0.54	0.25
		WS+FI	NDTI (1650;2215)	0.02	0.02	0.02	0.03	0.03	0.05	0.61	0.28
	GF	WS	NDWI (870;1260)	0.12	0.08	0.09	0.13	0.08	0.13	0.61	0.14
		FI	BI (460;660)	0.03	0.00	0.01	0.02	0.02	0.05	0.56	0.23
		WS+FI	NDSI (503;483)	0.03	0.01	0.01	0.06	0.06	0.07	0.58	0.17
Δ13Cm	AN	WS	TCARI (700;670;550)	0.11	0.05	0.10	0.10	0.15	0.25	0.58	0.20
		FI	SIPI (800;440;680)	0.05	0.01	0.02	0.11	0.12	0.10	0.58	0.13
		WS+FI	NDWI (970;920)	0.82	0.63	0.68	0.91	0.92	0.40	0.89	0.41

(Continued)

TABLE 3 | Continued

Trait[u]	Reflectance assessment[w]	Hydric condition[x]	Spectral reflectance indices		Regression models				Classification models		
					PCR	PLSR	RR	SVR	PCA-LDA	PLS-DA	kNN
			SRI[y]	r²	R²	R²	R²	R²	Pred. Rate (Class 2)	Pred. rate (Class 2)	Pred. rate (Class 2)
	GF	WS	NDWI (970;920)	0.26	0.12	0.16	0.23	0.34	0.42	0.66	0.27
		FI	NDWI (970;850)	0.06	0.02	0.03	0.10	0.10	0.04	0.58	0.22
		WS+FI	WI (970;900)	0.92	0.81	0.85	0.93	0.94	0.43	0.95	0.40
LAIan	AN	FI	Datt (850;710;680)	0.44	0.36	0.41	0.37	0.45	0.63	0.82	0.54
	GF	FI	SAVI 2 (800;670)	0.32	0.26	0.33	0.36	0.37	0.46	0.78	0.34

[u]SM2, spikes m^{-2}; KPS, kernels spike^{-1}; TKW, thousand kernels weight; GY, grain yield; Chl, SPAD index; water soluble carbohydrates concentration (WSC) and content (WSCC); $\Delta^{13}C$, isotopic discrimination of ^{13}C; LAI, leaf area index.

[v]Trait measurement at anthesis (an), grain filling (gf), or maturity (m).

[w]Spectral reflectance measurement at anthesis (AN) and grain filling (GF).

[x]Hydric conditions were water stress (WS), full irrigated (FI), and the combination (WS+FI).

[y]Spectral Reflectance Index: List at Lobos and Poblete-Echeverría (2017).

[z]In bold values (R² or prediction rate) between 0.40 and 0.79 (red) and ≥0.80 (blue).
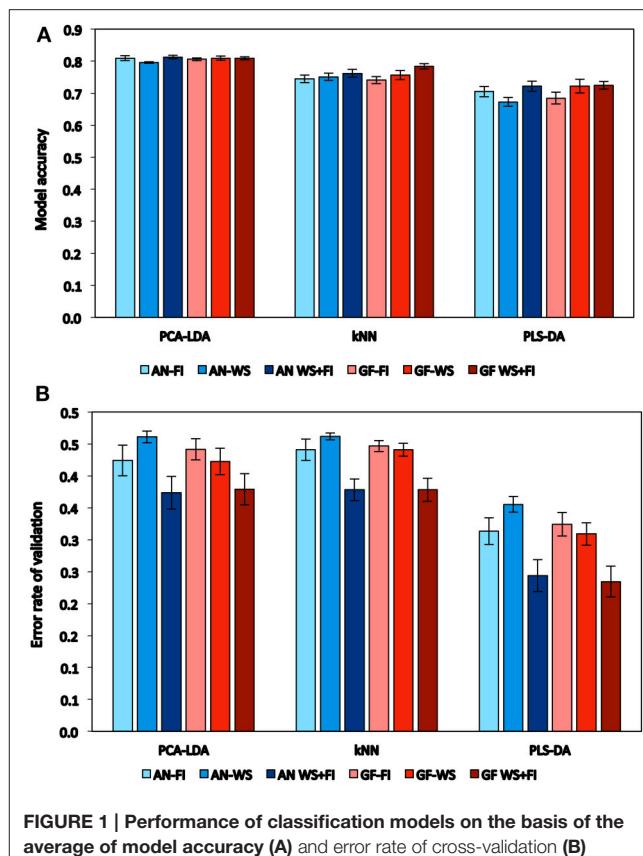
FIGURE 1 | Performance of classification models on the basis of the average of model accuracy (A) and error rate of cross-validation **(B)** calculated to all traits and estimated by spectral reflectance at anthesis (*AN*) and grain filling (*GF*). Wheat genotypes growing under two hydric conditions (FI: fully irrigated and WS: water stress); combination of both conditions (WS+FI) for modeling purposes. Vertical bars represent the standard error.

GF-WS, and LAI*an GF*-FI) and Class 2 (TKW*m AN*-FI and Chl*gf AN*-FI). Prediction levels greater than 0.80 for both classes under the individual hydric conditions were only achieved for LAI*an* with prediction rates of 0.85 and 0.82 for Class 1 and Class 2, respectively (**Table 3**; Supplementary Table 3).

DISCUSSION

Despite the present study being conducted in a single year, it generated an interesting database for testing approximation methodologies. This was due to the use of a large number of cultivars/advanced lines of wheat that were grown under two contrasting hydric conditions and evaluated for a large number of traits, with spectral reflectance assessed at two developmental stages (*AN* and *GF*).

Unlike other studies, this work covers a large proportion of the SRIs reported in the remote sensing literature (Lobos and Poblete-Echeverría, 2017). In general, the regression analysis between the traits and SRIs showed an important increase in predictive potential when the experimental data from both hydric conditions were combined. Nonetheless, when compared with previous reports (Lobos et al., 2014), lower coefficients of determination were found for GY*m* and $\Delta^{13}C$*m* when the individual environments were considered.

FIGURE 2 | Performance of each classification method (**A**: PCA-LDA, **B**: *k*NN, and **C**: PLS-DA) on the basis of the average of prediction rate of cross-validation calculated to all traits and classes, and estimated by spectral reflectance at anthesis (*AN*) and grain filling (*GF*). Wheat genotypes growing under two hydric conditions (FI: fully irrigated and WS: water stress); combination of both conditions (WS+FI) for modeling purposes. Vertical bars represent the standard error.

The developmental stage at which spectral reflectance was assessed influenced the relationship between the traits and the SRIs, which has been reported previously (Marti et al., 2007; Lobos et al., 2014). The best predicted traits, GY*m* and Δ^{13}C*m*, had a greater r^2 at *GF*, while TKW*m* and LAI*an* had a greater r^2 at *AN*; no major changes were observed for the other traits. The better prediction of GY*m* at *GF* could be related to the fact that the three main yield components (SM2, KPS, and TKW) are determined in the crop during this stage. In the case of Δ^{13}C*m*, both stomatal conductance and carboxylation rate influence the carbon isotope ratio (^{13}C/^{12}C) in kernels, and are affected by WS in Mediterranean environments at the *GF* stage (Condon et al., 2004).

Regarding to GY*m* and Δ^{13}C*m* assessed at *GF*, among the 255 SRIs tested, water indices were the ones explaining the highest variability on each environment: NWI-3 on WS and WI2 on WS+FI, while on FI WI and NWI-3 highlighted. Water indices, which combines near infrared spectra wavelengths (NIR), do not directly measure water content but instead detect the changes in leaf anatomy and cell structure that are induced by the state of hydration (Lobos et al., 2014), which influences the productivity of the plant. For GY*m*, similar results have been reported by other authors (Babar et al., 2006a; Prasad et al., 2007; Gutierrez et al., 2010; Lobos et al., 2014), while for Δ^{13}C*m*, this has been reported just once (Lobos et al., 2014). This highlights the effectiveness of water indices over vegetation indices. Traits such as SM2*m*, Chl*an*, Chl*gf*, WSC*an*, and WSCC*an* correlated better with vegetation and chlorophyll SRIs, which combine visible and NIR wavelengths, but also with water indices.

Although SRIs are easy to calculate, they are limited by the use of few wavelengths. Of the multivariate regression models studied (PCR, PLSR, RR, and SVR) (**Table 3**, Supplementary Figure 1, and Supplementary Table 2), PCR and PLSR generally performed the same, or worse than, the SRIs. Although PLSR, which is the most popular technique used in studies of this kind, has the ability to reduce the effect of the spectral signatures collinearity through a reduction in the dimensionality of the data (Hastie et al., 2005); our results show that SRIs may perform similarly, or better, when used in plant phenotyping. On the other hand, the RR and SVR models performed the same, or better than, the SRIs (e.g., GY*m* and Δ^{13}C*m* estimations increased by 8 and 10%, respectively, when using SVR). This highlights that there are multivariate regression analysis models other than PLSR that should be used in plant phenotyping to improve prediction statistics in plant breeding.

RR is a method of multivariate linear regression that includes a contraction of the multivariate model regression coefficients (Hastie et al., 2005). Although there are fewer reports of RR in remote sensing studies, this method performed better than PLSR for the estimation of plant biomass from satellite images (Cai et al., 2009) and was successfully used by Hernández et al. (2015) to predict GY in a large set of wheat genotypes. RR is known to be useful when the number of observations is much lower than the number of variables (James et al., 2013). In the present study, the number of observations is ∼800, while the number of variables (reflectance values space) is around 2,000. Furthermore, RR is recognized to be an effective prediction model when there is high collinearity in the data (James et al., 2013), which is typical of spectroscopy studies where a full spectrum of reflectance is used. Our results provide empirical evidence that RR performs well in the context of spectral reflectance data, and better than more popular methods such as PLSR.

Optimization of a SVR model does not depend on the dimensionality of the input space (Smola and Vapnik, 1997). Thus, it has the ability to handle complex non-linear dependencies in high-dimensional feature spaces (Smola and Vapnik, 1997; Hastie et al., 2005), such as those modeled in this study, where it performs better than PCR and PLSR. SVR has been the subject of several comparisons in spectral studies. Wang et al. (2011) achieved more accurate estimates of rice LAI when using LS-SVM (Least Squares Support Vector Machines) rather

than PLSR and MLR models. Similarly, better estimations of the nitrogen, phosphorus and potassium content in plant leaves were obtained by Zhai et al. (2013) when using SVR rather than PLSR. On the other hand, Yao et al. (2015) showed similar performance for PLSR and SVR in the prediction of nitrogen concentration in wheat leaves.

In addition to GYm and Δ^{13}Cm traits, improvements of between 6 and 22% were shown for the estimation of the traits SM2m, WSCan, and WSCCan, when assessed with spectral reflectance at AN, and SM2m, and TKWm, when assessed at GF. The best improvement in prediction was achieved for TKWm; this was 17 and 22% when using RR and SVR, respectively. Additionally, RR, SVR, and SRIs showed similar trends in their estimations of most traits when assessed with spectral data measured at AN or GF. The best predicted traits, GYm and Δ^{13}Cm, as well as SM2m, KPSm, Chlan, Chlgf, and WSCCan, were better-predicted using measurements of reflectance at GF. Meanwhile, LAIan was better-predicted with measurements of reflectance at AN, although this trait was evaluated only for well-watered plants.

The SRIs and multivariate regression models all performed much better when the data from both hydric conditions were combined (WS+FI). This situation was likely produced due to the increase in the number of samples but potentially the increase in the trait-range responses associated with the two contrasting environments was more important (**Table 2**). The increase in the trait-range by combining contrasting environments, and the effects on modeling improvement has been reported previously in wheat (Aparicio et al., 2002; Royo et al., 2003; Lobos et al., 2014). Because of this, special attention was paid to identification of elite genotypes (Class 2) in individual environments with a categorical approximation (PCA-LDA, PLS-DA, and kNN). Although the quality of a model is usually shown by its accuracy and error rate, the main difference in model performance was found to be the model's ability to identify samples from either individual (WS or FI) or combined (WS+FI) environments as Class 2, with PLS-DA being shown to be the stronger methodology (**Figure 2C**, **Table 3**). There are currently a few reports regarding the use of reflectance data and categorical methods for cereal breeding purposes (Delwiche et al., 1999, 2006, 2011; Lavine et al., 2014; Porker et al., 2017); however, all of these studies were carried out using the reflectance information from kernels or ground meal.

The selection of wheat genotypes suitable for water deficit-prone environments has traditionally been based on grain yield under irrigation conditions (yield potential) and under water deficit conditions (Araus et al., 2008; Cattivelli et al., 2008; Araus and Cairns, 2014). Even though both selected sites in this study where relatively close (70 km apart), the environmental conditions were different. Clearly the environment where the plants grew influenced the phenotype of each genotype, and therefore the spectral signature of a given genotype at each site. This could explain, in part, the differences in the estimation of each character; traits such as SM2m, TKWm, Chlgf, WSCan, and WSCCan showed higher predictions at FI, while GYm, WSCm, WSCCm, and Δ^{13}Cm were better estimated at WS.

The aim of this study was to assess the feasibility of using a categorical approximation to select featured genotypes, by comparing the performance of a large set of SRIs, multivariate regression models, and categorical models in the prediction of several traits using plant reflectance data. Even though information from only 1-year was considered, the data set used for modeling was large enough to determine the potential of each approach in plant breeding programs. Unfortunately, there are no previous studies contemplating the number of genotypes, SRIs, or the regression/categorical methods covered in this article.

Although data from additional studies and a greater number of years are needed, the present results suggest that future works oriented at plant breeding should focus on identification of elite genotypes in preference to predicting specific trait-values. The assessment of agricultural and physiological traits, such as those examined in this study, could contribute to the improvement of plant breeding programs and accelerate the selection and release of wheat genotypes/cultivars with greater adaptation to adverse environmental conditions.

CONCLUSIONS AND FUTURE PERSPECTIVES

Field measurement of canopy spectral reflectance is an efficient and fast way to collect plant status information for a large number of genotypes simultaneously. Analysis of reflectance data, gathered from different hydric conditions and developmental stages, by SRIs, multivariate regression, and categorical models, allows for the prediction of agricultural and physiological traits that are related to wheat yield and water deficit adaptation. The categorical model PLS-DA proved to be a useful tool for identifying elite genotypes grown under FI or WS conditions, improving upon genotype selection based on SRIs and multivariate regression methods.

Although GY and some of the other traits evaluated in this study were predicted using SRI, multivariate regression, and categorical models, there remains a need for assessing other secondary traits that have yet to be explored in plant breeding programs.

To improve trait prediction, it will be crucial to consider other tools, such as machine learning approaches (e.g., random forest or tree-based neural networks), or include other variables that are usually assessed by remote sensing (e.g., plant temperature).

AUTHOR CONTRIBUTIONS

AdP and IM designed the experiments, selected the germplasm, and participated on field evaluations. SR-B, FE, and AE contributed to experimental measurements and data collection. GL and AE contributed to the development of a tool for spectral analysis. CA was in charge of the implementation of modeling tools. MG and SR-B contributed to data analysis and development of the models. GL and MG were in charge of the writing up but all the authors contributed to the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2017.00280/full#supplementary-material

REFERENCES

Aparicio, N., Villegas, D., Araus, J. L., Casadesús, J., and Royo, C. (2002). Relationship between growth traits and spectral vegetation indices in durum wheat. *Crop Sci.* 42, 1547–1555. doi: 10.2135/cropsci2002.1547

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Araus, J. L., Slafer, G. A., Royo, C., and Serret, M. D. (2008). Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant Sci.* 27, 377–412. doi: 10.1080/07352680802467736

Ayeneh, A., Van Ginkel, M., Reynolds, M. P., and Ammar, K. (2002). Comparison of leaf, spike, peduncle and canopy temperature depression in wheat under heat stress. *Field Crop. Res.* 79, 173–184. doi: 10.1016/S0378-4290(02)00138-7

Azimi, M., Alizadeh, H., Salekdeh, G. H., and Naghavi, M. (2010). Association of yield and flag leaf photosynthesis among wheat recombinant inbred lines (RILs) under drought condition. *J. Food Agric. Environ.* 8, 861–863.

Babar, M. A., Reynolds, M. P., van Ginkel, M., Klatt, A. R., Raun, W. R., and Stone, M. L. (2006a). Spectral reflectance indices as a potential indirect selection criteria for wheat yield under irrigation. *Crop Sci.* 46, 578–588. doi: 10.2135/cropsci2005.0059

Babar, M. A., Van Ginkel, M., Klatt, A. R., Prasad, B., and Reynolds, M. P. (2006b). The potential of using spectral reflectance indices to estimate yield in wheat grown under reduced irrigation. *Euphytica* 150, 155–172. doi: 10.1007/s10681-006-9104-9

Ballabio, D., and Consonni, V. (2013). Classification tools in chemistry. Part 1: linear models. PLS-DA. *Anal. Methods* 5, 3790–3798. doi: 10.1039/c3ay40582f

Barker, M., and Rayens, W. (2003). Partial least squares for discrimination. *J. Chemometr.* 17, 166–173. doi: 10.1002/cem.785

Bowman, B. C., Chen, J., Zhang, J., Wheeler, J., Wang, Y., Zhao, W., et al. (2015). Evaluating grain yield in spring wheat with canopy spectral reflectance. *Crop Sci.* 55, 1881–1890. doi: 10.2135/cropsci2014.08.0533

Cabrera-Bosquet, L., Crossa, J., von Zitzewitz, J., Serret, M. D., and Araus, J. L. (2012). High-throughput phenotyping and genomic selection: the frontiers of crop breeding converge. *J. Integr. Plant Biol.* 54, 312–320. doi: 10.1111/j.1744-7909.2012.01116.x

Cai, T., Ju, C., and Yang, X. (2009). Comparison of ridge regression and partial least squares regression for estimating above-ground biomass with land sat images and terrain data in mu us sandy land, China. *Arid Land Res. Manag.* 23, 248–261. doi: 10.1080/15324980903038701

Camargo, A., and Lobos, G. A. (2016). Latin America: a development pole for phenomics. *Front. Plant Sci.* 7:1729. doi: 10.3389/fpls.2016.01729

Cattivelli, L., Rizza, F., Badeck, F. W., Mazzucotelli, E., Mastrangelo, A. M., Francia, E., et al. (2008). Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. *Field Crop Res.* 105, 1–14. doi: 10.1016/j.fcr.2007.07.004

Condon, A. G., Richards, R. A., Rebetzke, G. J., and Farquhar, G. D. (2004). Breeding for high water-use efficiency. *J. Exp. Bot.* 55, 2447–2460. doi: 10.1093/jxb/erh277

Crain, J. L., Wei, Y., Barker, J., Thompson, S. M., Alderman, P. D., Reynolds, M., et al. (2016). Development and deployment of a portable field phenotyping platform. *Crop Sci.* 56, 1–11. doi: 10.2135/cropsci2015.05.0290

Cunningham, P., and Delany, S. J. (2007). *k-Nearest Neighbor Classifiers.* Technical report UCD-CSI-2007-4, Technical report UCD-CSI-2007-4. School of Computer Science and Informatics. University College Dublin, Dublin.

del Pozo, A., Yáñez, A., Matus, I., Tapia, G., Castillo, D., Araus, J. L., et al. (2016). Physiological traits associated with wheat yield potential and performance under water-stress in a Mediterranean environment. *Front. Plant Sci.* 7:987. doi: 10.3389/fpls.2016.00987

Delwiche, S. R., Graybosch, R. A., Amand, P. S., and Bai, G. (2011). Starch waxiness in hexaploid wheat (*Triticum aestivum* L.) by NIR reflectance spectroscopy. *J. Agric. Food Chem.* 59, 4002–4008. doi: 10.1021/jf104528x

Delwiche, S. R., Graybosch, R. A., and Peterson, C. J. (1999). Identification of wheat lines possessing the 1AL. 1RS or 1BL. 1RS wheat-rye translocation by near-infrared reflectance spectroscopy. *Cereal Chem.* 76, 255–260. doi: 10.1094/CCHEM.1999.76.2.255

Delwiche, S. R., Graybosch, R. A., Hansen, L. E., Souza, E., and Dowell, F. E. (2006). Single kernel near-infrared analysis of tetraploid (durum) wheat for classification of the waxy condition. *Cereal Chem.* 83, 287–292. doi: 10.1094/CC-83-0287

Dixon, J., Braun, H. J., and Crouch, J. (2009). "Overview: Transitioning wheat research to serve the future needs of the developing world," in *Wheat Facts and Futures*, eds J. Dixon, H. J. Braun, P. Kosina, and J. Crouch (Mexico, DF: CIMMYT), 1–19.

Dreccer, M. F., Barnes, L. R., and Meder, R. (2014). Quantitative dynamics of stem water soluble carbohydrates in wheat can be monitored in the field using hyperspectral reflectance. *Field Crops Res.* 159, 70–80. doi: 10.1016/j.fcr.2014.01.001

FAO, IFAD, and WFP. (2015). *The State of Food Insecurity in the World: Meeting the 2015 International Hunger Targets: Taking Stock of Uneven Progress.* Rome.

FAOSTAT (2013). *Food and Agriculture Organization of the United Nations. Statistics Division.* Available online at: http://faostat3.fao.org/

Farquhar, G. D., Ehleringer, J. R., and Hubick, K. T. (1989). Carbon isotope discrimination and photosynthesis. *Annu. Rev. Plant Phys.* 40, 503–537. doi: 10.1146/annurev.pp.40.060189.002443

Gutierrez, M., Reynolds, M. P., Raun, W. R., Stone, M. L., and Klatt, A. R. (2010). Spectral water indices for assessing yield in elite bread wheat genotypes under well-irrigated, water-stressed, and high-temperature conditions. *Crop Sci.* 50, 197–214. doi: 10.2135/cropsci2009.07.0381

Hastie, T., Tibshirani, R., Friedman, J., and Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *Math. Intelligencer* 27, 83–85. doi: 10.1007/BF02985802

Hernández, J., Lobos, G. A., Matus, I., del Pozo, A., Silva, P., and Galleguillos, M. (2015). Using ridge regression models to estimate grain yield from field spectral data in bread wheat (*Triticum aestivum* L.) grown under three water regimes. *Remote Sens.* 7, 2109–2126. doi: 10.3390/rs70202109

Hernández-Barrera, S., Rodríguez-Puebla, C., and Challinor, A. J. (2016). Effects of diurnal temperature range and drought on wheat yield in Spain. *Theor. Appl. Climatol.* doi: 10.1007/s00704-016-1779-9. [Epub ahead of print].

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R.* New York, NY: Springer.

Jolliffe, I. (2002). *Principal Component Analysis.* New York, NY: Springer-Verlag.

Jones, R. W., Reinot, T., Frei, U. K., Tseng, Y., Lübberstedt, T., and McClelland, J. F. (2012). Selection of haploid maize kernels from hybrid kernels for

plant breeding using near-infrared spectroscopy and SIMCA analysis. *Appl. Spectrosc.* 66, 447–450. doi: 10.1366/11-06426

Lavine, B. K., Mirjankar, N., and Delwiche, S. (2014). Classification of the waxy condition of durum wheat by near infrared reflectance spectroscopy using wavelets and a genetic algorithm. *Microchem. J.* 117, 178–182. doi: 10.1016/j.microc.2014.06.030

Lehmann, J. R. K., Große-Stoltenberg, A., Römer, M., and Oldeland, J. (2015). Field spectroscopy in the vnir-swir region to discriminate between mediterranean native plants and exotic-invasive shrubs based on leaf tannin content. *Remote Sens.* 7, 1225–1241. doi: 10.3390/rs70201225

Li, F., Mistele, B., Hu, Y., Chen, X., and Schmidhalter, U. (2014). Reflectance estimation of canopy nitrogen content in winter wheat using optimised hyperspectral spectral indices and partial least squares regression. *Eur. J. Agron.* 52, 198–209. doi: 10.1016/j.eja.2013.09.006

Li, X., Zhang, Y., Bao, Y., Luo, J., Jin, X., Xu, X., et al. (2014). Exploring the best hyperspectral features for LAI estimation using partial least squares regression. *Remote Sens.* 6, 6221–6241. doi: 10.3390/rs6076221

Lobos, G. A., and Hancock, J. F. (2015). Breeding blueberries for a changing global environment: a review. *Front. Plant Sci.* 6:782. doi: 10.3389/fpls.2015.00782

Lobos, G. A., and Poblete-Echeverría, C. (2017). Spectral Knowledge (SK-UTALCA): software for exploratory analysis of high-resolution spectral reflectance data. *Front. Plant Sci.* 7:1996. doi: 10.3389/fpls.2016.01996

Lobos, G. A., Matus, I., Rodriguez, A., Romero-Bravo, S., Araus, J. L., and del Pozo, A. (2014). Wheat genotypic variability in grain yield and carbon isotope discrimination under Mediterranean conditions assessed by spectral reflectance. *J. Integr. Plant Biol.* 56, 470–479. doi: 10.1111/jipb.12114

Marti, J., Bort, J., Slafer, G. A., and Araus, J. L. (2007). Can wheat yield be assessed by early measurements of Normalized Difference Vegetation Index? *Ann. Appl. Biol.* 150, 253–257. doi: 10.1111/j.1744-7348.2007.00126.x

Pimstein, A., Karnieli, A., Bansal, S. K., and Bonfil, D. J. (2011). Exploring remotely sensed technologies for monitoring wheat potassium and phosphorus using field spectroscopy. *Field Crops Res.* 121, 125–135. doi: 10.1016/j.fcr.2010.12.001

Porker, K., Zerner, M., and Cozzolino, D. (2017). Classification and authentication of barley (*Hordeum vulgare*) malt varieties: combining attenuated total reflectance mid-infrared spectroscopy with chemometrics. *Food Anal. Methods* 10, 675–682. doi: 10.1007/s12161-016-0627-y

Prasad, B., Carver, B. F., Stone, M. L., Babar, M. A., Raun, W. R., and Klatt, A. R. (2007). Potential use of spectral reflectance indices as a selection tool for grain yield in winter wheat under great plains conditions. *Crop Sci.* 47, 1426–1440. doi: 10.2135/cropsci2006.07.0492

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Rebetzke, G. J., Rattey, A. R., Farquhar, G. D., Richards, R. A., and Condon, A. T. G. (2012). Genomic regions for canopy temperature and their genetic association with stomatal conductance and grain yield in wheat. *Funct. Plant Biol.* 40, 14–33. doi: 10.1071/FP12184

Royo, C., Aparicio, N., Villegas, D., Casadesus, J., Monneveux, P., and Araus, J. L. (2003). Usefulness of spectral reflectance indices as durum wheat yield predictors under contrasting Mediterranean conditions. *Int. J. Remote Sens.* 24, 4403–4419. doi: 10.1080/0143116031000150059

Shewry, P. R. (2009). Wheat. *J. Exp. Bot.* 60, 1537–1553. doi: 10.1093/jxb/erp058

Siegmann, B., and Jarmer, T. (2015). Comparison of different regression models and validation techniques for the assessment of wheat leaf area index from hyperspectral data. *Int. J. Remote Sens.* 36, 4519–4534. doi: 10.1080/01431161.2015.1084438

Smola, A., and Vapnik, V. (1997). Support vector regression machines. *Adv. Neur. In.* 9, 155–161.

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proc. Natl Acad. Sci.U.S.A.* 108, 20260. doi: 10.1073/pnas.1116437108

Velu, G., and Prakash, R. (2013). "Phenotyping in wheat breeding," in *Phenotyping for Plant Breeding,* eds S. K. Panguluri and A. A. Kumar (New York, NY: Springer), 41–71.

Wang, F. M., Huang, J. F., and Lou, Z. H. (2011). A comparison of three methods for estimating leaf area index of paddy rice from optimal hyperspectral bands. *Precis. Agric.* 12, 439–447. doi: 10.1007/s11119-010-9185-2

Willmott, C. J. (1981). On the validation of models. *Phys. Geogr.* 2, 184–194.

Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometr. Intell. Lab.* 58, 109–130. doi: 10.1016/S0169-7439(01)00155-1

Yao, X., Huang, Y., Shang, G., Zhou, C., Cheng, T., Tian, Y., et al. (2015). Evaluation of six algorithms to monitor wheat leaf nitrogen concentration. *Remote Sens.* 7, 14939–14966. doi: 10.3390/rs71114939

Yemm, E. W., and Willis, A. J. (1954). The estimation of carbohydrates in plant extracts by anthrone. *Biochem. J.* 57, 508–514. doi: 10.1042/bj0570508

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zhai, Y., Cui, L., Zhou, X., Gao, Y., Fei, T., and Gao, W. (2013). Estimation of nitrogen, phosphorus, and potassium contents in the leaves of different plants using laboratory-based visible and near-infrared reflectance spectroscopy: comparison of partial least-square regression and support vector machine regression methods. *Int. J. Remote Sens.* 34, 2502–2518. doi: 10.1080/01431161.2012.746484

Precision Automation of Cell Type Classification and Sub-Cellular Fluorescence Quantification from Laser Scanning Confocal Images

Hardy C. Hall[1]*[†], Azadeh Fakhrzadeh[2], Cris L. Luengo Hendriks[2] and Urs Fischer[1]

[1] Department of Forest Genetics and Plant Physiology, Umeå Plant Science Centre, Swedish University of Agricultural Sciences, Umeå, Sweden, [2] Centre for Image Analysis, Uppsala University, Uppsala, Sweden

While novel whole-plant phenotyping technologies have been successfully implemented into functional genomics and breeding programs, the potential of automated phenotyping with cellular resolution is largely unexploited. Laser scanning confocal microscopy has the potential to close this gap by providing spatially highly resolved images containing anatomic as well as chemical information on a subcellular basis. However, in the absence of automated methods, the assessment of the spatial patterns and abundance of fluorescent markers with subcellular resolution is still largely qualitative and time-consuming. Recent advances in image acquisition and analysis, coupled with improvements in microprocessor performance, have brought such automated methods within reach, so that information from thousands of cells per image for hundreds of images may be derived in an experimentally convenient time-frame. Here, we present a MATLAB-based analytical pipeline to (1) segment radial plant organs into individual cells, (2) classify cells into cell type categories based upon Random Forest classification, (3) divide each cell into sub-regions, and (4) quantify fluorescence intensity to a subcellular degree of precision for a separate fluorescence channel. In this research advance, we demonstrate the precision of this analytical process for the relatively complex tissues of Arabidopsis hypocotyls at various stages of development. High speed and robustness make our approach suitable for phenotyping of large collections of stem-like material and other tissue types.

Keywords: automated image analysis, confocal microscopy, Arabidopsis, hypocotyl, automated phenotyping, code:matlab

INTRODUCTION

Rapid and cheap sequencing technologies have dramatically changed plant breeding and functional genomics in the last decade. Availability of abundant genotyping data shifts the focus within the frame of genetic screening from more efficient genotyping to automated phenotyping technologies. Progress has been made on whole-plant phenotyping solutions, which for example record plant growth, photosynthesis rates, or stress markers (Furbank and Tester, 2011; Dhondt et al., 2013). Whole-plant phenotyping solutions have become commercially available for indoors and outdoors use and are now an integral part of

numerous large breeding programs (Cobb et al., 2013; Rahaman et al., 2015). While such whole-plant phenotyping technologies are useful to facilitate breeding for higher yield many important qualitative and developmental traits cannot be assessed by these macroscopic approaches. Especially, genetic screens for chemical composition, anatomical, and mechanical properties of plant raw materials still rely on laborious low-throughput manual phenotyping. A wide range of molecular markers enable the spatially highly resolved study of such traits. Many of these markers can be fluorescently imaged, either through their own inherent fluorescence, via fluorescent fusion proteins, stains, probes, or through immunofluorescence. With the wide range of fluorescence tools, Laser Scanning Confocal Microscopy (LSCM) has become the method of choice to localize and quantify fluorescent markers. LSCM imaging provides fast, sensitive, inexpensive, spatially highly resolved images where pixel intensity reflects target abundance over a wide dynamic range. Fluorescent imaging of morphogen gradients in the Drosophila embryo and of auxin transport proteins in the Arabidopsis shoot and root tip have, for example, greatly contributed to our understanding of pattern formation and development (Gregor et al., 2007; Kierzkowski et al., 2013). Nevertheless, many of these studies rely on comparison of fluorescence intensity between manually defined regions of interest (ROI; Nilufar and Perkins, 2014), e.g., between different cell types. Obviously, manual segmentation into ROIs is labor-intensive and underlies human subjectivity and inconsistency that may severely limit the interpretability of LSCM data.

Computer-assisted quantification of fluorescent targets on a cellular scale over large spatial ranges requires both accurate automatic segmentation and quantification of fluorescence in each individual segment (Luengo Hendriks et al., 2006). Multiple fluorescence sources, including a counterstain for segmentation, can be imaged simultaneously within a single field of view enabling the correlation of the segmented image with a host of other fluorescence targets. Whereas, for animal tissues the application of automated imaging analysis has become popular in the last decade and several software packages for such an approach are freely available (Wiesmann et al., 2015), the adaptation of automated image analysis is lagging behind in plants. This may be related to the limited optical transparency of most plant tissues and therefore a need of thin sectioning of plant specimens resulting in low throughput. In animal tissues the predominant strategy to segment tissues automatically makes use of fluorescently labeled nuclei and region growing algorithms, such an approach may fall short when a fluorescent target is localized to the plasma membrane or the cell wall. In plants, the few attempts to automatically segment tissue have made use of the cell wall stain propidium iodide and the plasma membrane marker FM4-64 (Federici et al., 2012; Pound et al., 2012; Band et al., 2014; Yoshida et al., 2014). These live stains are suitable for embryonic and meristematic tissue but not for mature plant tissues, as the Arabidopsis hypocotyl and stem, which contain terminally differentiated, dead cells with disrupted plasma membranes and which are in comparison to animal tissues of limited optical penetration depth. As an alternative, Sankar et al. (2014) suggested a protocol based on

non-fluorescent differential interference contrast images (DIC). However, applicability of this approach is limited by insufficient accuracy, extensively long computing times and incompatibility with confocal imaging.

In hypocotyl and stem, emerging models for wood formation and stem cell research in plants (Jouannet et al., 2015), derivatives of stem cells differentiate into several different cell types of the xylem (inner tissue) and phloem (outer tissue). New divisions of stem cells push daughter cells either toward the in- or out-side, and, with increasing distance from the stem cells, derivatives gradually differentiate. A morphogen-like gradient of the plant hormone auxin has been suggested to regulate stem cell activity and differentiation from cell expansion to cell wall thickening (Uggla et al., 1996; Bhalerao and Fischer, 2014). Changes in morphology and wall composition are indicative for the degree of differentiation and cell type (Liebsch et al., 2014). Compositional changes in the walls of stems have been successfully monitored with the help of monoclonal antibodies against specific wall epitopes (Hall et al., 2013). However, exploitation of such data is currently hampered by manual segmentation, classification and quantification of fluorescent signals. As a consequence, genetic improvement of woody feedstock, e.g., decreased lignin content in xylem fibers, is limited by the absence of automated high-throughput phenotyping tools.

Here, we provide an image analysis pipeline, which (i) accurately segments hypocotyls and stems into individual cells and subcellular regions, (ii) assigns each segment to a cell type, (iii) quantifies fluorescence intensity of the cell wall counterstain and, from a separate channel, quantifies several aspects of fluorescence intensity from cell wall epitopes for each individual segment and cell type, and also (iv) extracts a wealth of morphometric data. The pipeline is coded in a single software environment (MATLAB) and the data can easily be exported and, for example, be used in modeling or multivariate statistics. Short processing times permit large data sets, as required for mutant screens or association mapping, to be analyzed.

MATERIALS AND METHODS

Preparation of Plant Material

Wild-type (*Col-0*) and *knat1^{bp−9}* seeds were planted on soil with 18:6 h (light:dark) at 21 °C. Germination times were recorded and hypocotyls excised from plants at 21 and 31 days after germination (dag). Hypocotyls were identified as the 5 mm region below the cotyledons. The 5 mm hypocotyls were immersed in 150 μL 1X PME (stock 2X PME: 50 mM PIPES, 2 mM MgSO$_4$, 2 mM EGTA) fixation buffer, within 0.2 ml dome-cap thermal cycler tubes (Thermo Scientific, www.thermoscientificbio.com). Hypocotyls were then subjected to three consecutive 21°C cycles of 5 min vacuum infiltration at 68 kPa, and washed three times in 1X PME (21°C, 68 kPa) prior to storage at 4°C in 1X PME. Segments were individually encased in 1 cm^3 blocks of 5% agar at 65°C, and stored at 4°C to set. Transverse sections (40 μm thick) were cut from segments using a VT100S vibrating microtome (Leica), separated from agar encasement using a sable hair ("00") brush, then

blocked for at least 1 h in 5% bovine serum albumin in 1X TBST (10 mM Tris, 0.25 M NaCl, 0.1% Tween). Sections were mixed to randomize developmental difference, and randomly allocated from each biological replicate pool, together with 100 µl fresh blocking solution, to wells of a 96-well plate (Ibidi, www.bdbiosciences.com). Blocking solutions were swapped with 5 µl 1:36 dilutions of the LM10 antibody (Complex Carbohydrate Research Center, University of Georgia, US) using gel- loading tips, then sections were incubated at 4°C for 16 h. Hypocotyls were washed twice in 100 µL 1X TBST, then incubated for 1 h at 21°C in the dark in 10 µl of 2 µg/µl Alexa FluorTM 488 donkey anti-rat IgG (H + L) (Agrisera, Sweden). Sections were again washed twice in 40 µL 1X TBST prior to counter- staining with 0.015% Calcfluor White (Sigma-Aldrich, www.sigmaaldrich.com). Sections were again washed twice in 100 µL 1X TBST to remove excess counter-stain and unbound secondary antibody.

Confocal Imaging

Hypocotyl sections were imaged using a confocal laser scanning microscope Zeiss LSM780 point-scan system at 1024 × 1024 pixels (pixel size, 0.6–0.83 µm) with a 10X objective (a plan-apochromat objective with a numerical aperture of 0.45) within the 96-well plate (Ibidi, Germany) fitted with 180 µm-thick coverslip bottoms. Immunofluorescence of AlexaFluor 568 was excited with a 561 nm laser, and emitted light filtered at 575–600 nm. Calcfluor White was subsequently scanned on an independent channel with a 405 nm laser and emission observed at 420–430 nm. Images were saved as "LSM" files with file names that included plate location, antibody, genotype, tissue type, and biological replicate separated by underscores (ex. C08R6_LM10_Col_21-day-old_Hyp_BR1) to permit automated cataloging in the supplied MATLAB analysis pipeline.

Image Analysis–General

The following methods describe the analytical steps taken, and do not serve as an operation manual for processing images. Instead, refer to Supplemental Presentation 2 ("Precision Cell Classification and Quantification Manual") and Supplemental Video 1 in Presentation 1 (https://vimeo.com/148871821, password: Matlab4Segment) for details on system configuration, experimental setup, parameter optimization and data processing. The image analysis pipeline was implemented in MATLAB using the MATLAB Image Processing toolbox and the DIPimage toolbox (http://www.diplib.org/).

Experimental Setup for Image Analysis

A working title for the experiment was entered to automatically generate a time-stamped folder to deposit analysis output (Supplemental Figure 5A). The target image files for the experiment that existed within a user-defined source folder were automatically cataloged within the database "ExperInfo" (Supplemental Figure 5). This database recorded file location and levels for experimental factors (age, tissue type, genotype) for each image file.

Training Set Generation

Three to four training set images occupying a separate folder were selected that presented the range of morphological variation expected to be encountered in the experimental (testing) image set. These images were imported into MATLAB, and users prompted to enter parameters for image smoothing and segmentation (Supplemental Table 1). After data smoothing of the CFW channel (Supplemental Figure 5), tissue centers were manually selected via a user interface (Supplemental Figure 5), triggering the segmentation algorithm to generate ROIC (entire cell), ROIL (cell lumen), and ROIW (cell wall) for all objects. Prior to delineate the cell borders by Watershed segmentation, Gaussian filtering with a variance of 1 pixel was applied in order to remove background noise. Oversegmentation was corrected by merging regions where the difference in intensity between their minima and the first pixel on the watershed dam touching the two regions was <10. Lumen boundaries within each watershed region were precisely identified by applying Otsu's thresholding (Otsu, 1979). In some cases, we cropped the image to restrict the amount of tissue or the range of cell types to be examined. This restricted the ROICs to those that fell within the cropped region, and also restricted computation of ROIL and ROIW to those within the cropped region (Supplemental Figure 1). Measurements for all ROIs were saved for later access by the classification algorithm. The number of cell types and their names were subsequently defined, then selected within each training image (Supplemental Figure 5E). A graphical output of the selections was recorded (example in **Figure 2**) along with a MAT-file containing the locations and dimensions of those ROIs. In an iterative process (Supplemental Figures 5F–G), features (**Table 1**) and cell classes were selected and Random Forest classification (Breiman, 2001) executed for the chosen training set (cell selections). This model was then used to generate class predictions and confidence interval scores for all ROIs in the training set images. An overlay of the entire classification on the original cell selections was generated for each training image (data not shown), along with classification result at varying confidence interval thresholds (as in **Figures 6D–F**). Each training set iteration was outputted to a distinct time-stamped folder for subsequent evaluation. The "ExperInfo" database was updated to include a record of all iterations, their parameters, and the locations of the model data sufficient for classification of test images. Quantitative data for the CFW channel were saved but not further utilized in the experiment. Immunofluorescence quantitation was not performed in the training set generation.

Quantification of Test Images

Test images that were suitable for classification by a common training set iteration were processed in a similar manner as the training set (smoothing, defining tissue centers, segmenting) based upon parameters defined for the chosen training set (Supplemental Figures 5H–I). In an iterative approach, classification of the testing set was explored with different training set iterations (as in **Figure 3**). Quantification data and diagnostic plotting of the CFW channel similar to the training set iterations were stored in time-stamped folders. To

TABLE 1 | Features available in Random Forest classification analysis pipeline.

Feature[a]	Description[b]
m.cx	X coordinate of cell center, origin top-left corner of image (μm)
m.cy	Y coordinate of cell center, origin top-left corner of image (μm)
Xnew	X coordinate of cell center, origin center of image (μm)
Ynew	Y coordinate of cell center, origin center of image (μm)
radialV	Radial coordinate of cell center, origin center of image (μm)
angleV	Angular coordinate of cell center, origin center of image (radian)
m.majoraxes	Length of major axis (first principal component axis) of cell (μm)
m.eccenticity	Square root of [1-(Length of minor axis (second principal component axis)^2/(Length of major axis)^2)]
m.theta	The angle between major axis and horizontal axis (radian)
s.area	Number of cell pixels (μm^2)
perimeter	Length of cell perimeter (μm)
s.radius.mean	Average value of radius of cell (distance from object border to the center of object) (μm)
s.radius.min	Minimum value radius of cell (μm)
s.radius.max	Maximum value radius of cell (μm)
extv	Multiplication of Length of major and minor axes (μm^2)
inclV	Acute angle between radial vector of cell (originating form center of image to center of cell) and major axis of cell (radian)
P2A	Circularity of the cell (ratio of Perimeter to Area)
MedianROIC	Median of ROIC intensity
MeanROIC	Average of ROIC intensity
m.theta.real	Angle between radial vector of cell (originating form center of image to center of cell) and first (major) principal component of cell (radian)
MedianROIW	Median of ROIC intensity
MeanROIW	Average of ROIC intensity

[a]Feature name as it appears in the diagnostic plotting.
[b]Brief description of the measurement, including appropriate units.

quantify the signal intensities attributable to the immunolabeling, immunofluorescence channel images were then segmented using ROIC, ROIL, and ROIW generated from the CFW channel segmentation which acted as a mask. In addition, the ROIs were divided into four quadrants (**Figure 1E**) for higher resolution-based quantification of signal as detailed in **Table 2**. These data were exported in to separate time-stamped folders for later access. Importantly, the locations of MAT-files containing quantification data for each image were stored in the "ExperInfo" database for access during data assembly and export.

Data Review and Assembly

The final step in the pipeline is to assemble the data for export based upon information stored in "ExperInfo" regarding which images were processed, and where the associated data is stored. From a user prompt, we selected the levels of each factor (ex. specific antibodies from the "antibody" factor), and files with those properties were concatenated into a common file ("DataRawCompile") to be used in downstream (multivariate) analysis in MATLAB or another environment. We assembled a pipeline that facilitates iterative summary plotting of spatial maps of features for any image present in the assembled data set. "DataRawCompile" was then used to generate means and standard deviations for each cell class within each image (as in **Figures 5, 6**). The pipeline also permits iterative comparative plotting of the summary statistics (bar plots) for

any combination of images present in the output data set. The structure of the output file "DataRawCompile" is detailed in Supplemental File 2.

RESULTS

Automated Image Segmentation of Confocal Counterstain Channel

Embryonic and meristematic plant tissues have been successfully segmented with the help of propidium iodide. However, since propidium iodide is not retained in the cell wall of dead cells, this fluorescent stain is not suitable as a counterstain for segmentation of mature or fixed plant tissues. As an alternative to propidium iodide, we tested calcofluor white (CFW), which binds to cellulose and chitin in cell walls of plants, fungi and bacteria. In order to visualize boundaries between cells we counterstained cell walls of 21-day-old Arabidopsis hypocotyls with CFW and acquired images with a CLSM in a separate reference channel (Channel 1). CFW fluorescence was restricted to cell walls and not detected in cell lumen. After smoothing the reference channel, the watershed segmentation algorithm identified the outer boundaries of cells. Watershed dams matched cell-cell boundaries closely with little over-segmentation, and the ROIs for entire cells, denoted ROIC, matched morphology closely (**Figures 1A,B**). The lumen boundary within each watershed region was precisely identified using Otsu's threshold algorithm
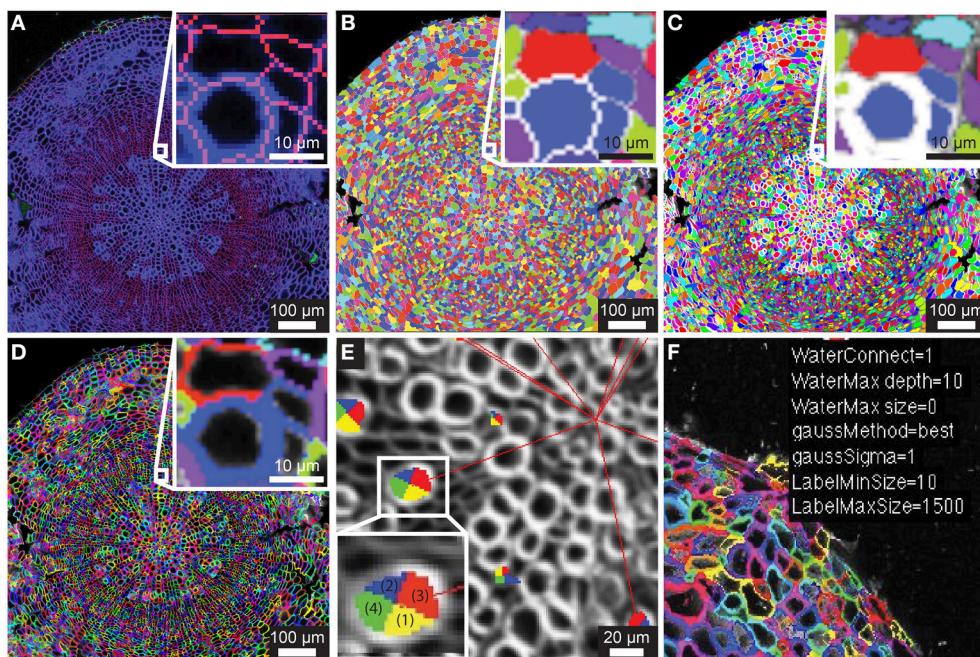
FIGURE 1 | Segmentation of counterstain channel. (A) Two-channel confocal image of transverse section with watershed boundaries overlaid for entire image and xylem to cambium transition (inset) **(B)** Image segmentation into ROIC (entire cell) objects with random color assignment **(C)** Segmentation of the lumen of each cell (ROIL) with zoomed region depicted in **(A**; inset). **(D)** Cell wall regions (ROIW) derived from difference of ROIC and ROIL with zoomed region depicted in **(A)**. **(E)** Example of quadrant definition for ROIC, with anticlinal quadrants "1" and "2" (yellow and blue) and periclinal quadrants "3" and "4" (green and red) for ROIL being defined by radial axis (red lines to tissue origin). **(F)** Sample of overlay of smoothing and segmentation parameters for diagnostic purposes.

to define ROIL (**Figure 1C**). Cell wall regions, referred to as ROIW, could be derived as the set difference of ROIC and ROIL (**Figures 1D,F**), thus giving us two distinct regions of the cell (wall/ROIW and lumen/ROIL). With the help of a manually selected center of tissue, ROIC, L, and W were subdivided into interior/exterior and left/right (lateral) quadrants in order to study polar or axial distribution of fluorescent signals (**Figure 1E**).

Manual Selection of Training Set Cells for Classification Model

Supervised learning algorithms have been shown successful for classifying image segments corresponding to individual cells into cell type categories (Field et al., 2010; Sankar et al., 2014). These algorithms require the creation of training sets, i.e., in our case manually assigned classification of cells into groups of user-defined cell types. ROICs of reference images provide sufficient context for manual classification of representative cells for each of the cell types. We defined six different cell types in 21-day-old hypocotyls: xylem vessels and parenchyma, cells of the cambial zone, phloem fibers, phloem, and cortical cells (**Figure 2A**). For each class, cells were chosen that best represented the class, avoiding selection of cells that lay on vague boundaries between cell types or exhibited morphology that were intermediate between cell classes.

During development, organ types such as the Arabidopsis hypocotyl undergo substantial change in tissue composition

(additional cell types) and architecture (e.g., cell morphology and relative position in the tissue context). With Arabidopsis hypocotyls, there is added complexity in 31-day-old hypocotyls with the addition of new cell types in the outer xylem (xylem II), demanding another, optimized training. For 31-day-old hypocotyls, we defined eight different cell types that included xylem I vessels, xylem I parenchyma, xylem II vessels, xylem II fibers, cambium, phloem fibers, phloem parenchyma, and cortex (**Figure 2B**). These selections were used as the basis for subsequent trials to determine the combination of cell types and features to use for classification.

Feature Computation and Pre-Selection

Given accurate identification of cell boundaries (ROIC), lumen (ROIL), and cell wall (ROIW), abundant, precise morphometric data can be derived that collectively provide a rich set of features to classify segments into different cell type categories. Using the DIPimage package, we measured 22 different features that could be derived from ROIs, covering aspects of position within the tissue, shape, and fluorescence intensity for each segment (**Table 1**). Many of these ROI measures correlate well with tissue morphology while some of these features, such as "m.cx" and "m.cy" (xy-coordinates), provide non-sense information with respect to cell identity in a radially symmetrical organ (Supplemental Figures 2A,B). Feature selection is generally considered an important step to improving the accuracy of classification (Janecek et al., 2008). With the exception of

TABLE 2 | Fluorescence measures available for quantification in analysis pipeline.

Measure[a]	Association[b]	Description[c]
Mean/radial(1)	Each ROI (C,L, or W)	Mean signal of radial region "1" (see **Figure 1E**), computed for each of ROI (C,L, or W)
Mean/radial(2)	Each ROI (C,L, or W)	Mean signal of radial region "2" (see **Figure 1E**), computed for each of ROI (C,L, or W)
Mean/tangentrail(3)	Each ROI (C,L, or W)	Mean signal of tangential region "3" (see **Figure 1E**), computed for each of ROI (C,L, or W)
Mean/tangentrail(4)	Each ROI (C,L, or W)	Mean signal of tangential region "4" (see **Figure 1E**), computed for each of ROI (C,L, or W)
Std/radial(1)	Each ROI (C,L, or W)	Standard deviation signal of radial region "1" (see **Figure 1E**), computed for each of ROI (C,L, or W)
Std/radial(2)	Each ROI (C,L, or W)	Standard deviation signal of radial region "2" (see **Figure 1E**), computed for each of ROI (C,L, or W)
Std/tangentrail(3)	Each ROI (C,L, or W)	Standard deviation signal of tangential region "3" (see **Figure 1E**), computed for each of ROI (C,L, or W)
Std/tangentrail(4)	Each ROI (C,L, or W)	Standard deviation signal of tangential region "4" (see **Figure 1E**), computed for each of ROI (C,L, or W)
Size/radial(1)	Each ROI (C,L, or W)	Area signal of radial region "1" (see **Figure 1E**), computed for each of ROI (C,L, or W); a morphological measure
Size/radial(2)'	Each ROI (C,L, or W)	Area signal of radial region "2" (see **Figure 1E**), computed for each of ROI (C,L, or W); a morphological measure
Size/tangentrail(3)	Each ROI (C,L, or W)	Area signal of tangential region "3" (see **Figure 1E**), computed for each of ROI (C,L, or W); a morphological measure
Size/tangentrail(4)	Each ROI (C,L, or W)	Area signal of tangential region "4" (see **Figure 1E**), computed for each of ROI (C,L, or W); a morphological measure
LumenRPA	Derived, ROIL	Ratio of periclinal to anticlinal ROIL
Lumensignal	Derived, ROIL	Total signal per ROIL
PvD	Derived, ROIL	Punctateness vs. diffuseness of ROIL
WallRPA	Derived, ROIW	Ratio of periclinal to anticlinal ROIW
Wallsignal	Derived, ROIW	Total signal per ROIW
CellRPA	Derived, ROIC	Ratio of periclinal to anticlinal ROIC
Cellsignal	Derived, RO1C	Total signal per ROIC
LumenRPAmean	Derived, ROIL	Ratio of periclinal to anticlinal ROIL
Lumensignalmean	Derived, ROIL	Total signal per ROIL
WallRPAmean'	Derived, ROIW	Ratio of periclinal to anticlinal ROIW
Wallsignalmean	Derived, ROIW	Total signal per ROIW
CellRPAmean	Derived, ROIW	Ratio of periclinal to anticlinal ROIC

[a]*Name of fluorescence measure as it appears in diagnostic plotting.*
[b]*Regions of interest (ROIs) to which the measure applies. Those marked as "derived" are computed from other measures generated from the ROIs.*
[c]*Description of how each measure is computed.*

excluding "m.cx" and "m.cy" measures, we lacked *a priori* evidence to eliminate other variables without first examining their importance to successful classification. It was therefore necessary to take an iterative approach of feature selection, based upon the output of the classification, to arrive at an optimal set of features.

Classification

We then chose to compare two different supervised learning algorithms: Support Vector Machine (SVM), originally designed for binary classification problems, and Random Forest, developed specifically for multiclass problems. We tested the accuracy of the classification outputs employing all the above-mentioned features. Random Forest outperformed SVM using normalized measures, distance-scaled measures, and untransformed measures (Supplemental Figure 3). Interestingly, the Random Forest model with the untransformed data resulted in the best fit. We therefore focused on Random Forest for the optimization of the classification procedure.

As a first step of optimization we assessed the impact of removing features on the classification result, using 21-day-old hypocotyls as a guide. In the first case we admitted the 18 features into the model, all except the Cartesian coordinates ("m.cx," "m.cy," "Xnew," and "Ynew"). The Random Forest model yielded

rank scores of the importance of these features (**Figure 3A**), indicating that the radial displacement from the center of the tissue ("radialV") was the most discriminate feature underlying the radial organization of the tissue types. Other features that contributed substantially to the discrimination between the different cell types were median fluorescence intensity of ROIC and ROIW ("medianROIC," "medianROIW") and the size of the luminal area ("s.area"). The incline angle ("inclV"), which was used as a discriminating feature by Sankar et al. (2014), played a minor role. We used spatial mapping of features (Supplemental Figure 2) to remove six features that we considered redundant with others. Again, "radialV" was dominant, followed by cell wall and cell intensity ("medianROIW," "medianROIC," respectively; **Figure 3A**). Finally, we reduced the selection to five features that were ranked highest in the 12-feature set. Again, "radial" was dominant, while rankings for the remaining features remained similar to those in the 12-feature set (**Figure 3A**).

The Random Forest algorithm, as with other classification methods, classifies all objects. This invariably results in misclassified objects. However, the Random Forest model assigns a "confidence interval score" to each object such that misclassifications can be largely avoided through filtering. We tested the performance of confidence filtering at 50, 70, and 90% confidence by examining misclassification in cells that were
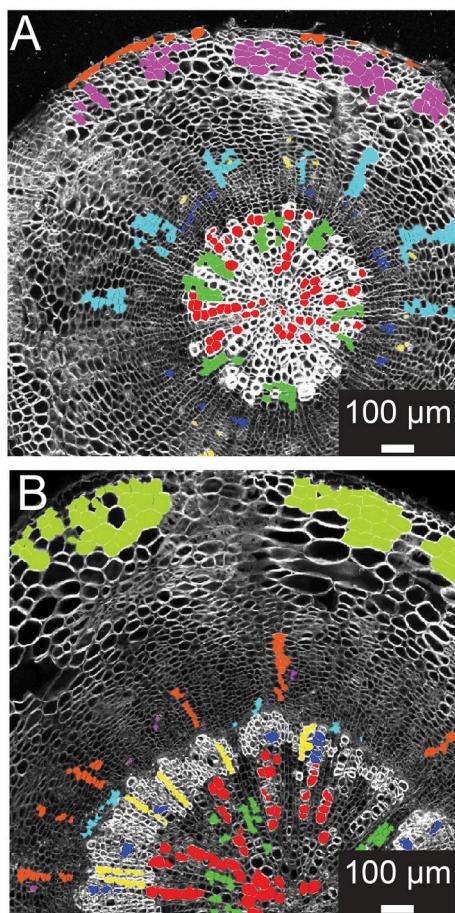
FIGURE 2 | Training set selection from ROIC segmentation result. (A) Representative 21-day-old wild-type hypocotyl tissues showing selections for xylem I vessels (red), xylem I parenchyma (green), cambium (navy blue), phloem fibers (yellow), phloem parenchyma (light blue), cork (purple), and epidermis (orange). **(B)** Representative 31-day-old wild-type hypocotyl tissues showing xylem I vessels (red), xylem I parenchyma (green), xylem II vessels (navy blue), xylem II fibers (yellow), cambium (light blue), phloem fibers (purple), phloem parenchyma (orange), and cortex (olive).

color-coded according to class in an overlay of the original reference channel, considering 18-, 12-, and 5-feature selection sets (**Figures 3B–J**). It is evident from these panels that increased confidence filtering reduces selection of cells in boundaries of differing cell types such as the cork and phloem parenchyma. The incidence of misclassifications is diminished by confidence filtering where cell types are interspersed, such as with xylem vessels and xylem parenchyma. Conversely, removing low-ranked and seemingly redundant features can lead to increased misclassification (inset of **Figures 3B–J**).

Training Set Versatility

A common scenario in developmental biology is the need to survey (to "phenotype") many genotypes. Automated quantitative morphometrics and fluorescence channel screening offer a means to circumvent the logistic bottleneck in quantifying

traits from microscopic tissues. Yet it is not efficient to develop distinct testing sets for each genotype (as with a screen of a mutagenized population). To examine the potential of using a common training set for genotypes with greatly differing tissue organization and morphometric characteristics, we chose to carry out a reciprocal examination between wild type (*Col-0*) and the $knat1^{bp-9}$ mutant which exhibits irregular radial organization of tissues in the hypocotyl (e.g., reduced xylem fiber formation) and altered luminal areas of xylem vessels (Liebsch et al., 2014). In this examination, we produced 12-feature training set models for each, and then compared result for each genotype (**Figure 4**). For the wild-type test image (**Figures 4A,B**), the $knat1^{bp-9}$ training set classified the vast majority of cortical, phloem parenchyma, phloem fiber, xylem I fiber and vessel cells correctly. However, at the boundary between xylem I and xylem II, relatively more misclassifications or absent classification (low confidence) of vessels and fibers occurred in the wild-type tissue with the $knat1^{bp-9}$ training set. Similarly, the wild-type training model on $knat1^{bp-9}$ performed well on the classification of all cell types except for xylem fibers and vessels at the boundary between xylem I and II relative to the $knat1^{bp-9}$ training set (**Figures 4C,D**). Generally, with the wild-type training set the boundary between xylem I and II is moved toward the outside of $knat1^{bp-9}$ hypocotyls (Liebsch et al., 2014) and misclassification at the boundary most likely represent the dominant nature of "radius" in the classification model. Yet, such misclassifications compose a small proportion of the classified cells that present a wealth of information for comparative morphometrics and fluorescence quantification.

In order to see if automated classification could be used as the basis to conduct comparative morphometrics of genotypes, we chose to examine cell area ("s.area") of xylem vessels of wild type (*Col-0*) and the $knat1^{bp-9}$ mutant, as $knat1^{bp-9}$ is known to have smaller cell areas (Liebsch et al., 2014). In this case, three separate sections from the same specimen were quantified, filtered on 70% confidence interval, and the means of all remaining regions of "s.area" measures computed with standard deviations (**Figure 5**). In line with previously published data (Liebsch et al., 2014), Xylem-II vessels of $knat1^{bp-9}$ were significantly smaller than in wild type providing evidence that multi-genotype morphometric comparisons are feasible.

Fluorescence Quantification

With robust, accurate classification and subdivision of ROIs to predefined classes, each ROI (and sub-ROI) provides a mask to conduct a variety of measurements of fluorescence (**Table 2**). While sub-region-specific ROIs provide a high resolution of fluorescence distribution, relative or summative distribution of these intensities can be more informative. As a result, "derived" measures are computed from values in sub-regions (1–4) of various ROIs. For the purpose of demonstration, we probed tissues with an antibody specific to xylan in the secondary cell wall. A fluorescent secondary antibody permitted visualization of epitope localization. Fluorescence was quantified by computing pixel-wise intensities of each ROI, then "derived" measures were subsequently plotted as spatial maps (Supplemental Figure 4). In 21-day-old hypocotyls,
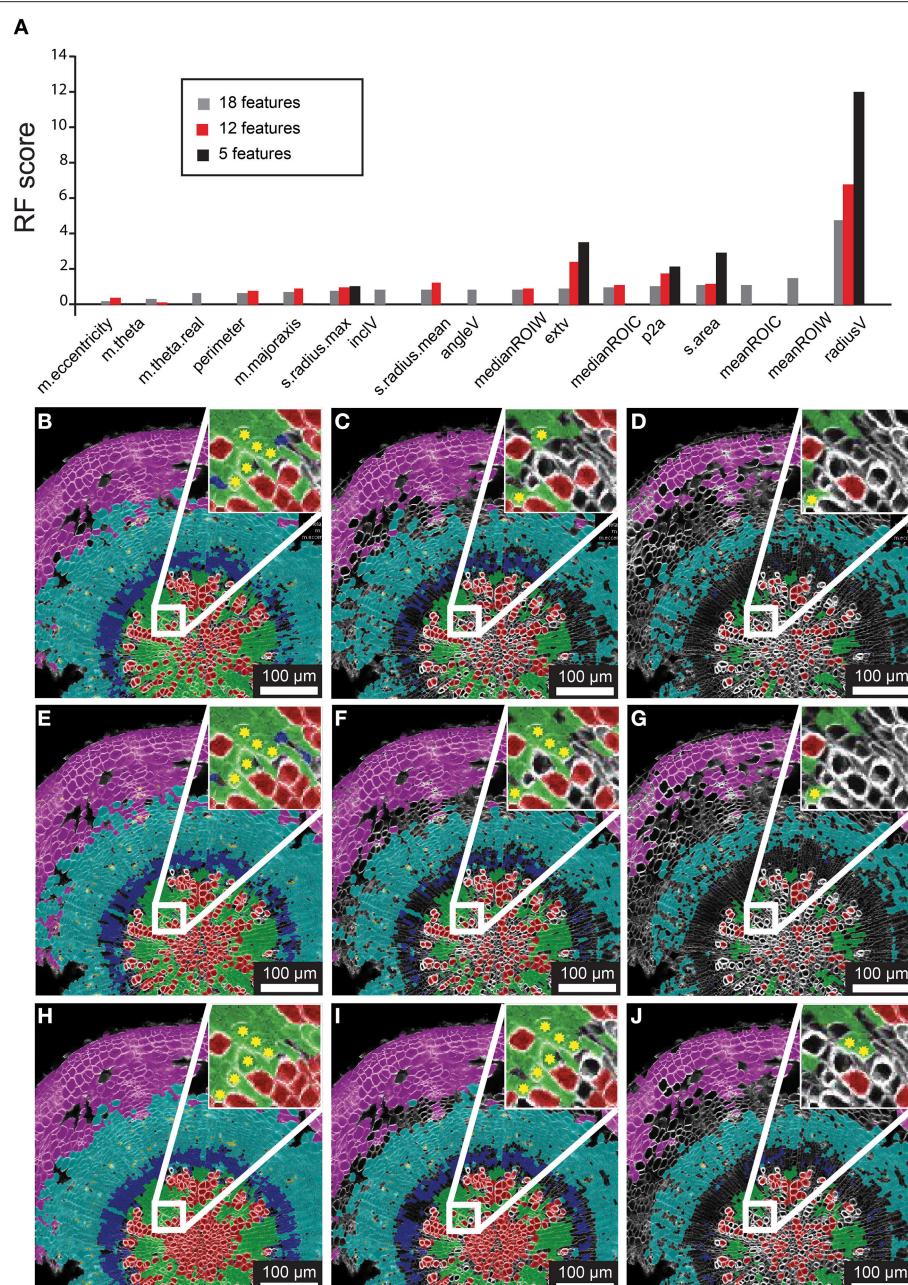
FIGURE 3 | Feature selection effects on classification for representative 21 dag wild-type hypocotyl. (A) Random Forest scores for features chosen in 18-, 12-, and 5-feature classification iterations. **(B–D)** Classification results for 18-feature classification with 50, 70, and 90% confidence interval filtering, respectively. **(E–G)** Classification results for 12-feature classification with filtering as in **(B–D)**. **(H–J)** Classification result for 5-feature classification with filtering as in **(B–D)**, and **(E–G)** sets. Inset for **(B–J)** depict a sub-region of tissue where misclassifications (asterisks) of xylem-I vessel elements occurs.

secondary cell walls occur exclusively in the xylem vessels, thereby providing a clear case of cell type-specific fluorescence to validate the quantification methodology (**Figure 6**). Comparison of fluorescence channel (**Figure 6B**) with a spatial heatmap of fluorescence intensity ("derived_wallsignal") mapped to ROIC (**Figure 6C**) demonstrates that the quantification replicates the spatial distribution of fluorescence in the source image. As evidence that the signal is predominantly in the walls, the "derived_lumensignal" values are only moderate in the outermost (youngest) vessels and completely absent in older vessels (Supplemental Figure 4). This is likely due to presence of epitope within the lumen, as it is defined by the segmentation process, of living xylem vessel cells during wall assembly.
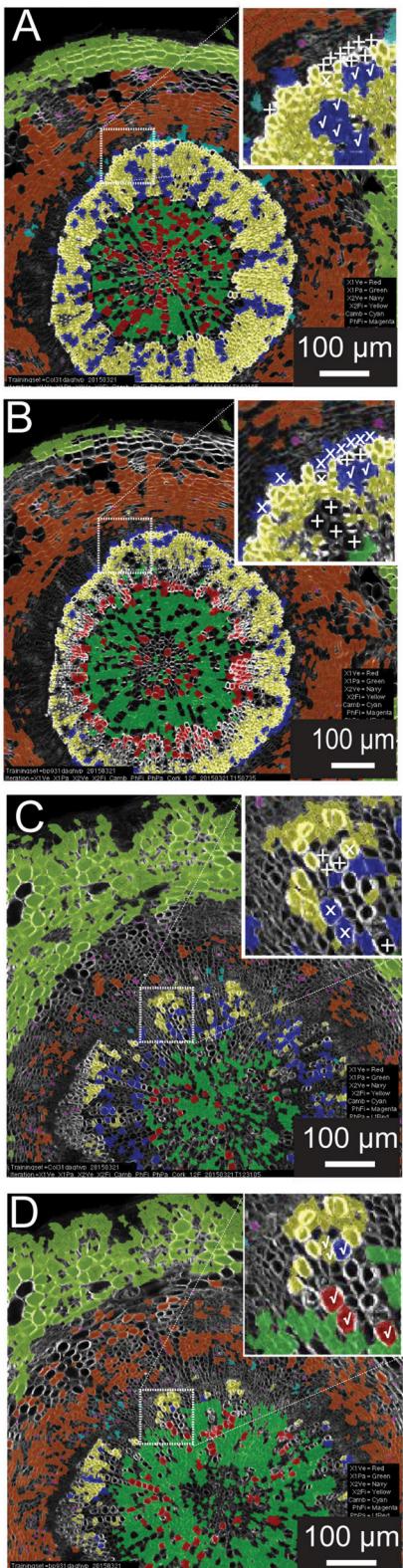
FIGURE 4 | Reciprocal classification of genotypes from training set iterations of either genotype. Thirty-one-day-old hypocotyls of wild-type

(Continued)

FIGURE 4 | Continued

(A,B) and knat1^{bp-9} (C,D) genotypes classified with training sets of either knat1^{bp-9} (B,D) and wild-type (A,C) 31-day-old hypocotyls. (A–D) Classifications passing the 70% confidence interval threshold. Insets in wild type (A,B) and knat1^{bp-9} (C,D) are arbitrarily chosen regions at the boundary between Xylem I and II that demonstrate the effect of selection of training set iteration on classification result (classification of selected cells of clear identity are indicated as correct [checkmarks], not classified ["+" sign], and misclassified ["x"]).
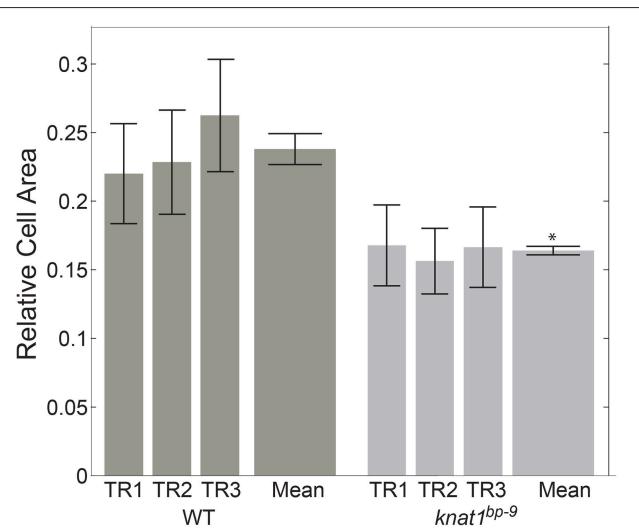
FIGURE 5 | Quantitative morphological difference between wild-type and knat1^{bp-9} genotypes. Relative sizes of secondary cell-walled xylem II vessels ("X-II-Ve") for technical replicates (separate immunolabeled sections) of representative 31-day-old hypocotyls of wild type and knat1^{bp-9}, with 12-feature classification iteration. Genotype-specific training sets were applied for classifications. Error bars represent standard deviations of relative areas of all cells passing the 70% confidence interval filtering. *t-test, wild type vs. mutant, $p < 0.05$.

As proof-of-concept that the classification provides a meaningful basis to group cells for fluorescence characterization, we next examined the means of fluorescence intensities ("derived_wallsignal") values for all cells of each cell class, filtering at 50, 70, and 90% (**Figures 6D–G**, respectively). From this series, it is clear that the xylem vessels are the dominant cell type that exhibits a fluorescence signal. Not evident with visual examination of the spatial map of "derived_wallsignal" (**Figure 6C**), the xylem parenchyma exhibits a weak signal, too. As increasing stringency reduces the contribution of the xylem parenchyma to the overall signal, we presume that misclassification is the primary cause of signal bleed into that cell type. Yet, stringent filtering (**Figure 6G**) does not eliminate this signal entirely and thus parenchymal cells of the xylem still retain signal where we expect none (no xylan in this cell type). One explanation for the signal is that there is inaccuracy in establishing the watershed boundaries between cell types that differ greatly in cell wall thickness (**Figure 1A**, inset).
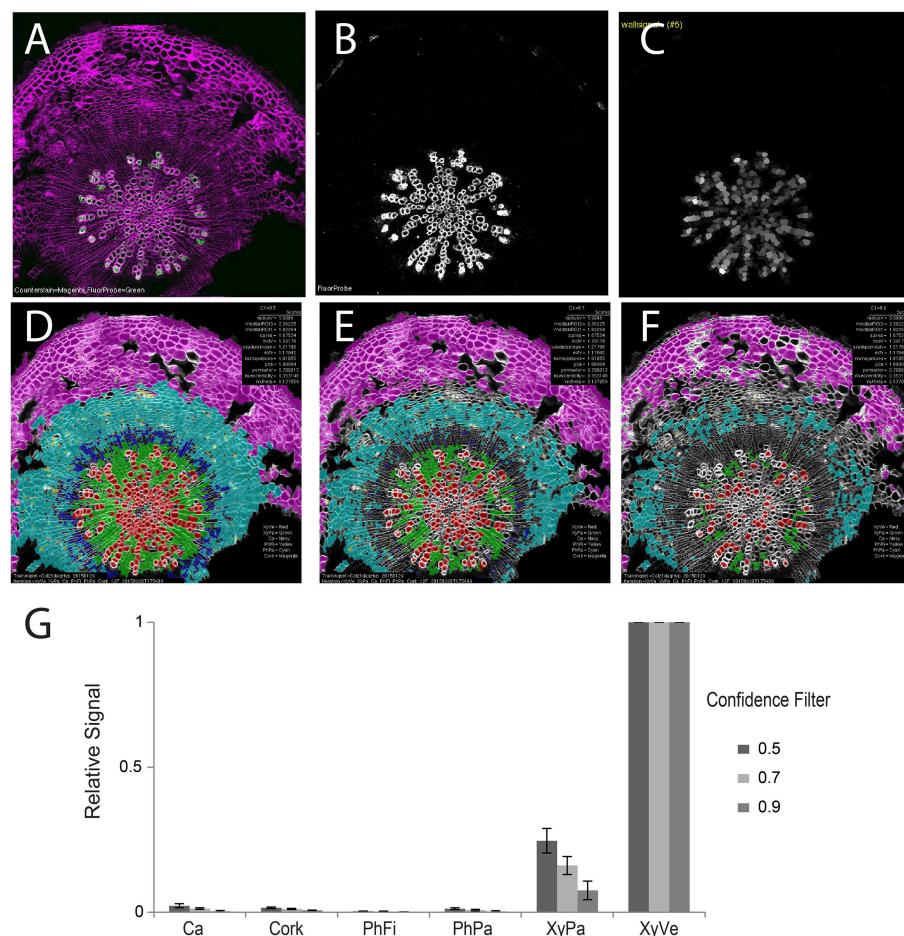
FIGURE 6 | Quantification of second channel (fluorescence) for a representative tissue (21-day-old wild-type hypocotyl) labeled with xylan-specific antibody (LM10). (A) Overlay image of CFW counterstain channel (magenta) and LM10 immunolabel (green) channel. **(B)** Isolated immunolabel channel after background fluorescence correction. **(C)** Heatmap of wall signal (grayscale; white, strong signal; black, no signal) for ROIL objects. Classification result and relative wall signal for 50 **(D)**, 70 **(E)**, and 90% **(F)** confidence filtering. **(G)** Quantification of relative fluorescence intensity. Error bars represent the standard deviations of means of relative signal intensities. Three biological replicates.

High Throughput Data Processing Package

This methodological proposal presents the reader with a MATLAB-based data analysis pipeline (Supplemental File 1) that provides the complete set of scripts necessary to prepare quantitative data from a complete fluorescence imaging experiment for downstream (multivariate statistical) analysis, while also providing a wealth of images and plots of diagnostic value (**Figure 7**, detailed in Supplemental Figure 5). Importantly, this set of scripts employs the MATLAB Image Processing toolbox as well as the DIPimage toolbox (http://www.diplib.org/) together providing flexibility and efficiency in a variety of standard image analysis procedures. However, it is the unique assemblage of novel custom scripts using the Random Forest classification model that provide the core analytical steps for generating quantitative data from raw images of plant tissue. Importantly, the core scripts are nested within a graphical user interface that minimizes command line interaction and prompts the user with all of the important parameters for

image processing, segmentation, classification and filtering (Supplemental Table 1).

DISCUSSION

Here, we provide an image analysis toolkit to accurately segment all the cells in transverse sections of hypocotyls, to precisely classify the individual segments into 6 to 8 cell type classes with a specified degree of certainty (confidence interval scores) and to extract precise morphometric data and fluorescence intensity of two different channels for each cell type. By analyzing the distribution of fluorescence emitted by a xylem-vessel-specific marker we obtained accurate segmentation and classification of xylem vessels and parenchymatic cells. This is a relevant improvement over previous attempts of automated cell type classification, which could not distinguish between xylem vessels from parenchyma (Sankar et al., 2014; Montenegro-Johnson et al., 2015). These recent methods relied heavily on spatial
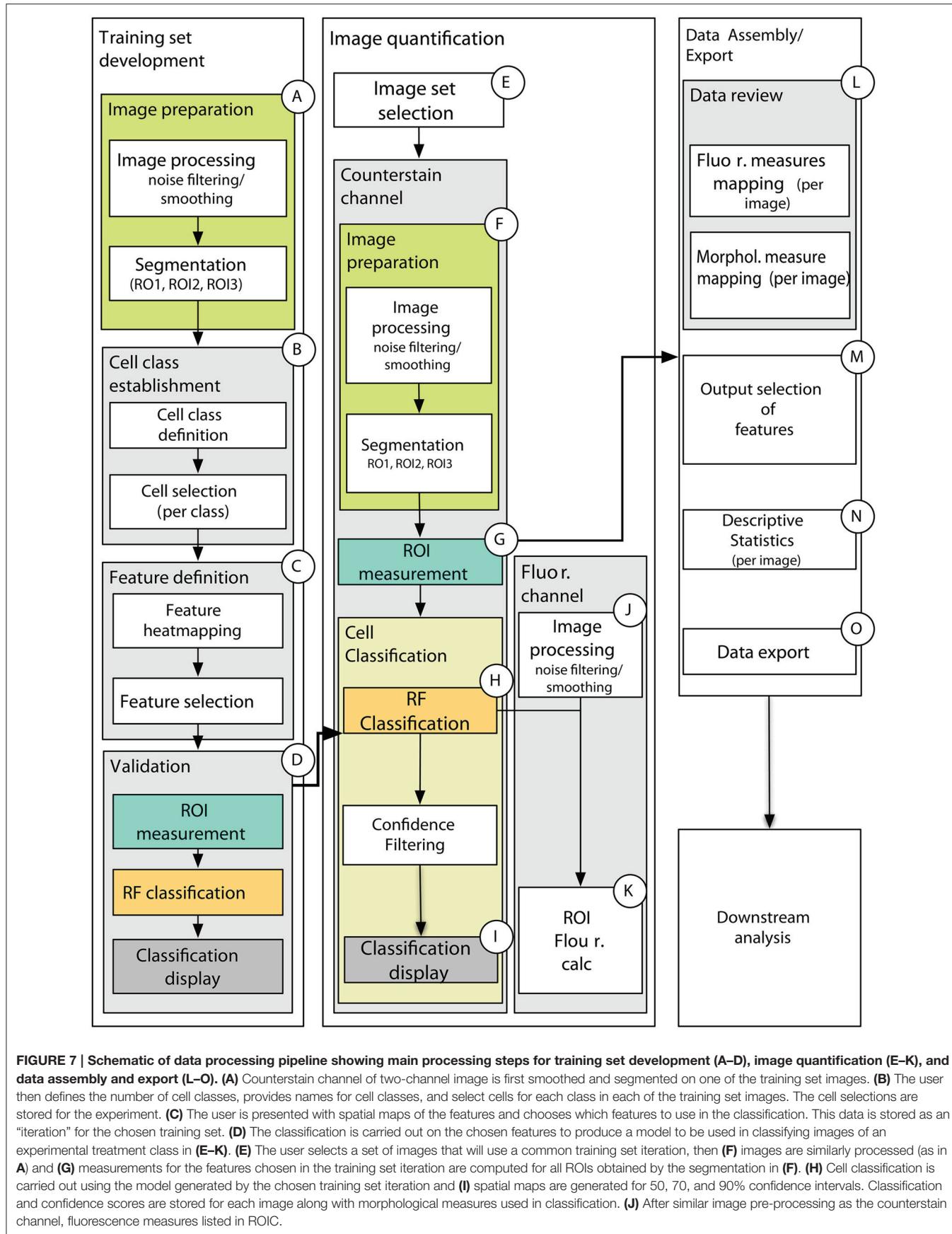
FIGURE 7 | Schematic of data processing pipeline showing main processing steps for training set development (A–D), image quantification (E–K), and data assembly and export (L–O). (A) Counterstain channel of two-channel image is first smoothed and segmented on one of the training set images. **(B)** The user then defines the number of cell classes, provides names for cell classes, and select cells for each class in each of the training set images. The cell selections are stored for the experiment. **(C)** The user is presented with spatial maps of the features and chooses which features to use in the classification. This data is stored as an "iteration" for the chosen training set. **(D)** The classification is carried out on the chosen features to produce a model to be used in classifying images of an experimental treatment class in **(E–K)**. **(E)** The user selects a set of images that will use a common training set iteration, then **(F)** images are similarly processed (as in **A**) and **(G)** measurements for the features chosen in the training set iteration are computed for all ROIs obtained by the segmentation in **(F)**. **(H)** Cell classification is carried out using the model generated by the chosen training set iteration and **(I)** spatial maps are generated for 50, 70, and 90% confidence intervals. Classification and confidence scores are stored for each image along with morphological measures used in classification. **(J)** After similar image pre-processing as the counterstain channel, fluorescence measures listed in ROIC.

localization of cell types within the tissue. However, using coordinates as a sole criterion for the classification can obviously not discriminate between different spatially dispersed cell types, as xylem vessels. Furthermore, the relative location of a cell to the manually selected origin of coordinates can change during development, e.g., the cambium is progressively pushed away from the center of the hypocotyl. Our approach is not only based on the position of each cell within the tissue context but also on features such as cell wall thickness or shape and is therefore expected to be less sensitive to variation in growth than strategies relying on coordinates only.

Automated image analysis is required for large screens involving hundreds of samples, like mutant screens and association mapping. Whereas, segmentation should not be affected by the morphologic variation across the samples, classification might require time-consuming extension of trainings sets. Application of a training set derived from wild type on hypocotyls of $knat1^{bp-9}$, a mutant which is characterized by severely distorted radial organization, indicated that applying stringent confidence filtering can reduce misclassifications in the mutant efficiently. However, the boundaries between xylem I and II cell types were inaccurately recognized when using a wild-type training set on mutant hypocotyls. This underlies the dominant nature of the radius feature, which measures the radial displacement of a segment from the center of the hypocotyl. Omission of the radius feature, improvements to training sets (image and cell choices) or simply scaling of "radius" are promising means to improve the classification of cells at the boundary between xylem I and II.

Although plant cell walls are considered to be dynamic structures, deposition of wall components is usually stable and irreversible; as opposed to changes in the abundance of short-lived gene products, which can occur in the range of minutes. This, together with easy fixation and efficient preservation of walls as compared to the cytoplasm, makes chemical cell wall properties a good marker for irreversible decisions in cell differentiation. The onset of cell death in xylem vessels is, for example, marked by the lignification of secondary vessel walls (Smith et al., 2013). Our image analysis toolkit permits the analysis of large numbers of commercially available antibodies against different cell wall epitopes. Alternative methods, which offer theoretically the same spatial resolution, to assess the chemical composition of walls of single cells are vibrational spectroscopy and ToF-SIMS (Gorzsás et al., 2011; Gerber et al., 2014; Felten et al., 2015). However, long acquisition times, rather poor image quality and absence of automated image analysis are obstacles that have not yet been overcome by these spectroscopic methods. While in specific cases methods of spatially resolved spectroscopy can provide important chemical information they are, in contrast to the here presented method, not suitable for larger genetic screens or highly resolved time courses.

Here, we present proof-of-concept studies employing images from Arabidopsis hypocotyls. While we expect that automated segmentation of similarly processed plant material, irrespective of the species, should be feasible with no or little adaptation to the script, there may be a need to optimize tissue fixation, sectioning and mounting for other materials than transverse sections of Arabidopsis hypocotyls or stems. On the other hand, application of our pipeline in the analysis of epitope distribution in whole-mount samples, e.g., root tip, or in live imaging of fluorescent markers should be within reach. Images derived from whole-mount samples or live imaging may, however, be more difficult to segment due to lower signal to noise ratios.

In its current form, our data analysis pipeline efficiently and accurately provides a wealth of morphometric data for automatically categorizing cell types of transverse sections of Arabidopsis hypocotyls at various growth stages. Furthermore, our pipeline provides a robust means to accurately quantify immunofluorescence for specific cell types, filterable by confidence scores for individual cells. This manuscript, along with the accompanying script package, thus presents an initial exploration into the application of this MATLAB-based analytical approach of segmentation, classification and quantification of confocal images; one that foreseeably quantifies any number of fluorescence targets on separate channels in more sophisticated fluorescence-based experiments on living or fixed tissues.

AUTHOR CONTRIBUTIONS

HH and UF designed the research. HH prepared the specimens and captured the images. HH, AF and CL wrote the code and analyzed the data. HH and UF wrote the manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00119

REFERENCES

Band, L. R., Wells, D. M., Fozard, J. A., Ghetiu, T., French, A. P., Pound, M. P., et al. (2014). Systems analysis of auxin transport in the *Arabidopsis* root apex. *Plant Cell* 26, 862–875. doi: 10.1105/tpc.113.119495

Bhalerao, R. P., and Fischer, U. (2014). Auxin gradients across wood-instructive or incidental? *Physiol. Plant.* 151, 43–51. doi: 10.1111/ppl.12134

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Cobb, J. N., Declerck, G., Greenberg, A., Clark, R., and McCouch, S. (2013). Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. *Theor. Appl. Genet.* 126, 867–887. doi: 10.1007/s00122-013-2066-0

Dhondt, S., Wuyts, N., and Inzé, D. (2013). Cell to whole-plant phenotyping: the best is yet to come. *Trends Plant Sci.* 18, 428–439. doi: 10.1016/j.tplants.2013.04.008

Federici, F., Dupuy, L., Laplaze, L., Heisler, M. G., and Haseloff, J. (2012). Integrated genetic and computation methods for in planta cytometry. *Nat. Methods* 9, 483. doi: 10.1038/nmeth.1940

Felten, J., Hall, H. C., Jaumot, J., Tauler, R., de Juan, A., and Gorzsás, A. (2015). Vibrational spectroscopic image analysis of biological material using multivariate curve resolution-alternating least squares (MCR-ALS). *Nat. Protoc.* 10, 217–240. doi: 10.1038/nprot.2015.008

Field, G. D., Gauthier, J. L., Sher, A., Greschner, M., Machado, T. A., Jepson, L. H., et al. (2010). Functional connectivity in the retina at the resolution of photoreceptors. *Nature* 467, 673–677. doi: 10.1038/nature09424

Furbank, R. T., and Tester, M. (2011). Phenomics-technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Gerber, L., Viet Mai, H., Linh, T., Hoang Anh Tuan, K., Malmberg, P., Hanrieder, J., et al. (2014). Using imaging ToF-SIMS data to determine the cell wall thickness of fibers in wood. *Surf. Interface Anal.* 46, 225–228. doi: 10.1002/sia.5661

Gorzsás, A., Stenlund, H., Persson, P., Trygg, J., and Sundberg, B. (2011). Cell-specific chemotyping and multivariate imaging by combined FT-IR microspectroscopy and orthogonal projections to latent structures (OPLS) analysis reveals the chemical landscape of secondary xylem. *Plant J.* 66, 903–914. doi: 10.1111/j.1365-313X.2011.04542.x

Gregor, T., Wieschaus, E. F., McGregor, A. P., Bialek, W., and Tank, D. W. (2007). Stability and nuclear dynamics of the bicoid morphogen gradient. *Cell* 130, 141–152. doi: 10.1016/j.cell.2007.05.026

Hall, H. C., Cheung, J., and Ellis, B. E. (2013). Immunoprofiling reveals unique cell-specific patterns of wall epitopes in the expanding *Arabidopsis* stem. *Plant J.* 74, 134–147. doi: 10.1111/tpj.12111

Janecek, A., Gansterer, W., Demel, M. A., and Ecker, G. F. (2008). On the relationship between feature selection and classification accuracy. *JMLR W&P.* 4, 90–105. Available online at: http://jmlr.csail.mit.edu/proceedings/papers/v4/janecek08a/janecek08a.pdf

Jouannet, V., Brackmann, K., and Greb, T. (2015). (Pro)cambium formation and proliferation: two sides of the same coin? *Curr. Opin. Plant Biol.* 23, 54–60. doi: 10.1016/j.pbi.2014.10.010

Kierzkowski, D., Lenhard, M., Smith, R., and Kuhlemeier, C. (2013). Interaction between meristem tissue layers controls phyllotaxis. *Dev. Cell* 26, 616–628. doi: 10.1016/j.devcel.2013.08.017

Liebsch, D., Sunaryo, W., Holmlund, M., Norberg, M., Zhang, J., Hall, H. C., et al. (2014). Class I KNOX transcription factors promote differentiation of cambial derivatives into xylem fibers in the *Arabidopsis* hypocotyl. *Development* 141, 4311–4319. doi: 10.1242/dev.111369

Luengo Hendriks, C. L., Keränen, S. V. E., Fowlkes, C. C., Simirenko, L., Weber, G. H., DePace, A. H., et al. (2006). Three-dimensional morphology and gene expression in the *Drosophila* blastoderm at cellular resolution I: data acquisition pipeline. *Genome Biol.* 7, R123. doi: 10.1186/gb-2006-7-12-r123

Montenegro-Johnson, T. D., Stamm, P., Strauss, S., Topham, A. T., Tsagris, M., Wood, A. T. A., et al. (2015). Digital single-cell analysis of plant organ development using 3DCellAtlas. *Plant Cell* 27, 1018–1033. doi: 10.1105/tpc.15.00175

Nilufar, S., and Perkins, T. J. (2014). Learning a cost function for microscope image segmentation. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2014, 5506–5509. doi: 10.1109/embc.2014.6944873

Otsu, N. (1979). Threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* 9, 62–66. doi: 10.1109/TSMC.1979.4310076

Pound, M. P., French, A. P., Wells, D. M., Bennett, M. J., and Pridmore, T. P. (2012). CellSeT: novel software to extract and analyze structured networks of plant cells from confocal images. *Plant Cell* 24, 1353–1361. doi: 10.1105/tpc.112.096289

Rahaman, M. M., Chen, D., Gillani, Z., Klukas, C., and Chen, M. (2015). Advanced phenotyping and phenotype data analysis for the study of plant growth and development. *Front. Plant Sci.* 6:619. doi: 10.3389/fpls.2015.00619

Sankar, M., Nieminen, K., Ragni, L., Xenarios, I., and Hardtke, C. S. (2014). Automated quantitative histology reveals vascular morphodynamics during *Arabidopsis* hypocotyl secondary growth. *Elife* 3:e01567. doi: 10.7554/eLife.01567

Smith, R. A., Schuetz, M., Roach, M., Mansfield, S. D., Ellis, B., and Samuels, L. (2013). Neighboring parenchyma cells contribute to *Arabidopsis* xylem lignification, while lignification of interfascicular fibers is cell autonomous. *Plant Cell* 25, 3988–3999. doi: 10.1105/tpc.113.117176

Uggla, C., Moritz, T., Sandberg, G., and Sundberg, B. (1996). Auxin as a positional signal in pattern formation in plants. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9282–9286. doi: 10.1073/pnas.93.17.9282

Wiesmann, V., Franz, D., Held, C., Münzenmayer, C., Palmisano, R., and Wittenberg, T. (2015). Review of free software tools for image analysis of fluorescence cell micrographs. *J. Microsc.* 257, 39–53. doi: 10.1111/jmi.12184

Yoshida, S., de Reuille, P. B., Lane, B., Bassel, G. W., Prusinkiewicz, P., Smith, R. S., et al. (2014). Genetic control of plant development by overriding a geometric division rule. *Dev. Cell* 29, 75–87. doi: 10.1016/j.devcel.2014.02.002

Phenotyping of Eggplant Wild Relatives and Interspecific Hybrids with Conventional and Phenomics Descriptors Provides Insight for Their Potential Utilization in Breeding

Prashant Kaushik, Jaime Prohens *, Santiago Vilanova, Pietro Gramazio and Mariola Plazas

Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, Valencia, Spain

Eggplant (*Solanum melongena*) is related to a large number of wild species that are a source of variation for breeding programmes, in particular for traits related to adaptation to climate change. However, wild species remain largely unexploited for eggplant breeding. Detailed phenotypic characterization of wild species and their hybrids with eggplant may allow identifying promising wild species and information on the genetic control and heterosis of relevant traits. We characterized six eggplant accessions, 21 accessions of 12 wild species (the only primary genepool species *S. insanum* and 11 secondary genepool species) and 45 interspecific hybrids of eggplant with wild species (18 with *S. insanum* and 27 with secondary genepool species) using 27 conventional morphological descriptors and 20 fruit morphometric descriptors obtained with the phenomics tool Tomato Analyzer. Significant differences were observed among cultivated, wild and interspecific hybrid groups for 18 conventional and 18 Tomato Analyzer descriptors, with hybrids generally having intermediate values. Wild species were generally more variable than cultivated accessions and interspecific hybrids displayed intermediate ranges of variation and coefficient of variation (CV) values, except for fruit shape traits in which the latter were the most variable. The multivariate principal components analysis (PCA) reveals a clear separation of wild species and cultivated accessions. Interspecific hybrids with *S. insanum* plotted closer to cultivated eggplant, while hybrids with secondary genepool species generally clustered together with wild species. Many differences were observed among wild species for traits of agronomic interest, which allowed identifying species of greatest potential interest for eggplant breeding. Heterosis values were positive for most vigor-related traits, while for fruit size values were close to zero for hybrids with *S. incanum* and highly negative for hybrids with secondary genepool species. Our results allowed the identification of potentially interesting wild species and interspecific hybrids for introgression breeding in eggplant. This is an important step for broadening the genetic base of eggplant and for breeding for adaptation to climate change in this crop.

Keywords: descriptors, genepools, intespecific hybrids, introgression breeding, phenomics, *Solanum melongena*, Tomato Analyzer

INTRODUCTION

Eggplant (*Solanum melongena* L.) is an important vegetable in tropical and subtropical regions across the world, where it is a source of dietary fiber, micronutrients and bioactive compounds (Mennella et al., 2010; Niño-Medina et al., 2014; San José et al., 2014). At present eggplant is the sixth most important vegetable after tomato, watermelon, onion, cabbage, and cucumber and the most important *Solanum* crop native to the Old World (FAO, 2016). At the global level, it has been one of the crops with the greatest increase in production in the last years, with total production rising by 59% in a decade, from $31.0 \cdot 10^6$ t in 2004 to $49.3 \cdot 10^6$ t in 2013 (FAO, 2016).

The narrow genetic base of eggplant, probably a consequence of a genetic bottleneck during its domestication in Southeast Asia (Meyer et al., 2012), is a limitation to obtain major breeding advances. This limited genetic diversity contrasts with the large morphological and genetic variation present in the eggplant wild relatives (Meyer et al., 2012; Vorontsova et al., 2013; Vorontsova and Knapp, in press). Phylogenetically, eggplant is a member of the so-called "spiny solanums" group (*Solanum* subgenus *Leptostemonum*), which contains many wild species from the Old World, most of them from Africa (Vorontsova et al., 2013; Vorontsova and Knapp, in press). These wild species could represent a source of variation for developing a new generation of eggplant cultivars with dramatically improved yield and quality, as well as for addressing the challenges posed by adaptation to the climate change. In this respect, resistance and tolerance to several major diseases and pests is found among wild eggplant relatives (Daunay and Hazra, 2012; Rotino et al., 2014) and they can also be found in a wide range of environmental conditions, including desertic and semi-desertic areas, environments with extreme temperatures (Knapp et al., 2013; Vorontsova and Knapp, in press). Some eggplant wild relatives are known to possess high levels of chlorogenic acid and other bioactive compounds of interest for human health (Mennella et al., 2010; Meyer et al., 2015). However, with a few exceptions (Rotino et al., 2014; Liu et al., 2015), eggplant breeders have largely neglected the potential of wild species for eggplant breeding, and contrarily to other crops like tomato (Díez and Nuez, 2008), wild relatives have not made a relevant contribution to the development of new eggplant cultivars.

Eggplant can be crossed with a large number of wild relatives (Daunay and Hazra, 2012; Rotino et al., 2014; Plazas et al., 2016). The closest wild relative of eggplant is *S. insanum* (Knapp et al., 2013; Vorontsova et al., 2013), which is naturally distributed in Southeast Asia, Madagascar and Mauritius (Knapp et al., 2013; Vorontsova and Knapp, in press), where it is frequently found as a weed (Mutegi et al., 2015). *Solanum insanum* is considered as the wild ancestor of eggplant and is the only species in the primary genepool of cultivated eggplant (Syfert et al., 2016). Hybrids of eggplant with *S. insanum* are easily obtained; fruits from interspecific hybridization have many seeds, which have high germination rates, and the hybrid plants are fully fertile (Davidar et al., 2015; Plazas et al., 2016). Interspecific hybrids have also been obtained with many wild species from

the secondary genepool (Daunay and Hazra, 2012; Rotino et al., 2014; Plazas et al., 2016), which includes some 50 African and Southeast Asian species (Vorontsova et al., 2013; Syfert et al., 2016). The degree of success of interspecific sexual hybridization between eggplant and secondary genepool species, as well as the hybrid fertility is variable depending on the species involved and the direction of the cross (Plazas et al., 2016).

The characterization of wild species and interspecific hybrids for traits of interest for breeders is a fundamental step for the efficient utilization of crop wild relatives in breeding. Combined data on the cultivated and wild species and their interspecific hybrids, not only allows identifying sources of variation and materials of potential interest, but also provides information on the inheritance of some traits present in the wild species, as has been demonstrated in crosses between *S. incanum* and eggplant (Prohens et al., 2013). Also, characterization of these materials for vigor traits may allow identification of materials potentially useful as rootstocks. In this respect, highly vigorous eggplant of wild relatives and interspecific hybrids are increasingly used for eggplant grafting, as they induce precocity and higher yield and many of them are tolerant to biotic and abiotic stresses (Gisbert et al., 2011; Daunay and Hazra, 2012). In the case of eggplant wild relatives there are a number of studies on their taxonomic and phylogenetic relationships (Vorontsova et al., 2013; Vorontsova and Knapp, in press), of resistance or tolerance to diseases and pests (Bubici and Cirulli, 2008; Daunay and Hazra, 2012; Naegele et al., 2014). However, to our knowledge there are no comprehensive studies on the morphological and agronomic traits of interest in a set of wild species of the primary and secondary genepools of eggplant and their interspecific hybrids with cultivated eggplant.

Several characterization studies in eggplant with standardized morphological and agronomic descriptors developed by the European Eggplant Genetic Resources Network (EGGNET; van der Weerden and Barendse, 2007) and the International Board for Plant Genetic Resources (IBPGR, 1990) have revealed that are suited for providing a useful morphological and agronomic characterization for eggplant breeders (Prohens et al., 2005; Muñoz-Falcón et al., 2009; Boyaci et al., 2015). EGGNET and IBPGR descriptors have been successfully used for evaluating segregating generations of interspecific crosses between eggplant and related species (Prohens et al., 2012, 2013). In addition to conventional morphological descriptors fruit phenomics data provide eggplant breeders with relevant information for evaluating the variation of the fruit morphology. In this respect, the phenomics tool Tomato Analyzer (Rodríguez et al., 2010) has revealed as useful for the detailed morphometric analysis of fruit size and shape of eggplant and related materials (Prohens et al., 2012; Hurtado et al., 2013).

Here we characterize cultivated eggplant, wild relatives from the primary and secondary genepools and interspecific hybrids between cultivated eggplant and wild relatives using conventional and Tomato Analyzer descriptors. Apart from providing a characterization of the three types of materials studied and

their differences, we aim to evaluate the interest for breeding of different wild relatives using characterization data of the wild relatives and of their interspecific hybrids with eggplant. The information obtained may also provide clues on the interest of wild species and hybrids as potential rootstocks for eggplant.

MATERIALS AND METHODS

Plant Material

The plant material included six accessions of cultivated eggplant (*S. melongena*), 21 accessions of a total of 12 wild species, and 45 interspecific hybrids between the eggplant accessions and seven of the wild species (**Table 1**). The eggplant accessions include materials from both the Occidental (Ivory Coast) and Oriental (Sri Lanka) cultivated genepools (Vilanova et al., 2012; Cericola et al., 2013). Among the wild relatives, three accessions belong

to the primary genepool (GP1) *S. insanum*, and 18 accessions to secondary genepool (GP2) species, namely *S. anguivi* ($n = 2$), *S. campylacanthum* ($n = 3$), *S. dasyphyllum* ($n = 1$), *S. incanum* ($n = 1$), *S. lichtensteinii* ($n = 2$), *S. lidii* ($n = 2$), *S. linnaeanum* ($n = 2$), *S. pyracanthos* ($n = 1$), *S. tomentosum* ($n = 1$), *S. vespertilio* ($n = 2$), and *S. violaceum* ($n = 1$). All the accessions are deposited at the germplasm bank of the Universitat Politècnica de València (València, Spain). The 45 interspecific hybrids were obtained after reciprocal crossings between cultivated eggplant and wild relatives (Plazas et al., 2016) resulting in 18 hybrids between eggplant and primary genepool species and 27 hybrids between eggplant and secondary genepool species (**Table 1**). Five plants per accession or interspecific hybrid were grown under open field conditions during the summer season of 2015 at the Universitat Politècnica de València (Valencia, Spain; GPS coordinates of the plot: 39° 28′ 55″ N, 0° 22′ 11″ W; altitude 7 m a.s.l.). Plants were spaced 1.2 m between

TABLE 1 | Accessions of cultivated eggplant (*Solanum melongena*) and wild relatives of the primary and secondary genepools, and interspecific hybrids between cultivated eggplant and wild relatives used for the morphological and phenomics characterization.

Species	Accession	Germplasm collection code	Country of origin	Interspecific hybrids with cultivated eggplant accessions					
				MEL1	MEL2	MEL3	MEL4	MEL5	MEL6
CULTIVATED EGGPLANT									
S. melongena	MEL1	BBS-118/B	Ivory Coast						
	MEL2	BBS-146	Ivory Coast						
	MEL3	BBS-175	Ivory Coast						
	MEL4	07145	Sri Lanka						
	MEL5	8104	Sri Lanka						
	MEL6	Ampara	Sri Lanka						
WILD PRIMARY GENEPOOL (GP1)									
S. insanum	INS1	SLKINS-1	Sri Lanka	MEL1 × INS1	MEL2 × INS1	MEL3 × INS1	MEL4 × INS1	INS1 × MEL5	MEL6 × INS1
	INS2	SLKINS-1	Sri Lanka	MEL1 × INS2	MEL2 × INS2	MEL3 × INS2	MEL4 × INS2	MEL5 × INS2	MEL6 × INS2
	INS3	MM498	Japan	INS3 × MEL1	INS3 × MEL2	INS3 × MEL3	INS3 × MEL4	MEL5 × INS3	INS3 × MEL6
WILD SECONDARY GENEPOOL (GP2)									
S. anguivi	ANG1	BBS119	Ivory Coast		MEL2 × ANG1	MEL3 × ANG1	MEL4 × ANG1	MEL5 × ANG1	
	ANG2	BBS125/B	Ivory Coast	MEL1 × ANG2	MEL2 × ANG2	ANG2 × MEL3	ANG2 × MEL4	MEL5 × ANG2	ANG2 × MEL6
S. campylacanthum	CAM5	MM680	Tanzania						
	CAM6	MM700	Kenya						
	CAM8	MM1426	Tanzania						
S. dasyphyllum	DAS1	MM1153	Uganda	MEL1 × DAS1	MEL2 × DAS1	MEL3 × DAS1		MEL5 × DAS1	
S. incanum	INC1	MM664	Israel	INC1 × MEL1		MEL3 × INC1		MEL5 × INC1	MEL6 × INC1
S. lichtensteinii	LIC1	MM674	South Africa	MEL1 × LIC1				MEL5 × LIC1	MEL6 × LIC1
	LIC2	MM677	Iran	MEL1 × LIC2		MEL3 × LIC2	MEL4 × LIC2		
S. lidii	LID1	4788	Spain						
	LID2	MM1005	Spain						
S. linnaeanum	LIN1	JPT0028	Spain						LIN1 × MEL6
	LIN3	MM195	Tunisia						
S. pyracanthos	PYR1	SOLN-66	Unknown						
S. tomentosum	TOM1	MM992	South Africa		MEL2 × TOM1	TOM1 × MEL3			
S. vespertilio	VES1	4601A	Spain						
	VES2	BGV-3218	Spain						
S. violaceum	VIO1	SLKVIL-1	Sri Lanka						

For the interspecific hybrids, the first and second parentals included in the hybrid code correspond to the female and male, respectively.

rows and 1.0 m within the row and distributed according to a completely randomized design. Drip irrigation was applied and 80 g plant^{-1} of a 10N–2.2P–24.9K plus micronutrients fertilizer (Hakaphos Naranja; Compo Agricultura, Barcelona, Spain) was applied during the whole cultivation period through the irrigation system. Plants were trained with bamboo canes and pruned when needed. Weeds were removed manually and no phytosanitary treatments were needed.

Characterization

All plants were characterized using 27 conventional morphological descriptors based on EGGNET (van der Weerden and Barendse, 2007) and IBPGR (IBPGR, 1990) descriptors (**Table 2**). These morphological descriptors describe different traits of the whole plant (4), leaf (7), inflorescence and flower (7) and fruit (9) and in general display limited GxE interaction (IBPGR, 1990). Except for descriptors concerning the whole plant (e.g., plant growth habit), for which one measurement was taken per plant (i.e., one measurement per replicate), five measurements were taken from each individual plant in order to obtain individual plant averages for the conventional morphological descriptors (i.e., five measurements per replicate). Using a similar approach, five fruits per plant (replicate), collected at the commercially ripe stage (i.e., physiologically immature) for cultivated eggplant and at a similar physiological stage (when they had attained full size but was not physiologically mature) in the case of wild species and interspecific hybrids, were cut opened longitudinally and scanned using an HP Scanjet G4010 photo scanner (Hewlett Packard, Palo Alto, CA, USA) at a resolution of 300 dpi. Scanned images were subjected to fruit morphometric analysis with the fruit shape phenomics tool Tomato Analyzer version 4 software (Rodríguez et al., 2010). A total of 20 fruit morphometric descriptors were recorded using this tool (**Table 2**).

Data Analyses

For each trait, the mean, range and coefficient of variation (CV, %) were calculated using average accession or hybrid values of cultivated eggplant ($n = 6$), wild relatives ($n = 21$) and interspecific hybrids ($n = 45$). Means of each accession or hybrid were subjected to analyses of variance (ANOVA) to detect differences among the three groups considered. Significance of differences among group means was evaluated using the Student-Newman-Keuls multiple range test at $P = 0.05$. Heterosis over mid parent (H; %) for the traits of greater agronomic importance was studied in the interspecific hybrids using formula H $= 100 \times ((F1 - MP)/MP)$, where F$_1$ = hybrid mean, and MP = mean of the parents. Values of H above 100% indicate that the hybrid is superior to the highest parent, and therefore present positive heterosis over the highest parent. Principal components analyses (PCA) were performed using pairwise Euclidean distances among accession or hybrid means for standardized characterization data. All the statistical analyses were performed using the Statgraphics Centurion XVI software (StatPoint Technologies, Warrenton, VA, USA).

RESULTS

Differences between Eggplant, Wild Relatives and Interspecific Hybrids

Significant differences ($P < 0.05$) were found among average values for the groups constituted by cultivated eggplant, wild relatives and interspecific hybrids for 18 out of the 27 conventional descriptors (**Table 3**). Generally, wild species and interspecific hybrids had larger plant size, greater leaf prickliness, more flowers per inflorescence, and less elongated fruits than the cultivated species. The cultivated species and interspecific hybrids had more anthocyanin pigmentation, larger leaf size, and greater number of flower parts than the wild species. Flower, fruit pedicel and fruit size had the greater average values in the cultivated species, while the smaller ones were for the wild species, with the interspecific hybrids having intermediate values. The three groups overlap for all conventional descriptors except for Leaf Pedicel Length, Corolla Diameter, Fruit Pedicel Length, Fruit Pedicel Diameter, and Fruit Weight, in which all the accessions of the cultivated species presented higher values than any of the wild species.

All Tomato Analyzer descriptors evaluated, except two (Rectangular and Shoulder Height) displayed significant ($P < 0.05$) differences among average values for the three groups (**Table 4**). For the eight Tomato Analyzer descriptors related to fruit size the cultivated eggplant presented significantly higher values than wild species, while for Ovoid it had lower values; interspecific hybrids presented intermediate values, in most cases being significantly different from both cultivated eggplant and wild species (**Table 4**). Cultivated eggplant had greater Distal Fruit Blockiness and Ellipsoid values than either wild species or interspecific hybrids, while wild species had higher values for Triangular than either cultivated species or interspecific hybrids. Similarly to conventional descriptors, the three groups overlap in the ranges of variation for all Tomato Analyzer descriptors except for Perimeter, Area, Height Mid-width, Maximum Height, Curved Height and Circular, in which there is no overlap between the range of variation of cultivated and wild species, with the values of the former being larger than those of the latter (**Table 4**).

Variation in Eggplant, Wild Relatives, and Interspecific Hybrids

Variation for the conventional and Tomato Analyzer descriptors was found in the materials studied (**Tables 3**, **4**; **Figure 1**). For most traits, more variation both in terms of range and CV was found in the wild species, compared to the cultivated eggplant accessions. For all conventional descriptors there was more variation in the wild species than in the cultivated eggplant, except for Shoot Tip Anthocyanin Intensity, the number of flower parts. Conversely, in the case of Tomato Analyzer descriptors, the range of variation was greater in wild species than in the cultivated eggplant for only six out of the 20 descriptors evaluated (Perimeter, Width Mid-height, Maximum Width, Rectangular, and Ovoid), while for the CV the wild species had a greater value than cultivated eggplant for nine of the descriptors, of which seven are related to fruit size (**Table 4**).

TABLE 2 | Descriptors used for phenotyping.

Descriptors	Units/Scale/Description
CONVENTIONAL MORPHOLOGICAL DESCRIPTORS	
Plant growth habit	3 = Upright; 7 = Prostrate
Plant height	cm
Stem diameter	mm
Shoot tip anthocyanin intensity	0 = Absent; 9 = Very strong
Leaf blade lobing	1 = Very weak (none); 9 = Very Strong
Leaf prickles (upper surface)	0 = None; 0 = Very many (>20)
Leaf surface shape	1 = Flat; 9 = Very convex or bullate
Leaf blade tip angle	$1 \leq 15°$; $9 \geq 160°$
Leaf pedicel length	cm
Leaf blade length	cm
Leaf blade width	cm
Number of flowers per inflorescence	—
Corolla color	1 = Greenish white; 9 = Bluish violet
Corolla diameter	mm
Number of flower prickles (calyx)	0 = None; 9 = Very many (>20)
Number of sepals	—
Number of petals	—
Number of stamens	—
Fruit pedicel length	mm
Fruit pedicel diameter	mm
Fruit length/Breadth ratio	1 = Broader than long; 9 = Several times as long as broad
Fruit cross section	1 = Circular, no grooves; 9 = Very irregular
Fruit apex shape	3 = Protruded; 7 = Depressed
Fruit weight	g
Fruit flesh density	1 = Very loose; 9 = Very dense
Fruit calyx length (relative)	1 = Very short (<10%); 9 = Very long (>75%)
Fruit calyx prickles	0 = None; 9 = Very many (>30)
TOMATO ANALYZER PHENOMICS FRUIT MORPHOMETRIC DESCRIPTORS	
Perimeter	cm
Area (A)	cm^2
Width mid-height	The width measured at 1/2 of the fruit's height (cm).
Maximum width	The maximum horizontal distance of the fruit (cm).
Height mid-width	The height measured at 1/2 of the fruit's width (cm).
Maximum height	The maximum vertical distance of the fruit (cm).
Curved height	The height measured along a curved line through the fruit (cm).
Fruit shape index external I	The ratio of maximum height to maximum width.
Fruit shape index external II	The ratio of height mid-width to width mid-height.
Curved fruit shape index	The ratio of curved height to the width of the fruit at mid-curved-height.
Proximal fruit blockiness	The ratio of the width at the upper blockiness position to width mid-height.
Distal fruit blockiness	The ratio of the width at the lower blockiness position to width mid-height.
Fruit shape triangle	The ratio of the width at the upper blockiness position to the width at the lower blockiness position.
Ellipsoid	The ratio of the error resulting from a best-fit ellipse to the area of the fruit. Smaller values indicate that the fruit is more ellipsoid.
Circular	The ratio of the error resulting from a best-fit circle to the area of the fruit. Smaller values indicate that the fruit is more circular.
Rectangular	The ratio of the area of the rectangle bounding the fruit to the area of the rectangle bounded by the fruit.
Shoulder height	The ratio of the average height of the shoulder points above the proximal end point to maximum height.
Obovoid	Calculated according to the formula provided in the Tomato Analyzer manual (Rodríguez et al., 2010). The higher the value, the greater is the area of the fruit below mid height.
Ovoid	Calculated according to the formula provided in the Tomato Analyzer manual (Rodríguez et al., 2010). The higher the value, the greater is the area of the fruit above mid height.
Fruit shape index internal	The ratio of the internal ellipse's height to its width.

The list displays conventional morphological descriptors based on EGGNET (van der Weerden and Barendse, 2007) and IBPGR (1990) descriptors list and phenomics fruit morphometric descriptors based on Tomato Analyzer software (Rodríguez et al., 2010) used for the characterization of accessions of cultivated eggplant (S. melongena; n = 6); wild relatives (n = 21) and interspecific hybrids between cultivated eggplant and wild relatives (n=45).

TABLE 3 | Variation parameters for conventional morphological descriptors.

Descriptors	Cultivated eggplant (n = 6)		Wild relatives (n = 21)		Interspecific hybrids (n = 45; 42 for fruit traits)		F-ratio	Probability
	Mean[a] (range)	CV (%)	Mean (range)	CV (%)	Mean (range)	CV (%)		
Plant growth habit	5.33 a (5.00–7.00)	15.3	4.71 a (3.00–7.00)	24.3	5.00 a (5.00–5.00)	0.0	2.51	0.0883
Plant height	97.1 a (69.7–111.7)	16.5	124.8 b (91.0–160.5)	17.5	141.9 b (91.0–199.0)	19.5	9.81	0.0002
Stem diameter (mm)	22.6 a (15.3–28.0)	20.3	24.3 ab (12.0–34.7)	25.9	27.8 b (18.3–38.3)	16.8	5.09	0.0087
Shoot tip anthocyanin intensity	3.33 b (0.00–7.00)	86.3	0.57 a (0.00–3.00)	211.2	2.06 ab (1.00–7.00)	112.7	5.43	0.0064
Leaf blade lobing	4.33 a (3.00–5.00)	23.8	4.81 ab (1.00–9.00)	52.4	6.02 b (3.00–9.00)	19.5	5.42	<0.0065
Leaf prickles (upper surface)	0.11 a (0.00–0.67)	244.9	3.38 b (0.00–9.00)	95.6	4.45 b (0.00–9.00)	66.7	6.03	0.0039
Leaf surface shape	5.67 a (5.00–9.00)	28.8	5.29 a (1.00–9.00)	45.2	6.33 a (5.00–9.00)	30.1	1.96	0.1489
Leaf blade tip angle	5.00 a (3.00–7.00)	25.3	4.48 a (3.00–7.00)	32.8	4.58 a (2.00–7.00)	33.1	0.29	0.7484
Leaf pedicel length (cm)	6.91 c (5.80–8.28)	14.1	2.74 a (0.63–4.61)	41.1	5.70 b (2.67–9.05)	25.5	42.27	<0.0001
Leaf blade length (cm)	22.0 b (19.7–24.9)	7.9	13.8 a (5.2–20.9)	33.7	21.0 b (15.0–31.9)	19.3	23.65	<0.0001
Leaf blade width (cm)	15.8 b (12.5–19.5)	18.3	8.7 a (3.3–18.7)	38.1	15.9 b (10.8–25.7)	21.3	34.41	<0.0001
Number of flowers per inflorescence	3.49 a (1.07–5.00)	42.2	8.33 b (1.00–16.10)	57.9	6.77 b (2.00–14.44)	43.3	4.58	<0.0135
Corolla color	5.67 a (5.00–7.00)	18.2	5.57 a (1.00–9.00)	37.9	6.02 a (3.00–7.00)	24.1	0.58	0.5620
Corolla diameter (mm)	43.3 c (37.2–49.9)	12.9	22.2 a (7.7–30.4)	30.4	35.8 b (20.4–49.9)	22.3	30.44	<0.0001
Number of flower prickles (calyx)	1.83 a (0.00–5.00)	100.1	3.62 a (0.00–9.00)	102.5	3.64 a (0.00–9.00)	85.2	0.86	0.4269
Number of sepals	5.57 b (5.00–7.00)	14.1	4.81 a (4.00–5.00)	8.4	5.25 b (5.00–6.00)	7.2	10.69	<0.0001
Number of petals	5.65 c (5.00–7.00)	13.1	4.81 a (4.00–5.00)	8.4	5.24 b (5.00–6.00)	6.5	13.24	<0.0001
Number of stamens	5.61 b (5.00–7.00)	13.7	4.80 a (4.00–5.00)	8.4	5.26 b (5.00–6.22)	7.9	10.70	<0.0001
Fruit pedicel length (mm)	43.8 c (33.0–52.2)	15.2	17.5 a (8.5–27.5)	30.1	28.2 b (8.6–50.3)	44.9	16.13	<0.0001
Fruit pedicel diameter (mm)	10.2 c (7.0–12.2)	20.7	2.84 a (1.0–5.1)	42.8	5.4 b (1.0–10.3)	51.3	23.92	<0.0001
Fruit length/breadth ratio	6.50 b (1.00–8.00)	42.1	2.71 a (1.00–5.00)	35.2	3.90 a (1.00–7.00)	44.1	12.81	<0.0001
Fruit cross section	5.67 a (5.00–7.00)	18.2	6.05 a (1.00–9.00)	47.5	5.45 a (2.00–9.00)	41.7	0.43	0.6537
Fruit apex shape	5.33 a (3.00–7.00)	36.9	5.19 a (3.00–7.00)	32	5.33 a (3.00–7.00)	30.9	0.05	0.9485

(Continued)

TABLE 3 | Continued

Descriptors	Cultivated eggplant (n = 6)		Wild relatives (n = 21)		Interspecific hybrids (n = 45; 42 for fruit traits)		F-ratio	Probability
	Mean[a] (range)	CV (%)	Mean (range)	CV (%)	Mean (range)	CV (%)		
Fruit weight (g)	244.7 c (94.4–354.5)	36.0	10.5 a (0.4–35.7)	111.6	58.4 b (0.6–224.2)	111.2	39.43	<0.0001
Fruit flesh density	6.33 b (5.00–7.00)	16.3	3.95 a (1.00–9.00)	63.2	5.38 ab (1.00–9.00)	44.4	3.60	0.0328
Fruit calyx length (relative)	2.67 a (1.00–3.00)	30.6	4.62 a (1.00–9.00)	57.5	4.05 a (1.00–9.00)	51.5	1.85	0.1647
Fruit calyx prickles	2.00 a (1.00–3.00)	54.8	3.48 a (0.00–9.00)	91.3	3.19 a (0.00–9.00)	95.0	0.58	0.5646

Values represent the mean, range (between brackets), and coefficient of variation (CV; %) for the conventional morphological descriptors studied in accessions of cultivated eggplant (S. melongena; n = 6), wild relatives (n = 21) and interspecific hybrids between cultivated eggplant and wild relatives (n = 45 except for fruit traits in which n = 42) and significance of mean differences among the three groups.
[a]Means within rows separated by different letters are significantly different according to the Student-Newman-Keuls test.

For interspecific hybrids a large range of variation was observed for many conventional descriptors, with variation parameters generally larger than those of the cultivated species and smaller than those of the wild species. In this respect, the range of variation was larger than that of the cultivated eggplant for all but nine conventional descriptors, while compared to wild species it was larger for 11 descriptors (**Table 3**). The coefficient of variation for conventional descriptors was also larger than in the cultivated species for all traits except nine (Plant Growth Habit, Stem Diameter, Leaf Blade Lobing, Leaf Prickles, Number of Flower Prickles, Number of Sepals, Number of Petals, Number of Stamens, and Fruit Apex Shape) and larger than that of the wild species for eight descriptors (Plant Height, Leaf Blade Tip Angle, Fruit Pedicel Length, Fruit Pedicel Diameter, Fruit Length/Breadth Ratio, Fruit Cross Section, Fruit Apex Shape, and Fruit Calyx Prickles; **Table 3**).

Regarding the variation for Tomato Analyzer traits, the range of variation in the interspecific hybrids was greater than those of cultivated eggplant and wild species for all traits except five in the case of cultivated eggplant, which correspond to fruit shape indexes and Circular, and only one (Ovoid) in the case of wild species (**Table 4**). Also, larger values were obtained in the CV for Tomato Analyzer descriptors in the interspecific hybrids compared to the cultivated species for all traits but seven. When compared to wild species the interspecific hybrids also presented higher CV for all traits, except four (**Table 4**).

Multivariate Analysis

The three first components of the principal components analysis made with all conventional and Tomato Analyzer descriptors accounted for 58.8% accounted of the total variation among accession means, with the first, second and third component accounting, respectively for 37.2, 12.0, and 9.5% of the total variation (**Table 5**). The first principal component was positively correlated to Corolla diameter, fruit size and to elongated fruit shape (**Table 5**). The second principal component was positively correlated to Plant Height and to obovoid fruit shape. The third principal component was positively correlated to Plant Growth Habit (i.e., prostrate habit), to multiple plant, leaf and corolla size traits, to a higher number of flower parts (sepals, petals and stamens) and to an increased prickliness in leaves, and flower and fruit calyces (**Table 5**).

The projection of eggplant, wild species and interspecific hybrids in the PCA plot reveals that although considerable diversity exists in both eggplant (black squares) and wild species (white symbols), the interspecific hybrids (gray symbols) present a more scattered distribution in the PCA plot (**Figures 2**, **3**). Interspecific hybrids with the primary genepool species *S. insanum* plot closer to the cultivated eggplant and are intermingled with it the PCA graphs. On the contrary, interspecific hybrids with secondary genepool species plot closer to the wild species and are also intermingled with them (**Figures 2**, **3**). The first component separates the group formed by eggplant and the interspecific hybrids with the primary genepool species *S. insanum*, which present positive values for this component, from the group formed by all the wild species and interspecific hybrids with secondary genepool species. Among the interspecific hybrids with secondary genepool species, those with *S. incanum* and *S. lichtensteinii* are the closest to eggplant in this first component (**Figures 2**, **3**). When considering the second component all eggplant accessions but one have positive values, while interspecific hybrids with *S. insanum* are equally distributed in the positive and negative values of this second component (**Figure 2**). Primary genepool wild species *S. insanum* and all secondary genepool species, except *S. campylacanthum*, *S. pyracanthos*, *S. tomentosum* and one accession of each of *S. anguivi* and *S. lidii* have negative values for this second component. When considering interspecific hybrids with secondary genepool species, although they are intermingled with the wild species for this second component most of the hybrids present positive values for this second component, with the exceptions being the hybrids with *S.*

TABLE 4 | Variation parameters for Tomato Analyzer phenomics fruit descriptors.

Descriptors	Cultivated eggplant (n = 6)		Wild relatives (n = 21)		Interspecifi hybrids (n = 42)		F-ratio	Probability
	Mean[a] (range)	CV (%)	Mean (range)	CV (%)	Mean (range)	CV (%)		
Perimeter	24.1 c (20.2–28.0)	12.1	6.1 a (2.1–16.2)	70.0	12.7 b (2.4–28.2)	73.0	13.45	<0.0001
Area	35.4 c (24.4–42.2)	20.5	3.8 a (0.3–17.2)	129.5	15.4 b (0.4–46.9)	109.0	13.47	<0.0001
Width mid-height	5.21 b (4.01–7.03)	22.8	1.87 a (0.63–4.93)	68.2	3.08 a (0.70–7.37)	66.0	8.80	0.0004
Maximum width	5.35 b (4.06–7.07)	21.7	1.88 a (0.64–4.96)	68.1	3.11 a (0.86–7.43)	66.1	9.22	0.0003
Height mid-width	8.17 c (6.39–10.51)	18.4	1.69 a (0.54–3.78)	68.9	4.09 b (0.74–10.41)	77.2	15.60	<0.0001
Maximum height	8.28 c (6.55–10.64)	18.1	1.72 a (0.55–3.90)	69.5	4.15 b (0.75–10.53)	77.0	15.57	<0.0001
Curved height	8.47 c (6.93–10.81)	17.2	1.95 a (0.85–4.52)	60.0	4.34 b (0.99–10.62)	73.2	15.53	<0.0001
Fruit shape index external I	1.64 c (0.93–2.23)	30.0	0.90 a (0.75–1.04)	8.4	1.22 b (0.75–1.91)	22.5	21.66	<0.0001
Fruit shape index external II	1.67 c (0.91–2.30)	31.3	0.89 a (0.74–1.03)	8.8	1.22 b (0.71–1.96)	23.2	21.99	<0.0001
Curved fruit shape index	1.72 c (0.99–2.36)	29.9	1.13 a (0.91–1.41)	13.2	1.35 b (0.89–1.99)	17.7	14.32	<0.0001
Proximal fruit blockiness	0.62 a (0.55–0.71)	9.1	0.66 a (0.58–0.78)	7.7	0.61 a (0.36–0.74)	12.4	5.04	0.0092
Distal fruit blockiness	0.73 b (0.65–0.77)	9.3	0.60 a (0.52–0.65)	6.5	0.64 a (0.52–0.75)	8.5	16.30	<0.0001
Fruit shape triangle	0.86 a (0.74–1.10)	16.6	1.12 b (0.91–1.49)	12.6	0.97 a (0.52–1.31)	16.5	9.91	0.0002
Ellipsoid	0.05 b (0.03–0.07)	29.7	0.02 a (0.01–0.03)	22.0	0.03 a (0.01–0.07)	39.8	10.98	<0.0001
Circular	0.16 c (0.08–0.25)	52.0	0.05 a (0.02–0.10)	41.7	0.09 b (0.03–0.21)	54.2	14.92	<0.0001
Rectangular	0.51 a (0.49–0.54)	3.7	0.51 a (0.48–0.54)	3.2	0.50 a (0.41–0.53)	5.3	2.75	0.0711
shoulder height	0.01 a (0.00–0.02)	56.7	0.01 a (0.00–0.03)	68	0.01 a (0.00–0.03)	74.1	0.23	0.7985
Obovoid	0.18 b (0.04–0.29)	55.5	0.05 a (0.00–0.18)	105.6	0.10 a (0.00–0.31)	74.4	8.63	0.0005
Ovoid	0.03 a (0.00–0.11)	160.0	0.09 b (0.00–0.21)	62.6	0.05 ab (0.00–0.17)	97.2	5.65	0.0054
Fruit shape index internal	1.67 c (0.91–2.30)	31.4	0.90 a (0.76–1.02)	8.5	1.22 b (0.72–1.96)	23.3	21.71	<0.0001

Mean, range (between brackets), and coefficient of variation (CV; %) for the Tomato Analyzer phenomics fruit morphometric descriptors studied in accessions of cultivated eggplant (S. melongena; n = 6), wild relatives (n = 21) and interspecific hybrids between cultivated eggplant and wild relatives (n = 42) and significance of mean differences among the three groups.
[a]Means within rows separated by different letters are significantly different according to the Student-Newman-Keuls test.

lichtensteinii (four out of five), *S. linnaeanum* and one of each of the interspecific hybrids with each of the species *S. anguivi* and *S. incanum* (this latter with a value very close to 0). Amazingly, the highest values for this second component correspond to interspecific hybrids with *S. anguivi* (**Figure 2**). For the third

component both eggplant and the interspecific hybrids with *S. insanum* are scattered and display positive or negative values (**Figure 3**). Most wild species accessions have negative values for this third component, except the accessions of *S. dasyphyllum*, *S. linnaeanum, S. pyracanthos,* and *S. violaceum*, as well as one
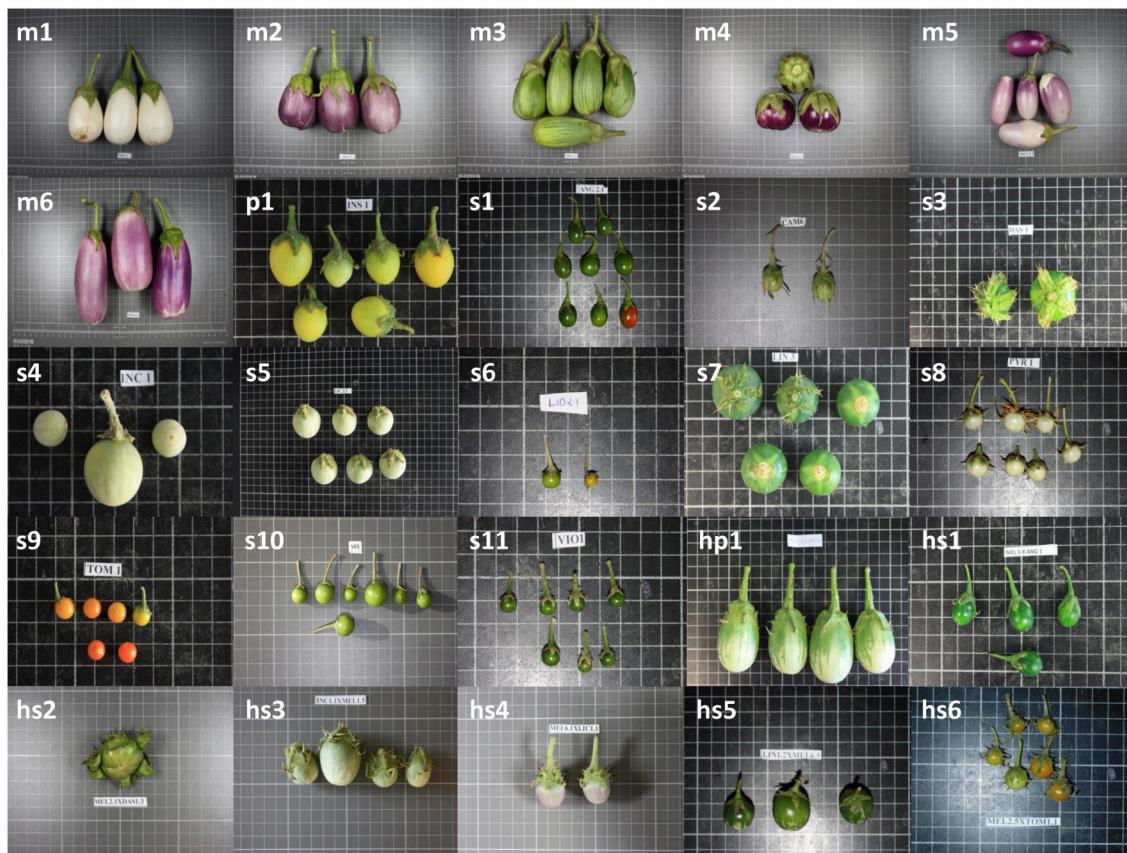
FIGURE 1 | Fruit samples of the materials used. This include: Cultivated eggplant (*S. melongena*) accessions MEL1 (m1) to MEL6 (m6); wild species of primary genepool *S. insanum* (p1); wild species of secondary genepool *S. anguivi* (s1), *S. campylacanthum* (s2), *S. dasyphyllum* (s3), *S. incanum* (s4), *S. lichtensteinii* (s5), *S. lidii* (s6), *S. linnaeanum* (s7), *S. pyracanthos* (s8), *S. tomentosum* (s9), *S. vespertilio* (s10), and *S. violaceum* (s11); interspecific hybrids between eggplant and primary genepool species *S. insanum* (hp1); and, interspecific hybrids between eggplant and secondary genepool species *S. anguivi* (hs1), *S. dasyphyllum* (hs2), *S. incanum* (hs3), *S. lichtensteinii* (hs4), *S. linnaeanum* (hs5), and *S. tomentosum* (hs6). Fruits are not depicted at the same scale; the size of the grid cells is 1 × 1cm.

accession of *S. incanum* (with values close to 0). The lowest values for this component are those of *S. lidii*, *S. vespertilio* and *S. tomentosum* (**Figure 3**). On the other hand all interspecific hybrids with secondary genepool species, with the exception of two interspecific hybrids with *S. anguivi*, present positive values for this third component. In this case, the highest values for the third component correspond to interspecific hybrids with *S. dasyphyllum*, *S. lichtensteinii,* and *S. incanum* (**Figure 3**).

Traits of Agronomic Interest in Wild Species

The 12 wild species evaluated presented considerable differences for traits of agronomic interest (**Table 6**). For example, important differences were found for vegetative traits. For example, the tallest plants were those of *S. anguivi*, which also presented thick stems (**Table 6**). Important differences were also found for Leaf Blade Lobing. The greatest leaf prickliness was observed *S. dasyphyllum*, *S. pyracanthos,* and *S. violaceum*, while *S. anguivi* and *S. tomentosusm* did not present prickles in the leaves. The largest leaf blades were those of *S. dasyphyllum* and *S. campylacanthum*, while the smallest were those of *S. tomentosum*,

with a Leaf Blade Length of 5.2 cm (**Table 6**). When considering flower and fruit traits, the two species with a larger number of flowers per inflorescence were *S. lidii* and *S. vespertilio*, with more than 13 flowers/inflorescence, while the smaller number was *S. insanum* (**Table 6**). Important differences were also observed for Corolla Color. All wild species had five petals (and sepals and stamens), except *S. lidii* and *S. vespertilio*, which had only four. The largest fruits were those of *S. incanum* and *S. lichtensteinii*, with average values above 25 g, more than 10-fold heavier than those of *S. anguivi*, *S. lidii*, *S. pyracanthos*, *S. tomentosum*, *S. vespertilio,* and *S. violaceum*. The highest calyx prickliness was observed in *S. linnaeanum*, *S. pyracanthos,* and *S. violaceum*, while *S. anguivi*, *S. lidii,* and *S. vespertilio* did not present calyx prickles (**Table 6**). The most elongated fruit were those of *S. incanum*, while the most flattened ones were those of *S. dasyphyllum* and *S. lidii* (**Table 6**).

Heterosis in Interspecific Hybrids

Interspecific hybrids between eggplant and its wild relatives generally displayed positive heterosis for plant size traits, with average heterosis values of up to 90.5% for Plant height and 46.2%

TABLE 5 | Correlation coefficients between morphological conventional and phenomics descriptors.

Descriptors	First principal component	Second principal component	Third principal component
Plant growth habit			0.151
Plant height (cm)		0.154	0.176
Stem diameter (mm)			0.266
Leaf blade lobing			0.258
Leaf prickles (upper surface)		−0.165	0.184
Leaf surface shape			0.236
Leaf blade length (cm)			0.291
Leaf blade width (cm)			0.306
Corolla diameter (mm)	0.184		0.153
Number of flower prickles (calyx)		−0.170	0.226
Number of sepals			0.275
Number of petals			0.267
Number of stamens			0.266
Fruit pedicel length (mm)	0.218		
Fruit pedicel diameter (mm)	0.218		
Fruit length/breadth ratio	0.191		
Fruit weight (g)	0.212		
Fruit calyx prickles		−0.190	0.253
Perimeter (cm)	0.225		
Area (cm^2)	0.219		
Width mid-height (cm)	0.204		
Maximum width (cm)	0.206		
Height mid-width (cm)	0.231		
Maximum height (cm)	0.231		
Curved height (cm)	0.231		
Fruit shape index external I	0.209		
Fruit shape index external II	0.209		
Curved fruit shape index	0.167		
Proximal fruit blockiness		−0.371	
Distal fruit blockiness	0.163	0.204	
Fruit shape triangle		−0.349	
Circular	0.189		
Rectangular		−0.245	
Shoulder height		0.159	
Obovoid		0.328	
Ovoid		−0.312	
Fruit shape index internal	0.208		
Eigenvalue	17.50	5.65	4.48
Variance explained (%)	37.23	12.04	9.53
Cumulative variance explained (%)	37.23	49.27	58.80

Values represent the correlation coefficients for the three first principal components in the collection of eggplant (S. melongena), wild relatives and interspecific hybrids evaluated. Only correlations with absolute values ≥0.150 have been listed.

for Stem diameter in the hybrids of eggplant with *S. dasyphyllum* (**Table 7**). The only negative value observed for these traits was for Stem Diameter in the interspecific hybrid with *S. linnaeanum*. Most interspecific hybrids presented higher prickliness than their parent species, and in consequence, very high average values for heterosis for Leaf Prickles are observed, with values between

91.0% for *S. dasyphyllum* and 800.0% for *S. tomentosum*. Leaf size traits were also, in general, heterotic in the interspecific hybrids, with the exception of Leaf Pedicel Length in *S. dasyphyllum* and *S. linnaeanum*. The same phenomenon was observed for the Number of Flowers per Inflorescence, with values of up to 87.7% in the hybrids with *S. tomentosum* (**Table 7**). The pigmentation of the corolla (Corolla Color) also presented average positive heterosis values in the hybrids of eggplant with five out of the seven wild species, the exception being interspecific hybrids with *S. anguivi* and *S. tomentosum*. The number of flower parts, represented by the Number of Petals, displayed low absolute values for heterosis in all cases (**Table 7**).

Regarding Fruit Weight, considerable differences were observed between the hybrids with the primary genepool species (*S. insanum*) on one hand, and the hybrids with secondary genepool species on the other. In this respect, while the hybrids with *S. insanum* displayed small negative average heterosis (−5.5%), not significantly different from 0, in the case of secondary genepool species, the heterosis for Fruit Weight is highly negative, with values between −60.4% for hybrids with *S. dasyphyllum* to −98.6% in hybrids with *S. tomentosum* (**Table 7**). As occurred for Leaf Prickles, positive heterosis values, although of smaller magnitude, were observed for Fruit Calyx Prickles, with the exception of the hybrids with *S. anguivi*, which did not present prickles in the calyx, and in consequence had a heterosis value of −100%. Finally, for fruit shape, the hybrids with primary genepool species *S. insanum* presented positive heterosis, while those with secondary genepool species had negative heterosis values (**Table 7**).

DISCUSSION

Crop wild relatives are widely recognized as an invaluable genetic resource for breeding, in particular for broadening the genetic base of crops with narrow genetic diversity, and as sources of variation for traits of interest in breeding crops, including adapting them to the challenges posed by climate change (Dempewolf et al., 2014). Modern varieties of many important crops carry introgressions from wild species resulting from breeding programmes performed in the last 100 years (Hajjar and Hodgkin, 2007). One of the most outstanding examples is tomato, where modern commercial hybrids carry different combinations of 15 different introgressions from different wild species (Díez and Nuez, 2008; Sabatini et al., 2013). However, in the case of eggplant, despite being one of the most important vegetables and being intercrossable with many wild relatives, there are few reports on the use of the variation available in the wild species for eggplant breeding (Daunay and Hazra, 2012; Rotino et al., 2014; Liu et al., 2015) and no modern commercial varieties of eggplant carrying introgressions from wild species are known to us.

In our study we have evaluated six accessions of cultivated eggplant, 21 accessions of 12 wild species, and 45 interspecific hybrids of cultivated eggplant with seven wild species. This
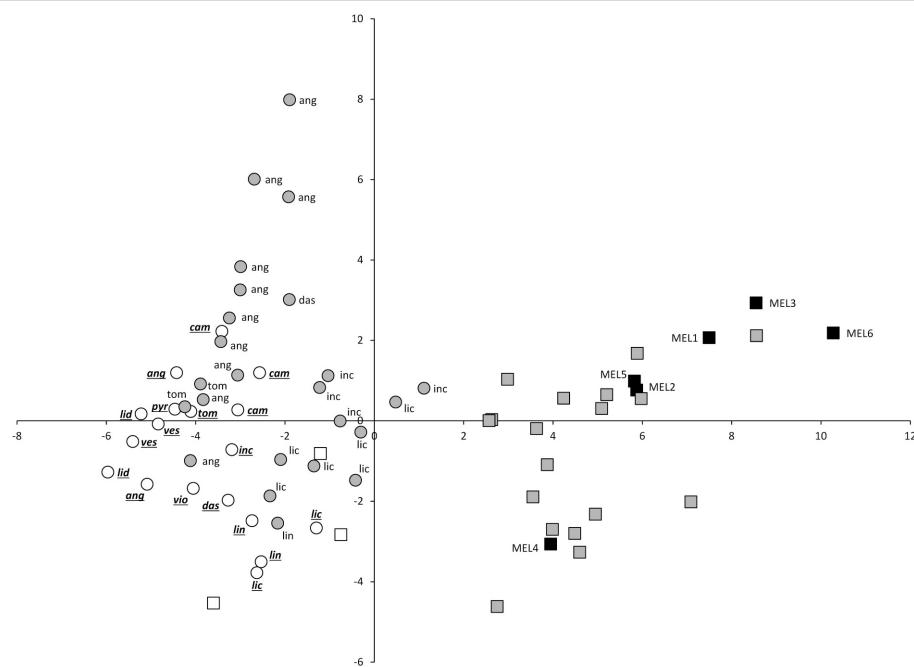
FIGURE 2 | First (X-axis) and second (Y-axis) principal components (37.2 and 12.0% of the total variation explained, respectively) scatterplot of cultivated eggplant, wild relatives and interspecific hybrids based on 27 conventional and 20 Tomato Analyzer morphological descriptors. Cultivated eggplant (*S. melongena*) is represented by black squares and the "MEL" accession code, primary genepool species *S. insanum* by white squares, interspecific hybrids between eggplant and *S. insanum* by gray squares, secondary genepool species by white circles (with species codes in underlined italics), and interspecific hybrids between eggplant and secondary genepool species by gray circles (with wild species codes in normal font). For secondary genepool species and their hybrids with eggplant, the following codes are used: ang (*S. anguivi*), cam (*S. campylacanthum*), das (*S. dasyphyllum*), inc (*S. incanum*), lic (*S. lichtensteinii*), lid (*S. lidii*), lin (*S. linnaeanum*), pyr (*S. pyracanthos*), tom (*S. tomentosum*), ves (*S. vespertilio*), vio (*S. violaceum*).

represents the largest study up to now on morphological and agronomic traits for breeding of this type of materials. As expected, many differences were found within and among cultivated eggplant, wild relatives and the interspecific hybrids for the conventional descriptors used, confirming the utility of the EGGNET (van der Weerden and Barendse, 2007) and IBPGR (1990) conventional morphological descriptors and Tomato Analyzer traits (Rodríguez et al., 2010) used for evaluating eggplant wild relatives and interspecific hybrids (Prohens et al., 2013).

Also, many differences were found for the traits studied among cultivated eggplant, wild species and interspecific hybrids. Although many of the wild species of eggplant thrive in arid and semi-arid conditions (Knapp et al., 2013; Vorontsova and Knapp, in press), when grown under the favorable conditions of cultivated environments, the wild species and their interspecific hybrids generally display a high vigor, expressed as average values for plant height and stem diameter above those of cultivated eggplant. This is of interest for developing new rootstocks, which generally require having high vigor (Gisbert et al., 2011), and opens the way to exploiting several to the wild species evaluated and interspecific hybrids as potential new rootstocks for eggplant. Another important trait of agronomic interest for which there were considerable differences among groups was prickliness, which was much greater in wild

species and interspecific hybrids, confirming that alleles from the cultivated eggplant are recessive (Doganlar et al., 2002; Gramazio et al., 2014; Portis et al., 2015). The number of flowers per inflorescence was also much greater in wild species and interspecific hybrids. This trait is very important in eggplant breeding, as a reduced value of this trait results in increased fruit size uniformity (Sękara and Bieniasz, 2008). Also, fruit size and shape, which are of great relevance for breeding (Daunay and Hazra, 2012; Portis et al., 2015), also differed considerably among the three groups, with the interspecific hybrids presenting intermediate values, although on most cases they were closer to those of the wild species, indicating dominance of the genes of the latter (Doganlar et al., 2002).

The much higher variation observed in wild species and interspecific hybrids for vegetative, flower and inflorescence traits compared to cultivated eggplant was expected, as we were comparing a single species with an admixture of different wild species or hybrids, which present a much higher genetic diversity (Meyer et al., 2012; Särkinen et al., 2013; Vorontsova et al., 2013). However, for traits related to the fruit size and shape much higher variation was observed in the cultivated eggplant than in the wild species, confirming the general observation that the morphological variation in the organ for which a crop is domesticated (in this case the fruit) increases during
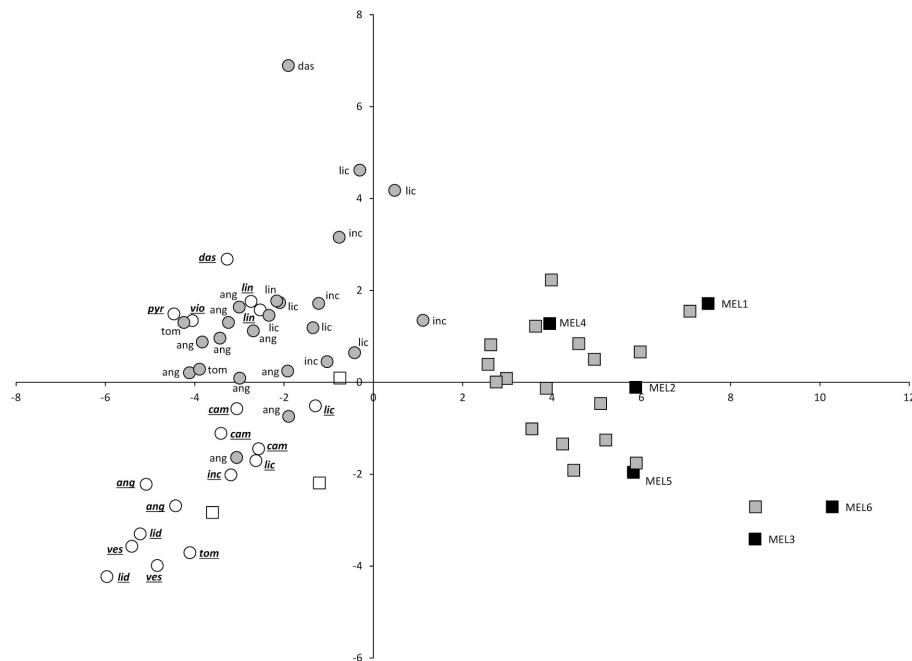
FIGURE 3 | First (X-axis) and thrid (Y-axis) principal components (37.2 and 9.5% of the total variation explained, respectively) scatterplot of cultivated eggplant, wild relatives and interspecific hybrids based on 27 conventional and 20 Tomato Analyzer morphological descriptors. Cultivated eggplant (*S. melongena*) is represented by black squares and the "MEL" accession code, primary genepool species *S. insanum* by white squares, interspecific hybrids between eggplant and *S. insanum* by gray squares, secondary genepool species by white circles (with species codes in underlined italics), and interspecific hybrids between eggplant and secondary genepool species by gray circles (with wild species codes in normal font). For secondary genepool species and their hybrids with eggplant, the following codes are used: ang (*S. anguivi*), cam (*S. campylacanthum*), das (*S. dasyphyllum*), inc (*S. incanum*), lic (*S. lichtensteinii*), lid (*S. lidii*), lin (*S. linnaeanum*), pyr (*S. pyracanthos*), tom (*S. tomentosum*), ves (*S. vespertilio*), vio (*S. violaceum*).

domestication (Meyer and Purugganan, 2014). Amazingly, in the case of interspecific hybrids a larger variation was found for most fruit size and shape traits than in the cultivated eggplant. Although most interspecific hybrids were more similar to the wild species, in some cases they were intermediate, revealing that different genic control mechanisms must exist for fruit size and shape among the wild relatives of eggplant. In this respect, the multivariate analysis clearly shows that interspecific hybrids with the primary genepool species *S. insanum* are morphologically closer to the cultivated eggplant, while the hybrids with secondary genepool species present a general morphology closer to that of the wild species. These results may support the hypothesis that *S. insanum* is the wild ancestor of cultivated eggplant (Knapp et al., 2013), as domestication should be easier when genes for domestication traits from the wild species display intermediate dominance rather than full dominance.

The study of individual wild species suggests that *S. anguivi*, *S. campylacanthum*, *S. pyracanthos,* and *S. violaceum* may be of interest for increasing the vigor of cultivated eggplant or for being used as rootstocks. Also, wild eggplant species use to have undesirable traits (e.g., prickliness, small fruit size, etc.) that have to be removed during the breeding (Rotino et al., 2014). In this case, the most desirable wild species are those that are most similar to the crop for these traits. For example, the lack of prickles or very low prickliness of *S. anguivi*, *S. campylacanthum*,

and *S. tomentosum* is a very favorable trait for breeders (Daunay and Hazra, 2012). Regarding fruit weight, the wild species with greater fruit weight should be the most interesting for breeders in order to recover fruit size in few backcross generations. In this case, *S. insanum*, *S. dasyphyllum,* and *S. lichtensteinii* should be the most interesting candidates if a rapid recovery of fruit size is desired. In any case, Prohens et al. (2013) showed that fruit size recovers quickly even in first backcrosses with the wild species *S. incanum*, which has an intermediate fruit size among wild species.

Although differences were observed among interspecific hybrids from different wild species, hybrids were in general vigorous, displaying heterosis for vigor traits. This phenomenon had already been described in interspecific hybrids with *S. incanum* (Gisbert et al., 2011; Prohens et al., 2013), and our results suggest that this is a common phenomenon in the hybrids between eggplant and wild relatives. Amazingly, most interspecific hybrids were highly heterotic for prickliness, with heterosis values over 100%. Prickles even appeared in interspecific hybrids with wild species that were not prickly, like *S. tomentosum*. In previous works, heterosis for prickliness had already been described in interspecific crosses in eggplant (Prohens et al., 2012; Devi et al., 2015; Plazas et al., 2016). Several studies with segregating populations of *S. linnaeanum* and *S. insanum* show that differences in prickliness between cultivated eggplant and wild relatives is under the control of

TABLE 6 | Average (±SE) values based on accession means for selected traits in the 12 wild species (one from the primary genepool GP1, S. *insanum*; and 11 from the secondary genepool GP2) evaluated.

Descriptors	GP1							GP2				
	S. insanum	S. anguivi	S. campylacanthum	S. dasyphyllum	S. incanum	S. lichtensteinii	S. lidii	S. linnaeanum	S. pyracanthos	S. tomentosum	S. vespertilio	S. violaceum
N	3	2	3	1	1	2	2	2	1	1	2	1
Plant height (cm)	108.7 ± 9.8	153.7 ± 6.3	150.2 ± 6.9	95.0	120.0	130.5 ± 12.5	108.6 ± 3.0	107.0 ± 0.0	141.7	104.0	115.5 ± 1.5	154.0
Stem diameter (mm)	22.8 ± 4.6	31.3 ± 2.0	21.5 ± 1.8	23.5	28.0	22.3 ± 2.4	14.0 ± 2.0	29.8 ± 0.2	34.7	19.5	28.6 ± 0.9	19.8
Leaf blade lobing	5.00 ± 0.00	2.00 ± 1.00	2.33 ± 0.67	9.00	3.00	5.00 ± 0.00	5.00 ± 0.00	9.00 ± 0.00	9.00	3.00	3.00 ± 0.00	7.00
Leaf prickles (upper surface)	3.33 ± 1.67	0.00 ± 0.00	0.67 ± 0.67	9.00	1.00	0.50 ± 0.50	6.00 ± 1.00	6.00 ± 1.00	7.00	0.00	4.00 ± 1.00	9.00
Leaf pedicel length (cm)	2.27 ± 0.51	1.98 ± 0.10	2.75 ± 0.71	1.2	2.3	2.75 ± 0.89	3.07 ± 0.33	2.99 ± 0.71	4.61	0.63	4.11 ± 0.13	3.95
Leaf blade length (cm)	8.9 ± 1.5	10.9 ± 1.6	19.1 ± 1.7	22.1	11.3	13.4 ± 2.8	14.7 ± 3.9	13.9 ± 2.4	16.9	5.2	14.3 ± 0.6	15.7
Leaf blade width (cm)	7.0 ± 1.0	7.4 ± 1.1	8.4 ± 1.4	18.7	7.8	9.3 ± 2.1	7.9 ± 3.5	9.5 ± 0.8	7.4	3.3	9.5 ± 0.7	12.9
Number of flowers per inflorescence	2.0 ± 1.0	8.2 ± 2.2	9.4 ± 1.7	10.6	9.1	5.1 ± 1.9	13.7 ± 0.7	3.0 ± 0.2	13.3	5.0	16.0 ± 0.2	10.7
Corolla color	5.67 ± 0.67	2.00 ± 1.00	7.67 ± 1.33	5.00	7.00	4.00 ± 1.00	7.00 ± 0.00	5.00 ± 0.00	9.00	3.00	5.00 ± 0.00	7.00
Number of petals	5.00 ± 0.00	5.00 ± 0.00	5.00 ± 0.00	5.00	5.00	5.00 ± 0.00	4.00 ± 0.00	5.00 ± 0.00	5.00	5.00	4.00 ± 0.00	5.00
Fruit weight (g)	26.5 ± 5.0	1.3 ± 0.5	4.6 ± 1.2	19.3	11.6	28.7 ± 6.2	0.4 ± 0.0	16.2 ± 3.3	1.0	0.5	1.2 ± 0.1	0.4
Fruit calyx prickles	3.00 ± 1.15	0.00 ± 0.00	1.67 ± 0.67	9.00	5.00	5.00 ± 0.00	0.00 ± 0.00	8.00 ± 1.00	7.00	5.00	0.00 ± 0.00	7.00
Fruit shape index external I	0.97 ± 0.39	0.95 ± 0.047	0.93 ± 0.04	0.79	1.04	0.93 ± 0.08	0.78 ± 0.04	0.94 ± 0.07	0.86	0.94	0.87 ± 0.01	0.85

Traits were selected so that they were relevant for breeding and useful to distinguish the different wild species.

TABLE 7 | Heterosis over mid parent values (%; ±SE) based on accession and interspecific hybrid means.

Descriptors	S. insanum	S. anguivi	S. dasyphyllum	S. incanum	S. lichtensteinii	S. linnaeanum	S. tomentosum
n	18	10	4/1[a]	4	6	1	2
Plant height (cm)	16.7 ± 4.6	34.4 ± 7.1	90.5 ± 7.6	36.8 ± 11.3	38.1 ± 4.4	2.3	23.3 ± 4.2
Stem diameter (mm)	10.5 ± 4.3	10.4 ± 3.8	46.2 ± 12.3	29.1 ± 11.0	39.8 ± 10.3	−18.7	23.8 ± 3.8
Leaf prickles (upper surface)	155.1 ± 34.5	260.0 ± 173.9	91.0 ± 5.4	733.3 ± 100.0	144.4 ± 92.9	100.0	800.0 ± 800.0
Leaf pedicel length (cm)	39.7 ± 6.5	22.5 ± 7.8	−21.6 ± 1.2	19.5 ± 2.7	24.9 ± 9.2	−13.3	56.3 ± 23.9
Leaf blade length (cm)	24.9 ± 4.1	22.2 ± 5.5	34.8 ± 5.7	47.6 ± 6.6	30.6 ± 6.3	3.9	22.8 ± 1.6
Leaf blade width (cm)	27.7 ± 4.5	38.2 ± 9.5	32.9 ± 5.0	67.7 ± 9.6	41.7 ± 8.5	7.1	22.4 ± 14.0
Number of flowers per inflorescence	70.1 ± 16.0	75.9 ± 16.3	36.9 ± 13.1	21.0 ± 9.4	42.7 ± 15.7	−1.8	87.7 ± 35.5
Corolla color	15.9 ± 4.3	−2.5 ± 4.6	18.9 ± 10.4	19.2 ± 3.0	16.2 ± 4.8	7.5	−0.1 ± 8.6
Number of petals	1.3 ± 2.1	−4.8 ± 1.6	1.9 ± 5.4	−4.4 ± 2.4	−2.2 ± 3.4	−3.2	−1.0 ± 1.0
Fruit weight (g)	−5.5 ± 6.9	−98.2 ± 0.3	−60.4	−86.6 ± 2.8	−89.4 ± 1.5	−89.9	−98.6 ± 0.3
Fruit calyx prickles	32.9 ± 25.2	−100.0 ± 0.0	80.0	27.1 ± 42.4	56.9 ± 27.6	80.0	29.1 ± 104.1
Fruit shape index external I	13.7 ± 3.5	−16.7 ± 6.9	−26.4	−13.6 ± 0.8	−15.0 ± 4.3	−40.8	−27.4 ± 8.0

Values are presented for traits of agronomic interest in the interspecific hybrids of eggplant with seven wild relatives (one from the primary genepool, S. incanum; and six from the secondary genepool).

[a]*For S. dasyphyllum data are available for four accessions for plant traits and only for one accession for fruit traits.*

a few QTL (Doganlar et al., 2002; Gramazio et al., 2014) and therefore prickliness should be easily removed in backcross generations. Although for fruit size traits negative heterosis was generally observed in the interspecific hybrids, indicating a greater similarity to the wild species, interspecific hybrids with primary genepool species *S. insanum* presented values close to zero, similarly to intraspecific hybrids of eggplant (Rodríguez-Burruezo et al., 2008), indicating intermediate dominance and values intermediate between both parental species. However, hybrids with wild species from the secondary genepool displayed highly negative heterosis, in some cases close to 100% like in interspecific hybrids with *S. anguivi* and *S. tomentosum*, suggesting that in these materials it may be more difficult to recover fruit size in the backcross generations.

In conclusion, the characterization with conventional descriptors and the Tomato Analyzer phenomics tool has allowed a detailed characterization of eggplant, close wild relatives and their interspecific hybrids. The high variation among wild species identified sources of variation and most promising species for traits of interest for eggplant breeding. The fact that interspecific hybrids with primary genepool species *S. insanum* are intermediate or close to eggplant for many traits, may facilitate the use of this species in introgression breeding and supports previous evidence that this species is the ancestor of cultivated eggplant. Also, the high vigor of most interspecific hybrids may be directly exploited by using them as rootstocks. The information obtained here on phenotypic characteristics and heterosis of wild species and interspecific hybrids is of interest for eggplant breeding. Given the adaptation of many wild species to stressful conditions, their

utilization in eggplant breeding may result in the development of a new generation of cultivars adapted to climate change challenges.

AUTHOR CONTRIBUTIONS

JP, SV, PG, and MP conceived and designed the research; PK and MP performed the phenotypic and phenomics characterization; PK, JP, and PG analyzed the data; JP, SV, PG, and MP wrote the manuscript. All authors read and approved the manuscript.

FUNDING

REFERENCES

Boyaci, H. F., Topcu, V., Tepe, A., Yildirim, I. K., Oten, M., and Aktas, A. (2015). Morphological and molecular characterization and relationships of Turkish local eggplant heirlooms. *Not. Bot. Horti Agrobot. Cluj Napoca* 43, 100–107. doi: 10.15835/nbha4319773

Bubici, G., and Cirulli, M. (2008). Screening and selection of eggplant and wild related species for resistance to *Leveillula taurica*. *Euphytica* 164, 339–345. doi: 10.1007/s10681-008-9663-z

Cericola, F., Portis, E., Toppino, L., Barchi, L., Acciarri, N., Ciriaci, T., et al. (2013). The population structure and diversity of eggplant from Asia and the Mediterranean basin. *PLOS ONE* 8:e73702. doi: 10.1371/jounal.pone.0073702

Daunay, M. C., and Hazra, P. (2012). "Eggplant," in *Handbook of Vegetables*, eds K. V. Peter and P. Hazra (Houston, TX: Studium Press), 257–322.

Davidar, P., Snow, A. A., Rajkumar, M., Pasquet, R., Daunay, M. C., and Mutegi, E. (2015). The potential for crop to wild hybridization in eggplant (*Solanum melongena*, Solanaceae) in southern India. *Am. J. Bot.* 102, 129–139. doi: 10.3732/ajb.1400404

Devi, C. P., Munshi, A. D., Behera, T. K., Choudhary, H., Vinod, Gurung, B., et al. (2015). Cross compatibility in interspecific hybridization of eggplant, Solanum melongena, with its wild relatives. *Sci. Hort.* 193, 353–358. doi: 10.1016/j.scienta.2015.07.024

Dempewolf, H., Eastwood, R. H., Guarino, L., Khoury, C. K., Müller, J. V., and Toll, J. (2014). Adapting agriculture to climate change: a global initiative to collect, conserve, and use crop wild relatives. *Agroecol. Sustain. Food* 38, 369–377. doi: 10.1080/21683565.2013.870629

Díez, M. J., and Nuez, F. (2008). "Tomato," in *Handbook of Plant Breeding: Vegetables II*, eds J. Prohens and F. Nuez (New York, NY: Springer), 249–323.

Doganlar, S., Frary, A., Daunay, M. C., Lester, R. N., and Tanksley, S. (2002). Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. *Genetics* 161, 1713–1726. doi: 10.1007/s10681-014-1234-x

FAO (2016). *FAOSTAT Production Databases*. Available online at: http://faostat3.fao.org/home/E (Accessed 11 February, 2016).

Gisbert, C., Prohens, J., Raigón, M. D., Stommel, J. R., and Nuez, F. (2011). Eggplant relatives as sources of variation for developing new rootstocks: effects of grafting on eggplant yield and fruit apparent quality and composition. *Sci. Hortic.* 128, 14–22. doi: 10.1016/j.scienta.2010.12.007

Gramazio, P., Prohens, J., Plazas, M., Andújar, I., Herraiz, F. J., Castillo, E., et al. (2014). Location of chlorogenic acid biosynthesis pathway and polyphenol oxidase genes in a new interspecific anchored linkage map in eggplant. *BMC Plant Biol.* 14:350. doi: 10.1186/s12870-014-0350-z

Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* 156, 1–13. doi: 10.1007/s10681-007-9363-0

Hurtado, M., Vilanova, S., Plazas, M., Gramazio, P., Herraiz, F. J., Andújar, I., et al. (2013). Phenomics of fruit shape in eggplant (*Solanum melongena* L.) using Tomato Analyzer software. *Sci. Hortic.* 164, 625–632. doi: 10.1016/j.scienta.2013.10.028

IBPGR (1990). *Descriptors for Eggplant*. Rome: International Board for Plant Genetic Resources.

Knapp, S., Vorontsova, M. S., and Prohens, J. (2013). Wild relatives of the eggplant (Solanum melongena L.: Solanaceae): new understanding of species names in a complex group. *PLoS One* 8:e57039. doi: 10.1371/journal.pone.0057039

Liu, J., Zheng, Z. S., Zhou, X. H., Feng, C., and Zhuang, Y. (2015). Improving the resistance of eggplant (*Solanum melongena*) to Verticillium wilt using wild species Solanum linnaeanum. *Euphytica* 201, 463–469. doi: 10.1007/s10681-014-1234-x

Mennella, G., Lo Scalzo, R., Fibiani, M., D'Alessandro, A., Francese, G., Toppino, L., et al. (2010). Chemical and bioactive quality traits during fruit ripening in eggplant (*S. melongena* L.) and allied species. *J. Agric. Food Chem.* 47, 11821–11831. doi: 10.1021/jf3037424

Meyer, R. S., Karol, K. G., Little, D. P., Nee, M. H., and Litt, A. (2012). Phylogeographic relationships among Asian eggplants and new perspectives on eggplant domestication. *Mol. Phylogenet. Evol.* 63, 685–701. doi: 10.1016/j.ympev.2012.02.006

Meyer, R. S., and Purugganan, M. D. (2014). Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* 14, 840–852. doi: 10.1038/nrg3605

Meyer, R. S., Whitaker, B. D., Little, D. P., Wu, S. B., Kennelly, E. J., Long, C. L., et al. (2015). Parallel reductions in phenolic constituents resulting from the domestication of eggplant. *Phytochemistry* 115, 194–206. doi: 10.1016/j.phytochem.2015.02.006

Muñoz-Falcón, J. E., Prohens, J., Vilanova, S., and Nuez, F. (2009). Diversity in commercial varieties and landraces of black eggplants and implications for broadening the breeders' gene pool. *Ann. Appl. Biol.* 154, 453–465. doi: 10.1111/j.1744-7348.2009.00314.x

Mutegi, E., Snow, A. A., Rajkumar, M., Pasquet, R., Ponniah, H., Daunay, M. C., et al. (2015). Genetic diversity and population structure of wild/weedy eggplant (*Solanum insanum*, Solanaceae) in southern India: implications for conservation. *Am. J. Bot.* 102, 140–148. doi: 10.3732/ajb.1400403

Naegele, R. P., Boyle, S., Quesada-Ocampo, L. M., and Hausbeck, M. K. (2014). Genetic diversity, population structure, and resistance to *Phytophthora capsici* of a worldwide collection of eggplant germplasm. *PLOS ONE* 9:e95930. doi: 10.1371/journal.pone.095930

Niño-Medina, G., Muy-Rangel, D., Gardea-Béjar, A., González-Aguilar, G., Heredia, B., Báez-Sanudo, M., et al. (2014). Nutritional and nutraceutical components of commercial eggplant types grown in Sinaloa, Mexico. *Not. Bot. Horti Agrobot. Cluj Napoca* 42, 538–544. doi: 10.15835/nbha4229573

Plazas, M., Vilanova, S., Gramazio, P., Rodríguez-Burruezo, A., Fita, A., Herraiz, F. J., et al. (2016). Interspecific hybridization between eggplant and wild relatives from different genepools. *J. Am. Soc. Hortic. Sci.* 141, 34–44.

Portis, E., Cericola, F., Barchi, L., Toppino, L., Acciarri, N., Pulcini, L., et al. (2015). Association mapping fro fruit, plant and leaf morphology traits in eggplant. *PLOS ONE* 10:e0135200. doi: 10.1371/jounal.pone.0135200

Prohens, J., Blanca, J. M., and Nuez, F. (2005). Morphological and molecular variation in a collection of eggplants from a secondary center of diversity: implications for conservation and breeding. *J. Am. Soc. Hort. Sci.* 130, 54–63.

Prohens, J., Plazas, M., Raigón, M. D., Seguí-Simarro, J. M., Stommel, J. R., and Vilanova, S. (2012). Characterization of interspecific hybrids and first backcross generations from crosses between two cultivated eggplants (*Solanum melongena* and *S. aethiopicum* Kumba group) and implications for eggplant breeding. *Euphytica* 186, 517–538. doi: 10.1007/s10681-012-0652-x

Prohens, J., Whitaker, B. D., Plazas, M., Vilanova, S., Hurtado, M., Blasco, M., et al. (2013). Genetic diversity in morphological characters and phenolic acids content resulting from an interspecific cross between eggplant, *Solanum melongena*, and its wild ancestor (*S. incanum*). *Ann. Appl. Biol.* 162, 242–257. doi: 10.1111/aab.12017

Rodríguez, G. R., Moyseenko, J. B., Robbins, M. D., Huarachi Morejón, N., Francis, D. M., and van der Knaap, E. (2010). Tomato analyzer: a useful software application to collect accurate and detailed morphological and colorimetric data from two-dimensional objects. *J. Vis. Exp.* 37:e1856. doi: 10.3791/1856

Rodríguez-Burruezo, A., Prohens, J., and Nuez, F. (2008). Performance of hybrids between local varieties of eggplant (*Solanum melongena*) and its relation to the mean of parents and to morphological and genetic distances among parents. *Eur. J. Hortic. Sci.* 73, 76–83.

Rotino, G. L., Sala, T., and Toppino, L. (2014). "Eggplant," in *Alien Gene Transfer in Crop Plants*, Vol. 2, eds A. Pratap and J. Kumar (New York, NY: Springer), 381–409.

Sabatini, E., Beretta, M., Sala, T., Acciarri, N., Milc, J., and Pecchioni, N. (2013). "Molecular breeding," in *Genetics, Genomics and Breeding of Tomato*, eds B. E. Liedl, J. A. Labate, J. R. Stommel, A. Slade, and C. Kole (Boca Raton, FL: CRC Press), 228–303.

San José, R., Sánchez-Mata, M. C., Cámara, M., and Prohens, J. (2014). Eggplant fruit composition as affected by the cultivation environment and genetic constitution. *J. Sci. Food Agric.* 94, 2774–2784. doi: 10.1002/jsfa.6623

Särkinen, T., Bohs, L., Olmstead, R. G., and Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC Evol. Biol.* 13:214. doi: 10.1186/1471-2148-13-214

Sękara, A., and Bieniasz, M. (2008). Pollination, fertilization and fruit formation in eggplant (*Solanum melongena* L.). *Acta Agrobot.* 61, 107–113. doi: 10.5586/aa.2008.014

Syfert, M. M., Castañeda-Álvarez, N. P., Khoury, C. K., Särkinen, T., Sosa, C. C., Achicanoy, H. A., et al. (2016). Crop wild relatives of the brinjal eggplant (*Solanum melongena*): poorly represented in genebanks and many species at risk of extinction. *Am. J. Bot.* 103, 635–651. doi: 10.3732/ajb.1500539

van der Weerden, G. M., and Barendse, G. W. M. (2007). A web-searchable database developed for the EGGNET project and applied to the Radboud University Solanaceae database. *Acta Hortic.* 745, 503–506. doi: 10.17660/actahortic.2007.745.37

Vilanova, S., Manzur, J. P., and Prohens, J. (2012). Development and characterization of genomic SSR markers in eggplant and their application to the study of diversity and relationships in a collection of different cultivar types and origins. *Mol. Breed.* 30, 647–660. doi: 10.1007/s11032-011-9650-2

Vorontsova, M. S., and Knapp, S. (in press). A revision of the spiny solanums, *Solanum* subgenus *Leptostemonum* (Solanaceae) in Africa and Madagascar. *Syst. Bot Monogr.*

Vorontsova, M. S., Stern, S., Bohs, L., and Knapp, S. (2013). African spiny *Solanum* subgenus *Leptostemonum*, Solanaceae): a thorny phylogenetic tangle. *Bot. J. Linn. Soc.* 173, 176–193. doi: 10.1111/boj.12053

A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding

Maria Tattaris[1], Matthew P. Reynolds[1]* and Scott C. Chapman[2]

[1] International Maize and Wheat Improvement Center, Texcoco, Mexico, [2] CSIRO Agriculture, Queensland Bioscience Precinct, Queensland, QLD, Australia

Remote sensing (RS) of plant canopies permits non-intrusive, high-throughput monitoring of plant physiological characteristics. This study compared three RS approaches using a low flying UAV (unmanned aerial vehicle), with that of proximal sensing, and satellite-based imagery. Two physiological traits were considered, canopy temperature (CT) and a vegetation index (NDVI), to determine the most viable approaches for large scale crop genetic improvement. The UAV-based platform achieves plot-level resolution while measuring several hundred plots in one mission via high-resolution thermal and multispectral imagery measured at altitudes of 30–100 m. The satellite measures multispectral imagery from an altitude of 770 km. Information was compared with proximal measurements using IR thermometers and an NDVI sensor at a distance of 0.5–1 m above plots. For robust comparisons, CT and NDVI were assessed on panels of elite cultivars under irrigated and drought conditions, in different thermal regimes, and on un-adapted genetic resources under water deficit. Correlations between airborne data and yield/biomass at maturity were generally higher than equivalent proximal correlations. NDVI was derived from high-resolution satellite imagery for only larger sized plots (8.5 × 2.4 m) due to restricted pixel density. Results support use of UAV-based RS techniques for high-throughput phenotyping for both precision and efficiency.

Keywords: UAV, multispectral, thermal, indices, airborne imagery, high-throughput phenotyping

INTRODUCTION

High-throughput phenotyping, particularly through the application of remote sensing tools, offers a rapid and non-destructive approach to plant screening (White et al., 2012). Recent advances in remote sensing technologies as well as in data processing has increased applications in both field and controlled growing conditions (Leinonen and Jones, 2004; Jones et al., 2007; Möller et al., 2007; Swain and Zaman, 2012; Araus and Cairns, 2014) with important consequences for crop improvement.

Remotely sensed spectral readings are based on the interaction between incoming radiation and target objects, resulting in a characteristic signature of reflected light. Such signatures are typically used to calculate spectral indices, which are a function of the light absorption properties of the plant at given wavelengths (e.g., see Tables 7.1–7.3 in Mullan, 2012

and Table 2 in Zarco-Tejada et al., 2013). Two commonly used traits for high-throughput screening are the Normalized Difference Vegetation Index (NDVI), and canopy temperature (CT). NDVI is calculated using wavelengths within the NIR (near infrared) and VIS (visible) regions of the electromagnetic spectrum. NDVI relates to chlorophyll content due to absorption features of the molecule, and hence the photosynthetic capacity of the plant. CT, which is measured from emitted infra-red radiation, can be used as a tool to indirectly evaluate the transpiration rate of a plant (Berliner et al., 1984; Peñuelas et al., 1992). Based mainly on ground based proximal sensing approaches, CT shows a robust association with plant performance, especially under stress, being intimately associated with water status and stomatal conductance (Blum et al., 1982; Berliner et al., 1984; Amani et al., 1996) while NDVI can estimate relative crop biomass at different growth stages (Babar et al., 2006) as well as N deficiency and crop senescence rate (Blum et al., 1982; Reynolds et al., 1994, 1998; Raun et al., 2001; Babar et al., 2006; Olivares-villegas et al., 2007).

Notwithstanding the examples cited above, proximal remote sensing methods can lose precision at high-throughput due to changes in environmental conditions between the start and end of measurements (typically a time period of one to several hours for breeding trials). Satellite imagery has the advantage of covering large areas instantaneously, but generally does not offer the spatial (sub-meter) and temporal (weekly/daily) resolution required for breeding experiments. Low level, airborne remote sensing measurements have the advantage in that resolution is at plot level while at the same time providing the possibility of instantaneously capturing multiple plots at a practical breeding scale at a high temporal resolution (Araus and Cairns, 2014; Chapman et al., 2014).

While a body of literature has shown the value of airborne derived spectral indices to estimate environmentally determined performance traits for a number of crops (Shanahan et al., 2001; Champagne et al., 2002; González-Dugo et al., 2006; Berni et al., 2009; Zhang et al., 2009; Dupin et al., 2011; Swain and Zaman, 2012; Zarco-Tejada et al., 2012), the use of UAVs to increase throughput for breeding purposes and the focus on genetic effects within one agronomic treatment is relatively new (Lelong et al., 2008; Chapman et al., 2014; Díaz-Varela et al., 2015; Zaman-Allah et al., 2015). The UAV approach has obvious potential to increase throughput but the issue of precision relative to other approaches has not been examined. Moreover, data from UAVs have not been compared with satellite derived imagery for phenotyping applications.

The work presented here aims to demonstrate the potential of low level thermal and multispectral UAV imagery and high-resolution multispectral satellite imagery for the derivation of spectral indices of experiments that comprise of 100 s of plots (**Table 1**) growing in realistic field environments. A methodology was developed with three main objectives. The first was to compare data derived from the UAV with proximal sensors, to determine how well they relate to each other, and their relative ability to predict biomass and yield of wheat. A second objective was to compare data derived from the UAV with satellite imagery for different sized experimental breeding plots, as well as with equivalent data at ground level. The focus of the UAV measurements was the derivation of NDVI and a spectral index relating to canopy temperature, and similarly NDVI and CT were measured using proximal instruments on the ground. For the satellite imagery, NDVI was calculated. A third objective was to evaluate the robustness of the UAV derived indices as selection tools by examining their relationship with crop performance characteristics of different classes of breeding material growing in different simulated target environments. Specifically, the screening traits were measured on advanced breeding lines under optimal, heat stressed, and water deficit conditions, while un-adapted genetic resources were evaluated under water deficit.

MATERIALS AND METHODS

Study Site
Trials were located at an experiment station of the International Maize and Wheat Improvement Centre (CIMMYT) in the Sonoran desert, close to Ciudad Obregon, NW Mexico (27°20′ N; 109°54′ W; and 38 m above sea level). Environmental and management details of this area are given in (Sayre et al., 1997). Five trials made up of elite lines and un-adapted genetic resources were studied in three different environments (**Table 1**); optimal irrigated (OPT), drought stress (DRT) (Gutiérrez-Rodríguez et al., 2004), and hot-irrigated (HOT) (Pinto et al., 2010). The trials named Elite OPT, Elite HOT 1, Elite HOT 2, and Elite DRT are made up of advanced spring wheat lines from CIMMYT adapted to the optimal, hot-irrigated and drought stress environments, respectively. The trial denoted as Gen Res DRT, sown under drought stress, is made up of landraces mainly from Mexico, northern Africa and western central Asia, chosen for potential expression of drought adaptive traits. All trials were sown under an alpha-lattice design, with either two or three replications.

Proximal Data Collection
Grain yield (gm^{-2}) and dry biomass weight (gm^{-2}) were estimated at maturity for each plot following the methods described in Pask et al. (2012) (see **Table 1** for harvest dates). Key phenological stages of emergence, heading, anthesis, and physiological maturity were recorded for each plot (Pask et al., 2012).

Canopy temperature (CT) was recorded at ground level using the Sixth Sense LT300 handheld infrared thermometer. Measurements were made along each of the plots from a distance of ∼0.5 m above canopy, angled to avoid bare soil (about 60° to nadir) and directed specifically at the part of the plot most exposed to the sun (i.e., with the sun behind observer), when cloud cover was minimal and at times of low wind speed (Pask et al., 2012).

NDVI was measured at ground level with the Trimble Greenseeker 505 Hand-Held active sensor. This instrument emits and measures light at 656 and 774 nm. Measurements were made close to noon, when the plant canopy and soil surface are dry, at about 0.5 m horizontally above the canopy such that the FOV

Trial	Germplasm	Env	Plot size (m)	No. of lines/reps	Sowing date	Harvest date	Variable	Alternative remote sensing approach		
								Proximal date	UAV date	Satellite date
Elite OPT	Elite	OPT	8.5 × 2.4	27/3	Nov 2012	May 2013	NDVI	11/3/13	11/3/13	6/4/13
								26/3/13	25/3/13	
Elite HOT 1	Elite	HOT	2.0 × 0.8	30/2	Feb 2012	May 2012	NDVI	15/5/12	11/5/12	
							CT	14/5/12	17/5/12	
Elite HOT 2	Elite	HOT	2.0 × 0.8	60/2	March 2014		NDVI	23/5/14	15/5/14	
								1/6/14	3/6/14	
							CT	13/5/14	13/5/14	
								15/5/14	15/5/14	
								20/5/14	16/5/14	
Elite DRT	Elite	DRT	2.0 × 0.8	50/3	Dec 2012	June 2013	NDVI	14/2/13	13/2/13	
								7/3/13	4/3/13	
							CT	18/2/14	21/2/14	
Gen Res DRT	Un-adapted genetic resources	DRT	2.0 × 0.8	208/2	Dec 2012	June 2013	NDVI	21/2/13	26/2/13	
							CT	31/3/13	25/3/13	
								7/2/13	22/2/13	

Harvest date indicates the approximate date at which harvest was made for yield and biomass estimates. Measurement dates of ground-based, UAV, and satellite data used for comparisons.

is directly above the plot and centered over the middle row (Pask et al., 2012). NDVI allows for the estimation of vegetation present in each measurement via Equation 1 (Rouse et al., 1973):

$$NDVI = \frac{NIR - R}{NIR + R} \qquad (1)$$

where *NIR* and *R* are the measured reflectance in the NIR and red spectral bands respectively, (774 and 656 nm for the case of the Greenseeker). **Table 1** details the measurement dates for the proximal instruments.

UAV Data Collection

Aerial imagery was collected via the AscTec Falcon 8 Unmanned Aerial Vehicle (UAV) (**Figure 1**). The 8-rotor UAV has a maximum 750 g payload; hence it has the ability to fly small, lightweight instruments. The flight system includes an on-board in-built GPS and a Mobile Ground Station (**Figure 1**, inset). Aerial images were collected with two cameras mounted separately on the UAV; the Tetracam ADC Lite multispectral camera (2048 × 1536 pixels for Red Green and NIR bands together) and the FLIR Tau 640 LWIR uncooled thermal imaging camera (640 × 512 pixels). See **Table 2** for specifications of the cameras. An 8000 mAh lithium battery powers the UAV and cameras, providing ∼15-min flight time. Several batteries allow for multiple flights in one session. The ADC Lite Tetracam takes photos in the green, red and NIR regions of the electromagnetic spectrum (**Figure 2A**), allowing for the calculation of NDVI,


FIGURE 1 | The airborne remote sensing platform used in this study: The AscTec Falcon 8 Unmanned Aerial Vehicle (UAV), operated with the Mobile Ground Station (inset).

while the FLIR thermal camera is used to derive a thermal index relating to the CT of the target plots. In the specification used, the thermal camera records analog video (integrated over 7.5–13 μm), which is subsequently converted to still images for processing (**Figure 2B**).

Instrument	Dims (in)	Weight (g)	Resolution	Spectral range	Lens (mm)	Pixel size (μm)
Tetracam ADC lite multispectral camera	4.5 × 3.0 × 2.38	200	2048 × 1536	Green, red, and NIR (TM2, TM3, and TM4)	8.0	3.2
FLIR Tau 640 LWIR uncooled thermal imaging camera	1.74 × 1.75 × 1.18	<72	640 × 512	7.5–13 μm	25	17
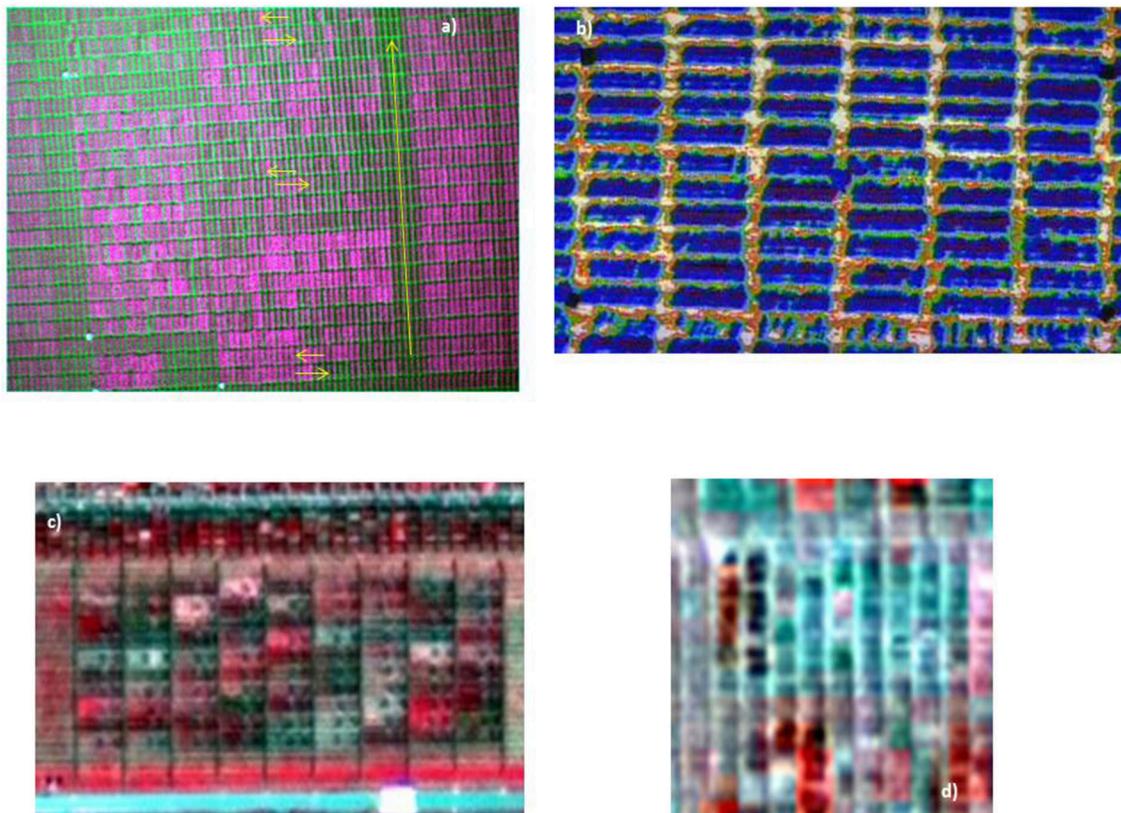
FIGURE 2 | (A). Raw image of Gen Res DRT trial within the drought environment, taken using the ADC Lite Tetracam on the UAV, approximately at 100 m height. Ground dimensions of plots are 2 × 0.8 m, with arrows representing direction of proximal measurements. Assuming a measurement time of 10 s per plot, the time taken to complete measurements using proximal sensors is ∼69 min for this trial, compared to several seconds with the UAV. **(B)** Raw image of a "HOT" trial extracted from video footage from the FLIR Tau thermal camera. Flight altitude was ∼30 m. Ground dimensions of plots are 2 × 0.8 m. **(C)** Pan-sharpened WV-2 imagery of Elite OPT. Pan-sharpened imagery of a trial containing smaller sized plots in **(D)** did not allow for the extraction of NDVI as plots were mixed within pixels.

Satellite Imagery Description

Satellite imagery was obtained from the commercial Digital Globe WorldView-2 (WV-2) satellite, taken on 6th April 2013. The imagery includes an 8-band multispectral image (bands between 396 and 1043 nm) and a panchromatic image (447–808 nm), with a spatial resolution of 0.46 and 1.85 m respectively. The georeferenced, unprocessed images cover ∼25 km^2, including the whole of the CIMMYT research station.

Table 1 presents the measurement dates for the relevant remotely sensed data for each of the trials. Proximal dates were compared based on closest dates available between proximal and airborne data collection.

Image Processing and Analysis
UAV Imagery

Processing was carried out using ENVI version 5.0 (Exelis Visual Information Solutions, Boulder, Colorado). Radiometric distortions, e.g., lens vignetting, were solved by applying a cross track illumination correction, removing any broadband variation without affecting narrowband features. Geometric distortions are corrected using a "warping" procedure, by which an image is resampled to match the geometry of a "base" image or a vector map via the selection of Ground Control Points (GCPs). Images are subsequently mosaicked together by identifying overlapping regions within images.
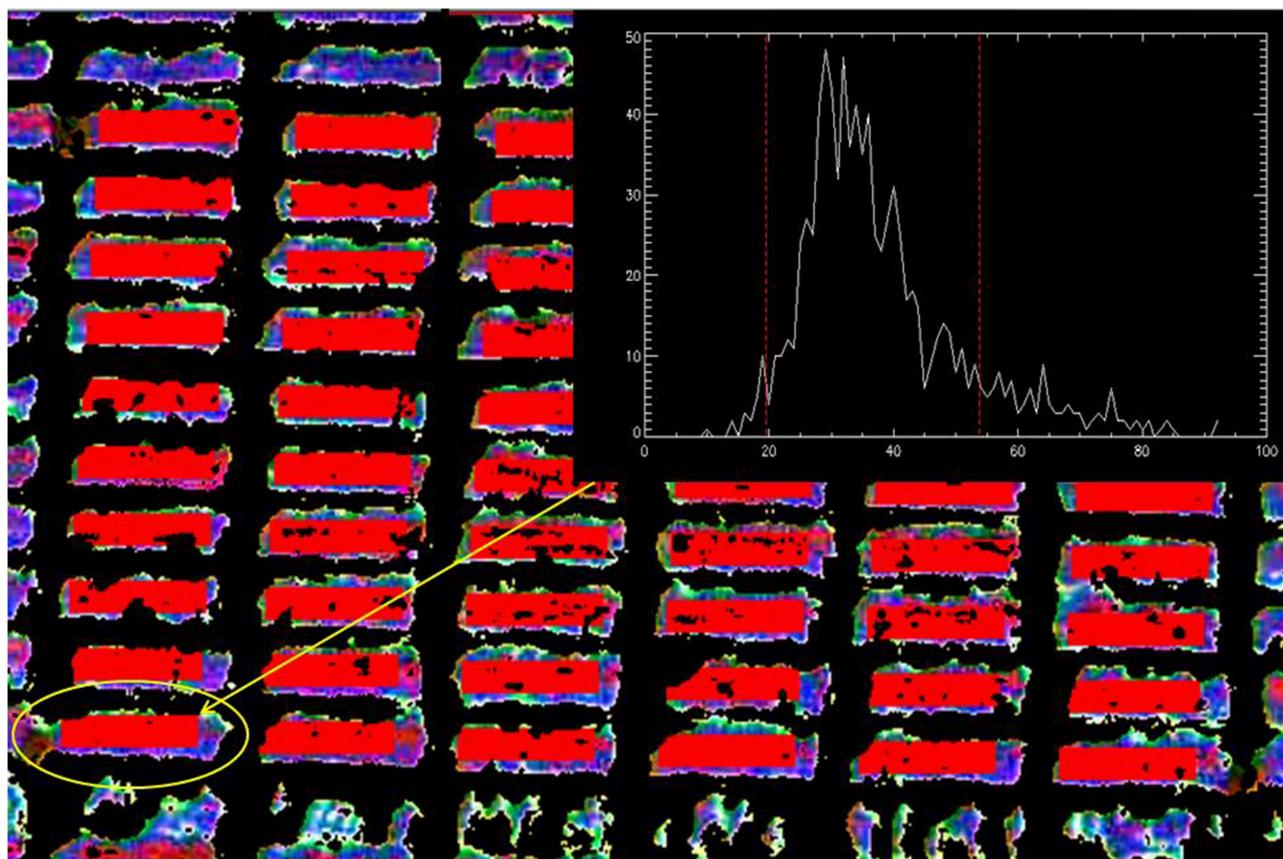
FIGURE 3 | Example of the image processing using UAV-mounted FLIR Tau image of "HOT" trial shown in Figure 2B where a mask is applied to remove any non-vegetation pixels by applying a threshold for each pixel value. This is followed by the detection of each plot using pre-defined location parameters (red rectangles) and the removal of high variance pixels (using histogram of the pixel values of each plot). An average of pixel values over each band is taken to get a value per band per plot. This value is then subsequently used to calculate the target indices.

A camera specific mask is applied to the image/mosaic of each trial, via pixel band ratio thresholds, to differentiate between vegetation and non-vegetation pixels, aiming to remove any non-vegetation pixels, such as soil, as different materials can be distinguished by their pixel signal. An algorithm is then applied for the automatic detection of plots using pre-defined parameters by the user, for example the plot size in pixels, and distance between plots in pixels. An average across all bands for each pixel is calculated and pixels within each plot that exhibit high variance are removed, to eliminate any non-vegetation pixels that the mask may have missed, as well as pixel mixing effects. The average of each plot at each band is then taken to derive the target indices at plot level (**Figure 3**). For the ADC Lite multispectral camera, the NDVI index is calculated as, $\frac{TM4-TM3}{TM4+TM3}$, where $TM4$ ($\approx$ 760–900 nm) and $TM3$ ($\approx$ 630–690 nm) denote the Landsat bands.

The processed images collected from the thermal camera onboard the UAV were used to derive a temperature index, which relates to the CT of each plot. The temperature index T_I was calculated using the sum of the green and blue bands of the plot averaged values of the processed images acquired from the recorded analog video:

$$T_I = T_G + T_B \qquad (2)$$

where T_G and T_B are the averaged "plot" values at the green and blue bands respectively.

Satellite Imagery

The Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH), an ENVI atmospheric correction tool, was applied to the satellite imagery used here. FLAASH incorporates the Moderate Resolution Atmospheric Transmission V4 (MODTRAN4) radiative transfer model to simulate the spectral radiance at pixel sensor level using user defined input variables. For a detailed description on the methods used by FLAASH see Adler-Golden et al. (1999).

The panchromatic (spatial resolution 0.46 m) and the multispectral imagery (spatial resolution 1.8 m) were fused together to create a single high resolution multispectral image, via the ESRI "pan-sharpening" algorithm, using ESRI ArcMap 10.1 (Mishra and Zhang, 2013). The pan-sharpened image, with

a spatial resolution of 0.46 m, was used to derive NDVI (using 833 nm and 659 nm as the NIR and red wavelengths, respectively) as this can then be compared with the NDVI derived from the UAV (ADC Lite), and proximal (GreenSeeker) measurements.

Statistical Analysis

All satellite, airborne and proximal data was spatially corrected for row and column variation across the experiments using Multi Environmental Trial Analysis (META) for SAS (Vargas et al., 2013). Adjusted means were computed individually at each measurement date for the given traits based on the lattice design of the trials. Phenotypic correlations among adjusted genotype means per trial per date were determined to compare the relationship between the airborne, proximal, and agronomic traits. When more than one reading of data was available for both airborne and proximal data, multiple correlations are presented. Statistical analysis was carried out using R 3.1.2 (R Core Team, 2014). Differences between the phenotypic correlations of the proximal and UAV indices against yield or biomass were tested for significance with a Student t-test. The Holm-Bonferroni method was applied to the p-values of the t-tests to account for multiple comparisons (Holm, 1979). In addition, in order to investigate the interactions of the UAV and proximal phenotypic

correlations against yield and biomass under the three different environments, for the CT/thermal index and NDVI, a multi-factor analysis of variance (ANOVA) was performed using R.3.1.2.

RESULTS

Comparison of Data from Airborne and Proximal Sensing Approaches

Table 3 shows phenotypic correlations between airborne and proximal sensed data, using mean values of genotypes for both the thermal index and NDVI measured from the UAV compared with the equivalent traits -CT and NDVI- measured using proximal sensors. Correlations between the UAV derived thermal index and the proximal CT are, in general, significant for all trials. This adds confidence to the use of the airborne thermal index. Difficulties can arise when comparing the two different methodologies due to the sensitivity of CT to external environmental factors (time of day, temperature, radiation, wind, irrigation status, VPD, etc.), particularly wind speed. Nonetheless of all the comparisons made, only one did not show a correlation, that between CT and the thermal index for the trial Elite DRT, probably due to variations in wind speed during measurements

TABLE 3 | Phenotypic correlations between genotype means for the airborne/satellite derived thermal index/NDVI, against the corresponding ground-based CT/NDVI and between the genotypic means for the aerial derived indices and yield/biomass.

Trial	ENV	Variable	Alternative remote sensing approach		Phenotypic correlation between methodologies	Phenotypic correlation with yield		Phenotypic correlation with biomass	
			Proximal date	UAV date	Proximal VS. UAV	UAV VS. yield	Proximal VS. yield	UAV VS. biomass	Proximal VS. biomass
Elite OPT	OPT	NDVI	11/3/13	11/3/13	0.78**	0.36+	0.37+	0.45*	0.39*
			26/3/13	25/3/13	0.79**	0.29+	0.38+	0.52**	0.41*
Elite HOT 1	HOT	NDVI	15/5/12	11/5/12	0.85**	0.75**	0.52**	0.79**	0.58**
		CT	14/5/12	17/5/12	0.78**	−0.73**	−0.56**	−0.78**	−0.6**
Elite HOT 2	HOT	NDVI	23/5/14	15/5/14	0.86**	0.51**	0.54**	0.59**	0.65**
			1/6/14	3/6/14	0.87**	0.64**	0.64**	0.74**	0.73**
		CT	13/5/14	13/5/14	0.36**	−0.45**	−0.34**	−0.56**	−0.37**
			15/5/14	15/5/14	0.4**	−0.57**	−0.43**	−0.57**	−0.49**
			20/5/14	16/5/14	0.75**	−0.62**	−0.48**	−0.62**	−0.55**
Elite DRT	DRT	NDVI	14/2/13	13/2/13	0.41**	0.56**	0.1	–	–
			7/3/13	4/3/13	0.80**	0.42**	0.27+	–	–
		CT	18/2/14	21/2/14	−0.04	−0.41**	−0.24+	–	–
Gen Res DRT	DRT	NDVI	21/2/13	26/2/13	0.72**	0.16**	0.25**	0.62**	0.46**
			31/3/13	25/3/13	0.91**	0.21**	0.14*	0.72**	0.69**
		CT	7/2/13	22/2/13	0.57**	−0.44**	−0.25**	0.1	0.18+
SATELLITE									
Trial	ENV	Variable	Proximal date	UAV date	Satellite date	SAT VS. proximal	SAT VS. UAV	SAT VS. yield	SAT VS. biomass
Elite OPT	OPT	NDVI	26/3/13	25/3/13	6/4/13	0.85**	0.84**	0.53**	0.58**

Also shown are equivalent correlations with ground-based indices. $^{+}$, *, ** *represent significant levels of 0.1, 0.05, and 0.01 respectively.*

(as was noted at the time of observation). Significant correlations were also observed between the UAV and proximal NDVI measurements for all trials.

Association of Traits with Yield and Biomass Comparing Airborne and Proximal Sensing Approaches

Phenotypic correlations were estimated between UAV derived NDVI and the thermal index with both biomass and yield of genotypes (**Table 3**). For comparison, the corresponding correlation between proximal NDVI and CT with yield and biomass is also shown. Correlations between the UAV derived thermal index and yield/biomass are significant for almost all trials, and are generally larger than the corresponding proximal CT correlations with the yield and biomass. Note that negative correlations were observed between CT/UAV thermal index and yield/biomass as cooler canopies are generally associated with better adaptation. Similarly, the UAV derived NDVI index generally shows stronger correlations with biomass and yield compared with the respective proximal NDVI.

The "Gen Res DRT" trial is made up of diverse genetic resources expressing non-homogeneous height and which are not necessarily well adapted to the photoperiod and other conditions of the screening environment. This could help explain the relatively lower, although still significant, correlations between NDVI and yield for this trial compared to the hot irrigated and drought data sets from elite material, i.e., there was large variation in development stage and morphology (**Table 3**). The lower correlation between CT and biomass for this trial compared to yield could also be attributed to confounding effects due to variations in height within the trial and their attendant influence on boundary air layers that affect transpiration rate when there is no breeze. Note that this is not the case for NDVI, which is free from such confounding effects.

For the Elite OPT trial, correlations with yield and NDVI are of lower significance ($p < 0.1$) compared to those of the hot irrigated trials, also made up of elite lines (**Table 3**). However, these results are consistent with previous observations that these techniques are most effective as a selection tools under abiotic stress (Pinto et al., 2010).

When considered together, correlations between the UAV derived indices and yield/biomass were significantly different to the equivalent proximal sensed correlations (t-test, $P = 0.01$). When separated into groups, there was significant difference between the UAV and proximal derived correlations with yield and biomass for the following groups CT/thermal index (t-test, $P = 0.01$), yield (t-test, $P = 0.05$), biomass (t-test, $P = 0.05$), and NDVI (t-test, $P = 0.1$).

The phenotypic correlations were separated into the three environments (OPT, DRT, and HOT) and a multi-factor ANOVA was performed. For the OPT environment, a significant interaction ($P = 0.06$) was observed between the proximal and UAV phenotypic correlations between yield and biomass, probably associated with the greater biomass correlations, particularly those of the UAV (see **Table 3**). For the HOT trials, there was a significant difference ($P = 0.08$) between the

proximal and UAV phenotypic correlations between yield and biomass for both NDVI and the thermal index/CT. This can be attributed to the higher correlations for the UAV observations. No interaction was observed for the DRT environment, this can be partly explained by un-adapted material in the genetic resource trial, as explained above.

Satellite Imagery

Figure 2C shows the pan-sharpened WV-2 extracted image of Elite OPT. It can be seen that the satellite image provides sufficient resolution for multiple (~20) pixels for each plot and hence the NDVI index was able to be calculated.

Table 3 compares the NDVI calculated from the three methods: space-borne collected WV-2 imagery, low level airborne collected imagery via the UAV and proximal measurements. The proximal and UAV measurements were chosen to be as close as possible to the satellite imagery collection date. It can be seen that the NDVI derived from all methods are well correlated with each other. The correlation between the NDVI from the satellite image and the NDVI from the other two methods gives confidence to the calculation of NDVI from high resolution satellite imagery for plots of the size of those in Elite OPT trial (8.5 × 2.4 m). Also compared in **Table 3** is the relationship between each of the NDVI indices and the dry biomass weight and yield measured at maturity for Elite OPT. The NDVI derived from the satellite provide the best correlation with biomass and yield, and proximal NDVI the lowest.

An attempt was made to retrieve NDVI from trials of smaller sized plots. **Figure 2D** shows an extract of the pan-sharpened WV-2 imagery from an OPT trial with plot size at 2 m × 0.8 m. The resolution of the image prevented the separation of plots due to pixel mixing; hence it was not possible to distinguish between plots.

DISCUSSION

The results of the current study demonstrate the advantage of airborne remote sensing as a tool to estimate a range of physiological and agronomic traits on a large scale in experimental plots. Proximal measurements have already been proven to predict yield and biomass in wheat (Reynolds et al., 1994, 1998; Aparicio et al., 2000) and are beginning to be used routinely in breeding (Pask et al., 2014). The generally strong correlations presented here between airborne indices and equivalent ground-based CT and NDVI, as well as significant correlations between the airborne indices and yield/biomass, that were generally greater than the equivalent correlations with ground-based measurements, suggest that increased precision results from the use of the indices derived from imagery, particularly in the stressed environments. This is a promising result given the impacts of changing climate and its implications for food security.

Most published work that attempts to thoroughly validate multispectral indices is based on proximal measurements. Errors may be introduced when moving from proximal to aerial measurements at a spatially larger scale, for example atmospheric scattering may cause absorption features of light by pigments

to alter, as well as affects related to canopy architecture (angle and area of leaves), water vapor in the atmosphere, background noise and measurement geometry (Suarez et al., 2008; Garbulsky et al., 2011). However, the results presented here demonstrate the potential of low level UAV measurements to indirectly measure yield and biomass in field conditions. The relative precision of airborne measurements can be associated with two main factors. The first is related to reduced errors linked to the ability to remove non-vegetation pixels and other statistical outliers during image analysis (**Figure 3**). The second is through limiting confounding effects caused by environmental drift, such as changes in temperature, sun angle etc., typically associated with the time taken to make ground based measurements on large trials (**Figure 2**).

Given that the operation of UAVs is less labor intensive than proximal readings, as well as being free from restrictions associated with access to plots (due to irrigation or application of pesticides, for example), the approach lends itself well to routine measurements including for growth analysis, to measure the evolution of stress, and the application of regression e.g., (Lopes and Reynolds, 2012) or spline (Hurtado et al., 2011) models over time from which additional parameters can be derived to compare treatments and genotypes.

Despite the promising results presented here for NDVI derived from the satellite measurements, it is probably not the most effective tool for this application. While satellite imagery has the advantage of covering vast areas, resolution restricts its application to measurements in which target objects are of a larger scale than the small plots typical of genotypic screening. Furthermore, it is difficult to obtain satellite imagery at frequent time intervals and the option to adjust timing of measurement to avoid cloud cover or other inclement weather conditions is absent.

The fact that the estimates of CT and NDVI were generally better associated with performance traits when measured by UAVs compared to proximal data, under both heat and drought stressed conditions, and in advanced lines as well as unimproved genetic backgrounds, confirms the value of the UAV approach in breeding for climate change, where a new generation of breeding lines must be developed based on extensive screening of plant genetic resources.

AUTHOR CONTRIBUTIONS

MT collected data, carried out the data analysis, drafted the article and carried out revisions. MR created the experimental design, made suggestions for the data analysis, helped with the draft and carried out critical revisions. SC made suggestions for the data analysis, helped with the draft and carried out critical revisions. All authors gave final approval for publication.

ACKNOWLEDGMENTS

REFERENCES

Adler-Golden, S. M., Matthew, M. W., Bernstein, L. S., Levine, R. Y., Berk, A., Richtsmeier, S. C., et al. (1999). Atmospheric correction for shortwave spectral imagery based on MODTRAN4. *Proc. SPIE 3753 Imaging Spectrometry V.* 3753, 61–69. doi: 10.1117/12.366315

Amani, I., Fischer, A., and Reynolds, M. P. (1996). Canopy temperature dpression association with yield of irrigated wheat cultivars in a hot climate. *J. Agron. Crop Sci.* 176, 119–129. doi: 10.1111/j.1439-037X.1996.tb00454.x

Aparicio, N., Villegas, D., Casadesus, J., Araus, J. L., and Royo, C. (2000). Spectral vegetation indices as nondestructive tools for determining durum wheat yield. *Agron. J.* 92, 83–91. doi: 10.2134/agronj2000.92183x

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Babar, M. A., Reynolds, M. P., Van Ginkel, M., Klatt, A. R., Raun, W. R., and Stone, M. L. (2006). Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. *Crop Sci.* 46, 1046–1057. doi: 10.2135/cropsci2005.0211

Berliner, P., Oosterhuis, D. M., and Green, G. C. (1984). Evaluation of the infrared thermometer as a crop stress detector. *Agric. For. Meteorol.* 31, 219–230. doi: 10.1016/0168-1923(84)90036-4

Berni, J. A. J., Zarco-tejada, P. P. J., Suarez, L., Fereres, E., Member, S., and Suárez, L. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Trans. Geosci. Remote Sens.* 47, 722–738. doi: 10.1109/TGRS.2008.2010457

Blum, A., Mayer, J., and Gozlan, G. (1982). Infrared thermal sensing of plant canopies as a screening technique for dehydration avoidance in wheat. *Field Crops Res.* 5, 137–146. doi: 10.1016/0378-4290(82)90014-4

Champagne, C. M., Staenz, K., Bannari, A., White, H. P., Deguise, J., and McNairn, H. (2002). "Estimation of plant water content of agricultural canopies using hyperspectral remote sensing," in *1st International Symposium on Recent Advances in Quantitative Remote Sensing* (Valencia).

Chapman, S. C., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, M. F., et al. (2014). Pheno-copter: a low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping. *Agronomy* 4, 279–301. doi: 10.3390/agronomy4020279

Díaz-Varela, R., de la Rosa, R., León, L., and Zarco-Tejada, P. (2015). High-resolution airborne UAV imagery to assess olive tree crown parameters using 3D photo reconstruction: application in breeding trials. *Remote Sens.* 7, 4213–4232. doi: 10.3390/rs70404213

Dupin, S., Gobrecht, A., and Tisseyre, B. (2011). "Airborne thermography of vines canopy: effect of the atmosphere and mixed pixels on observed canopy temperature," in *8 ème Conférence Européenne sur l'Agriculture de Précision* (Prague), 1–9.

Garbulsky, M. F., Peñuelas, J., Gamon, J., Inoue, Y., and Filella, I. (2011). The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies. A review and meta-analysis. *Remote Sens. Environ.* 115, 281–297. doi: 10.1016/j.rse.2010.08.023

González-Dugo, M. P., Moran, M. S., Mateos, L., and Bryant, R. (2006). Canopy temperature variability as an indicator of crop water stress severity. *Irrigation Sci.* 24, 233–240. doi: 10.1007/s00271-005-0022-8

Gutiérrez-Rodríguez, M., Reynolds, M. P., Escalante-Estrada, J. A., and Rodríguez-González, M. T. (2004). Association between canopy reflectance indices and yield and physiological traits in bread wheat under drought and well-irrigated conditions. *Aust. J. Agric. Res.* 55, 1139–1147. doi: 10.1071/AR04214

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Hurtado, P. X., Schnabel, S. K., Zaban, A., Vetelainen, M., Virtanen, E., Eilers, P. H. C., et al. (2011). Dynamics of senescence-related related QTLs in potato. *Euphytica* 183, 289–302. doi: 10.1007/s10681-011-0464-4

Jones, C. L., Weckler, P. R., Maness, N. O., Jayasekara, R., Stone, M. L., and Chrz, D. (2007). Remote censing to estimate chlorophyll concentration in spinach using multi-spectral plant reflectance. *Am. Soc. Agric. Biol. Eng.* 50, 2267–2273. doi: 10.13031/2013.24079

Leinonen, I., and Jones, H. G. (2004). Combining thermal and visible imagery for estimating canopy temperature and identifying plant stress. *J. Exp. Bot.* 55, 1423–1431. doi: 10.1093/jxb/erh146

Lelong, C. C. D., Burger, P., Jubelin, G., Roux, B., Labbé, S., and Baret, F. (2008). Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. *Sensors* 8, 3557–3585. doi: 10.3390/s8053557

Lopes, M. S., and Reynolds, M. P. (2012). Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. *J. Exp. Bot.* 63, 3789–3798. doi: 10.1093/jxb/ers071

Mishra, R. K., and Zhang, Y. (2013). "A comparison of commercial Pan-sharpening techniques for HR Satellite imagery," in *2013 Esri International User Conference (Esri UC)* (San Diego, CA).

Möller, M., Alchanatis, V., Cohen, Y., Meron, M., Tsipris, J., Naor, A., et al. (2007). Use of thermal and visible imagery for estimating crop water status of irrigated grapevine. *J. Exp. Bot.* 58, 827–838. doi: 10.1093/jxb/erl115

Mullan, D. (2012). "Spectral radiometry," in *Physiological Breeding I: Interdisciplinary Approaches to Improve Crop Adaptation,* eds M. P. Reynolds, A. Pask, and D. Mullan (CIMMYT), 69–80.

Olivares-villegas, J. J., Reynolds, M. P., and McDonald, G. K. (2007). Drought-adaptive attributes in the Seri/Babax hexaploid wheat population. *Funct. Plant Biol.* 34, 189–203. doi: 10.1071/FP06148

Pask, A., Joshi, A. K., Manes, Y., Sharma, I., Chatrath, R., Singh, G. P., et al. (2014). A wheat phenotyping network to incorporate physiological traits for climate change in South Asia. *Field Crops Res.* 168, 156–167. doi: 10.1016/j.fcr.2014.07.004

Pask, A., Pietragalla, J., and Mullan, D. (2012). *Physiological Breeding II: A Field Guide to Wheat Phenotyping.* Mexico: CIMMYT.

Peñuelas, J., Savé, R., Marfa, O., and Serrano, L. (1992). Remotely measured canopy temperature of greenhouse strawberries as indicator of water status and yield under mild and very mild water stress condictions. *Agric. For. Meteorol.* 58, 63–77.

Pinto, R. S., Reynolds, M. P., Mathews, K. L., McIntyre, C. L., Olivares-Villegas, J. J., and Chapman, S. C. (2010). Heat and drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. *Theor. Appl. Genet.* 121, 1001–1021. doi: 10.1007/s00122-010-1351-4

Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., Lukina, E. V., Thomason, W. E., et al. (2001). In-season prediction of potential grain yield in winter wheat using canopy reflectance. *Agron. J.* 93, 131. doi: 10.2134/agronj2001.931131x

R Core Team (2014). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing. Available online at: http://www.R-project.org/.

Reynolds, M. P., Balota, M., Delgado, M. I. B., Amani, I., and Fischer, R. A. (1994). Physiological and morphological traits associated with spring wheat yield under hot, irrigated conditions. *Aust. J. Plant Physiol.* 21, 717–730. doi: 10.1071/PP9940717

Reynolds, M. P., Singh, R. P., Ibrahim, A., Ageeb, O. A. A., Larque-Saavedra, A., and Quick, J. S. (1998). Evaluating physiological traits to complement empirical selection for wheat. *Euphytica* 100, 85–94. doi: 10.1023/A:1018355906553

Rouse, J., Haas, R., Schell, J., and Deering, D. (1973). "Monitoring vegetation system in the great plains with ERTs," in *Third ERTS Symposium* (Washington, DC), 309–317.

Sayre, K., Rajaram, S., and Fischer, R. A. (1997). Yield potential progress in short bread wheats in Northwest Mexico. *Crop Sci.* 37, 36–42. doi: 10.2135/cropsci1997.0011183X003700010006x

Shanahan, J. F., Schepers, J. S., Francis, D. D., Varvel, G. E., and Wilhelm, W. (2001). Use of remote-sensing imagery to estimate corn grain yield. *Agronomy* 93, 583–589. doi: 10.2134/agronj2001.933583x

Suarez, L., Zarco-Tejada, P. J., Sepulcre-Cantó, G., Perez-Priego, O., Miller, J. R., Jiménez-Muñoz, J. C., et al. (2008). Assesing canopy PRI for water stress detection with diurnal airborne imagery. *Remote Sens. Environ.* 112, 560–575. doi: 10.1016/j.rse.2007.05.009

Swain, K. C., and Zaman, Q. U. (2012). Rice "Crop monitoring with unmanned helicopter remote sensing images," in *Remote Sensing of Biomass - Principles and Applications*, ed T. Fatoyinbo (Rijeka: InTech), 254–272.

Vargas, M., Combs, E., Alvarado, G., Atlin, G., Mathews, K., and Crossa, J. (2013). Meta: a suite of sas programs to analyze multienvironment breeding trials. *Agron. J.* 105, 11–19. doi: 10.2134/agronj2012.0016

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crops Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003

Zaman-Allah, M., Vergara, O., Araus, J. L., Tarekegne, A., Magorokosho, C., Zarco-Tejada, P. J., et al. (2015). Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* 11, 35. doi: 10.1186/s13007-015-0078-2

Zarco-Tejada, P. J., González-Dugo, V., and Berni, J. A. J. (2012). Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sens. Environ.* 117, 322–337. doi: 10.1016/j.rse.2011.10.007

Zarco-Tejada, P. J., Morales, A., Testi, L., and Villalobos, F. J. (2013). Spatio-temporal patterns of chlorophyll fluorescence and physiological and structural indices acquired from hyperspectral imagery as compared with carbon fluxes measured with eddy covariance. *Remote Sens. Environ.* 133, 102–115. doi: 10.1016/j.rse.2013.02.003

Zhang, H., Lan, Y., Lacey, R. E., Hoffmann, W. C., and Huang, Y. (2009). Analysis of vegetation indices derived from aerial multispectral and ground hyperspectral data. *Int. J. Agric. Biol. Eng.* 2, 33–40. doi: 10.3965/j.issn.1934-6344.2009.03.033-040

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

A Novel Remote Sensing Approach for Prediction of Maize Yield Under Different Conditions of Nitrogen Fertilization

Omar Vergara-Díaz [1], Mainassara A. Zaman-Allah [2], Benhildah Masuka [2], Alberto Hornero [3], Pablo Zarco-Tejada [3], Boddupalli M. Prasanna [2], Jill E. Cairns [2] and José L. Araus [1]*

[1] Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, Barcelona, Spain, [2] International Maize and Wheat Improvement Center, CIMMYT Southern Africa Regional Office, Harare, Zimbabwe, [3] Laboratory for Research Methods in Quantitative Remote Sensing, QuantaLab, Institute for Sustainable Agriculture, National Research Council, Cordoba, Spain

Maize crop production is constrained worldwide by nitrogen (N) availability and particularly in poor tropical and subtropical soils. The development of affordable high-throughput crop monitoring and phenotyping techniques is key to improving maize cultivation under low-N fertilization. In this study several vegetation indices (VIs) derived from Red-Green-Blue (RGB) digital images at the leaf and canopy levels are proposed as low-cost tools for plant breeding and fertilization management. They were compared with the performance of the normalized difference vegetation index (NDVI) measured at ground level and from an aerial platform, as well as with leaf chlorophyll content (LCC) and other leaf composition and structural parameters at flowering stage. A set of 10 hybrids grown under five different nitrogen regimes and adequate water conditions were tested at the CIMMYT station of Harare (Zimbabwe). Grain yield and leaf N concentration across N fertilization levels were strongly predicted by most of these RGB indices (with $R^2 \sim 0.7$), outperforming the prediction power of the NDVI and LCC. RGB indices also outperformed the NDVI when assessing genotypic differences in grain yield and leaf N concentration within a given level of N fertilization. The best predictor of leaf N concentration across the five N regimes was LCC but its performance within N treatments was inefficient. The leaf traits evaluated also seemed inefficient as phenotyping parameters. It is concluded that the adoption of RGB-based phenotyping techniques may significantly contribute to the progress of plant breeding and the appropriate management of fertilization.

Keywords: breeding, crop management, field phenotyping, maize, nitrogen fertilization, NDVI, RGB indices

INTRODUCTION

Low soil fertility, alongside drought and heat, is a major stress factor limiting crop productivity on a world scale (Stewart et al., 2005). In the case of sub-Saharan Africa, the lack of nitrogen (N) is the main constraint on cereal yields in areas with more than 400 mm average annual rainfall (Buerkert et al., 2001). Therefore, an optimization of N use is critical for increased grain production, especially in the low productive regions. On the other hand, on the basis of environmental and economic

sustainability, a more restricted and reasonable use of fertilizers is necessary. Plant scientists, especially breeders and agronomists, face the challenge of solving these limitations while taking into account the additional implications of climate change on food security (Cairns et al., 2012, 2013).

Maize is the second most cultivated cereal worldwide and the most commonly cultivated cereal in Africa in terms of land area and production (FAO, 2013). In particular, agricultural productivity of sub-Saharan Africa remains the lowest in the world partly due to low soil fertility (Cairns et al., 2013; Fischer et al., 2014). Therefore, improving tolerance of maize to low N will increase yields and impact positively on livelihoods and food security (Masuka et al., 2012).

In this sense, two strategies are considered paramount for crop scientists: (i) breeding to improve varieties toward higher nutrient use efficiency and tolerance to nutrient-deficiency (ii) and appropriate fertilization management (Wezel et al., 2014), including precision agriculture (PA; Hatfield, 2000; Chen et al., 2014). Thereby, the implementation of such improvements may increase farmers' profits by maintaining crop yield and reducing the use of resources while preventing further degradation to the environment (Hergert et al., 1996; Delgado et al., 2001; Roberts et al., 2001; Wang et al., 2003). In that sense, technologies for crop monitoring and breeding must be high performing, broad-use and affordable, particularly (but not only) when national agricultural systems, seed companies, or small farmers from developing countries are the targets. Moreover, in the case of breeding, improvements are needed to overcome the field phenotyping bottleneck that limits breeding and advances in PA (Araus et al., 2008; Furbank and Tester, 2011; Araus and Cairns, 2014).

Remote proximal sensing technologies are being used currently for precise management of crops, whereas its potential application for field high throughput phenotyping has gathered increasing interest in recent years (Araus and Cairns, 2014; Liebisch et al., 2015). The classical approach has involved the use of multispectral sensors and the development of numerous vegetation indices associated with vegetation parameters such as above-ground biomass, water and nutrient-deficiency and crop yield (Petropoulos and Kalaitzidisz, 2012). Among the indices, the Normalized Difference Vegetation Index (NDVI) is the most widely used. Concerning crop N performance, several studies have shown that it is possible to quantify it satisfactorily using multispectral data at both the aerial and ground levels (Barnes et al., 2000; Boegh et al., 2002). However, multi and hyper-spectral imagers are relatively expensive and complex from the operational point of view.

As a low-cost alternative, vegetation indices derived from Red-Green-Blue (RGB) cameras have been employed for remote sensing assessment in field conditions, providing a wide-range of phenomic data about genotypic performance under different stress conditions and species, including water stress and foliar diseases in bread wheat, durum wheat and tritordeum and triticale (Casadesus et al., 2007; Casadesús and Villegas, 2014; Vergara-Diaz et al., 2015; Zhou et al., 2015). Moreover, digital sensors have been successfully integrated on board unmanned aerial vehicles (UAV) to assess crop vigor, vegetation coverage, and greenness (White et al., 2012; Andrade-Sanchez et al., 2014;

Svensgaard et al., 2014). For example, digital indices derived from RGB images have been proposed for grain yield (GY) assessment in water limiting conditions (Casadesus et al., 2007) and for quantifying leaf N concentration (Rorie et al., 2011). However, the use of RGB images to assess genotypic performance in terms of yield and crop N accumulation in response to different levels of soil fertility has not yet been assessed. RGB images may represent a proper alternative to spectroradiometric approaches at different levels: at the whole trial level from aerial platforms, at the plot level from ground-based measurements or even at the single leaf level replacing leaf chlorophyll meters.

Information derived from plant samples may also be relevant for crop monitoring and phenotyping (Araus and Cairns, 2014). For example the stable isotope composition in plant matter constitutes an integrative selection criterion because it can describe the behavior of the crop under stress (Masuka et al., 2012). Nitrogen isotope composition ($\delta^{15}N$) can be employed to characterize the efficiency in using N fertilizers (Evans, 2001; Serret et al., 2008). For its part, implementing carbon isotope composition ($\delta^{13}C$) in maize is not clear for assessing genotypic differences due to the C_4-photosynthetic metabolism of this species, but still appears responsive to differences in growing conditions (Monneveux et al., 2007; Araus et al., 2010). Finally, some other morphological and compositional traits such as the specific leaf area (SLA), N concentration, N per unit leaf area (N/LA), and carbon to nitrogen ratio (C/N), which are in turn related to nitrogen use efficiency, leaf construction, and primary metabolism (Poorter and Evans, 1998; Feng et al., 2008), have the potential to be useful for breeding, but knowledge about their association with crop yield is scarce.

The main goal of this study is to develop affordable easy-to-use new phenotyping tools that increase selection efficiency for grain yield and leaf N concentration under different N fertilization conditions in maize. To accomplish this objective, we compared the accuracy of field-spectroradiometer data vs. RGB-derived vegetation indices assessing GY and leaf N concentration in a set of ten maize hybrids grown in the field under five N-fertilizer levels. Firstly, we assessed the performance of these parameters for all the N-treatments together, and subsequently we dissected the correlations within each N-level for further discussion of phenotyping. Additionally, simple regression models were made for GY prediction and these models were tested and validated against the experimental yield of another trial. The performance of the leaf parameters N/LA, C/N, SLA, and $\delta^{13}C$ and $\delta^{15}N$ were also studied with the aim of relating these structural and compositional leaf properties with crop performance and phenotyping data. All RGB and UAV imagery were obtained at flowering stage in order to integrate the differences in crop performance from plant emergence to flowering stage, when the number of kernels per ear is determined.

MATERIALS AND METHODS

Experimental Design and Growing Conditions

Field trials were carried out at the Southern Africa regional station of CIMMYT (International Maize and Wheat

Improvement Center) located in Harare (17°43′32″S, 31°00′59″E) where two field experiments were studied. Before sowing, soil pH, total soluble salts (TSS), nitrogen as nitrate (NO_3^-) and phosphorus (P_2O_5) were analyzed in three soil depth ranges (0–30, 30–60, and 60–90 cm) and six replicates for each depth range were produced. Mean values for the full soil profile were pH = 5.8, TSS = 240.9 ppm, NO_3^- = 4.12 ppm, and P_2O_5 = 18.93 ppm.

Ten maize hybrids were sown, three of them were commercial hybrids (PAN7M-81, SC635, SC537) and the other seven were maize hybrids developed at CIMMYT (TH11894, TH127591, TH127053, TH127618, TH13466, CZHH1155, TH127004). These maize hybrids cover a big range of agronomical sensitivity to low nitrogen conditions. A split-plot arrangement in a randomized block design was set up and five nitrogen fertilization levels (0, 10, 20, 80, and 160 kg·ha^{-1} NH_4NO_3) were applied in both trials. Two and three replicates were set for the first and second trials, with 100 and 150 being the respective number of plots in each trial (trials S and P, respectively). A two-row border was sown between fertilization treatments and on the edges of the trial to prevent spatial variability.

Seeds were sown during the wet season, on December 23th 2013, in two rows per plot; rows were 4 m long and 75 cm apart (6 m^2/plot), with 17 planting points per row and 25 cm between plants within a row. All trials were homogeneously fertilized with 400 kg·ha^{-1} of super-phosphate and potassium oxide fertilizer (P_2O_5 14% and K_2O 7%). Weather conditions throughout growing season were recorded with a weather station. The mean temperature was 18.9°C, mean humidity 81.2 and total rainfall during the crop period was 563.1 mm, therefore, preventing the water deficit in these rainfed conditions.

The trials were harvested on May 20th 2014. The central 3.5 m of each row was harvested discarding 2 plants at each end, thus the collected weight corresponded to 5.25 m^2 (0.75 m apart × 2 rows × 3.5 m long). The cobs were threshed and the grains dried until they reached around 12% moisture, and then the grain from each plot was weighed. Grain yield (GY, Mg·ha^{-1}) was calculated as follows: (X kg plot^{-1} × 10)/5.25 where X is the grain weight per plot.

NDVI Calculation

The normalized difference vegetation index (NDVI) was calculated using the equation:

$$NDVI = (NIR - R)/(NIR + R)$$

where R is the reflectance in the red band and NIR is the reflectance in the near-infrared band. NDVI was obtained around the flowering stage by using two different approaches: using ground measurements and from aerial multispectral images (**Figure 1**).

The NDVI of individual plots at ground level (NDVI$_{ground}$) was determined with a ground-based portable spectroradiometer with an active sensor (GreenSeeker handheld crop sensor, Trimble, USA). This equipment uses the spectral wavelengths 650–670 nm as the red band and 765–795 nm as the near infrared. The distance between the sensor and the plots was kept constant
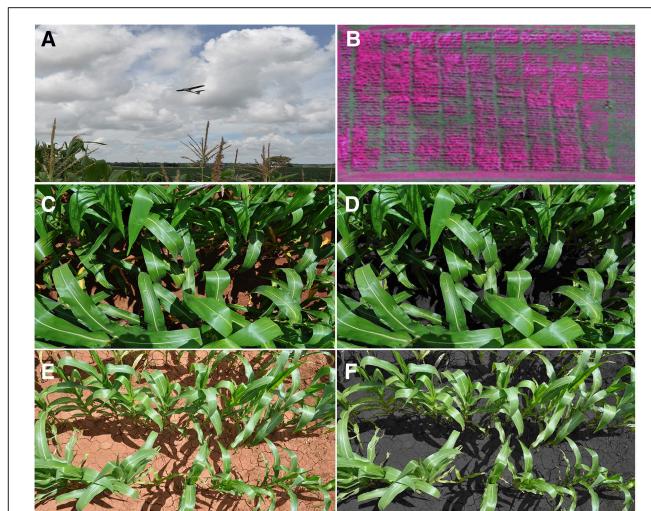

FIGURE 1 | **(A)** The unmanned aerial vehicle flying over the maize crops; **(B)** Multispectral false-color image at the aerial level showing near infrared (800 nm) as red, green (550 nm) as blue and red (670 nm) as green, spatial resolution of 10 cm/px; **(C)** RGB digital image from the high-nitrogen fertilization treatment at the canopy level; **(D)** and its resulting processed image with BreedPix; **(E)** RGB image from the low-nitrogen treatment and **(F)** its respective processed image.

using a ladder, around 0.5–0.6 m above and perpendicular to the canopy. The whole areas of the two trials were measured from 12 to 14 h on March 3rd and 4th, 2014.

The aerial NDVI index (NDVI$_{aerial}$) was obtained using a UAV-based remote sensing platform developed by Airelectronics (Montegancedo campus, Spain) in collaboration with the Crop Breeding Institute-Zimbabwe, CIMMYT, QuantaLab at the Institute for Sustainable Agriculture (IAS-CSIC, Spain) and the University of Barcelona. This aerial platform was equipped with a multispectral camera (ADC-Lite, Tetracam, Inc., Chatsworth, CA, US), which provides spectral images on the green, the red and the near-infrared bands, with a final ground resolution of 10 cm per pixel when flying at an object distance of 150 m. These bands are approximately equal to the Landsat Thematic Mapper (TM) bands TM2, TM3, and TM4, respectively, so that the spectral wavelengths from 630 to 690 nm represent the red band and 760 to 920 nm the near infrared band. The flight was conducted at an altitude of 150 m at midday on a sunny day when crops were around the flowering stage. The collected images covered 220 out of the total 250 plots, completely covering the block S trial (100 plots) and partially covering block P (120 of the total of 150 plots). Aerial images were subsequently corrected and calibrated with ImapQ (QuantaLab-IAS-CSIC, Cordoba, Spain) which converts images to radiance. Mosaicking and rectifying processes were applied with Autopano (Kolor SARL, Francin, France) by applying the image stitching technique (SIFT algorithm) in addition to a manual orthorectification from several checkpoints selected. NDVI values were finally extracted from the images using ENVI software (Exelis Visual Information Solutions, Boulder, Colorado, USA).

RGB Indices

Vegetation indices derived from red-green-blue (RGB) images were evaluated at the plot and the single leaf level (RGB$_{canopy}$ and RGB$_{leaf}$ indices, respectively; **Figure 1**). In the case of RGB$_{canopy}$, one digital RGB picture was taken per plot by holding the camera about 0.8–1.0 m above the canopy, in a zenithal plane and focusing near the center of each plot. Plot images were taken on the same days as the measurements with the ground spectroradiometer using a Nikon COOLPIX S8000 digital compact camera without flash and with a focal length of 54 mm and were saved in a 4288 × 2848 pixel JPEG format. Later, six leaves per plot were taken from the S trial (100 plots) and were subsequently scanned with a Dell 2155 cdn multifunction color printer (Round Rock, TX, USA). Finally, scanned images were saved in the same format with a resolution of 2338 × 1653 pixels and RGB$_{leaf}$ indices calculated as below.

Subsequently, images were analyzed with the open source Breedpix 0.2 software (Casadesus et al., 2007) designed to process digital images. This software enables calculation of several RGB vegetation indices based on the different properties of color inherent in RGB images. RGB VIs were obtained either from the average color of the whole image or from the hue histogram in each image. BreedPix produces several automatic conversions of the original RGB image to other color spaces (i.e., each model that numerically represents the color in terms of different coordinates). Four VIs (a*, b*, u*, and v*) belonging to CIE (from the French abbreviation of International Commission on Illumination) color spaces were calculated and used in this study. The software require the use of Java Advanced Imaging (JAI) for the conversion of RGB color space to CIE-XYZ color space and the resulting coordinates are subsequently converted to other color spaces. First, the VIs a* and b* belong to CIE-Lab color space, being L* the lightness dimension and a* and b* the color-opponent coordinates. Red/green opponent-colors are represented along a* axis, whereas b* axis represent the yellow/blue opponent colors. Similarly, u* and v* indices represent the axis in the chromaticity diagram of CIE-Luv color space. Thereby the software obtains the average values of these components of color for each one of the processed images. Hue component is calculated using the JAI functions which employ the formulae described in Seul et al. (2000) whereas the components of CIE-Lab and CIE-Luv color spaces are calculated as described in Trussell et al. (2005). The relative green area (GA) and the relative "greener area" (GGA) are based on the sum of frequencies of the histogram classes included in a certain range of hue in the image. GA is the percentage of pixels in the image in the hue range from 60 to 180°, that is, from yellow to bluish green. On the other hand GGA is somewhat more restrictive since the range of hue considered by this index is from 80 to 180°, excluding yellowish-green tones and therefore, it more accurately describes the amount of photosynthetically active biomass and leaf senescence.

Analysis of Leaf Parameters

The leaf portions in the RGB$_{leaf}$ indices were also used the subsequent measures. Firstly, immediately before being scanned, a handheld spectroradiometer developed for leaf chlorophyll measurements (Minolta SPAD-502, Spectrum Technologies Inc, Plainfield, IL, USA) was used to measure the index related to leaf chlorophyll content (LCC). Four measurements were made for each leaf segment. Secondly, the leaves were oven dried at 70°C for 24 h and the dry weight was measured. Then the specific leaf area (SLA) was calculated using the equation

$$SLA = LA / DW$$

where LA is the total leaf area (m^2) measured previously from the scanned images using the open-source Java-based software ImageJ (http://rsb.info.nih.gov.sire.ub.edu/ij/) and DW is the corresponding dry weight (kg).

Finally, dry leaves were ground to a fine powder and 0.7–0.9 mg of leaf dry matter from each plot was weighed and sealed into tin capsules. Stable carbon (^{13}C/^{12}C) and nitrogen (^{15}N/^{14}N) isotope ratios as well as the leaf N and C concentrations (%) were measured using an elemental analyser (Flash 1112 EA; Thermo Finnigan, Bremen, Germany) coupled with an isotope ratio mass spectrometer (Delta C IRMS, Thermo Finnigan) operating in a continuous flow mode. Samples were loaded into a sampler and analyzed. Measurements were conducted at the Scientific Facilities of the University of Barcelona. Isotopic values were expressed as a composition notation (δ) as follow:

$$\delta\ (‰) = (^{13}C/^{12}C)_{sample}/ (^{13}C/^{12}C)\ _{standard} - 1$$

where "sample" refers to plant material and "standard" to international secondary standards of known ^{13}C/^{12}C ratios (IAEA CH7 polyethylene foil, IAEA CH6 sucrose and USGS 40 L-glutamic acid) calibrated against Vienna Pee Dee Belemnite calcium carbonate with an analytical precision (standard deviation) of 0.15‰. The same δ notation was used for the ^{15}N/^{14}N ratio expression but with the standard referring to air. For nitrogen, international isotope secondary standards IAEA N1, IAEA N2, IAEA NO3, and USGS 40 were used with a precision of 0.3‰. Further, the C/N ratio was obtained from these analyses and total nitrogen concentration per unit leaf area (N/LA) was calculated with the formula:

$$N\ /LA = \left(\frac{DW}{LA}\right) x\ N$$

where LA is the total leaf area (m^2), DW is the corresponding dry weight (g) and N is its nitrogen concentration (in % dry matter).

Statistical Analysis

Statistical analyses were conducted using SPSS 21 (IBM SPSS Statistics 21, Inc., Chicago, IL, USA). Multiple variance analyses, the multiple comparison Duncan *post-hoc* test and bivariate correlations were performed. The presented leaf parameters (LCC, N, δ^{15}C, N/LA, SLA, δ^{13}C, C/N) and RGB$_{leaf}$ indices from scanned leaves were only analyzed for the S trial, whereas the NDVI$_{ground}$, NDVI$_{aerial}$, and RGB$_{canopy}$ indices were studied for both trials. The determination coefficients of the linear relationships of GY and leaf N concentration with the vegetation indices NDVI and RGB were calculated for the entire trials and within each N fertilization treatment. All graphs were

performed with SigmaPlot 10.0 (Systat Software Inc., San Jose, California, US).

RESULTS

Significant differences in GY between genotypes and nitrogen-fertilization levels were observed in this study (**Table 1**) with GY increasing in response to N fertilization (**Table 2**). Differences within nitrogen-input levels were also detected with both (ground and aerial) NDVI approaches and with all RGB$_{canopy}$ indices except by a*. Genotypic differences were detected by the RGB$_{canopy}$ indices GGA, GA, a*, u*, and hue, whereas among the spectroradiometric indices only the NDVI at the ground level detected genotypic differences.

Leaf N concentration varied significantly between genotypes and the effect of N-fertilization levels was highly significant

TABLE 1 | P-values from multivariate analysis of variance with two fixed factors: genotype and nitrogen level and its interaction (GxN).

	Genotype	Nitrogen level	G × N
GY	< 0.001	< 0.001	0.884
SPECTRAL INDICES			
NDVI$_{aerial}$	0.509	< 0.001	0.745
NDVI$_{ground}$	< 0.001	< 0.001	0.500
RGB$_{canopy}$ INDICES			
hue	0.032	< 0.001	0.997
a*	0.006	0.125	0.727
b*	0.590	< 0.001	0.997
u*	0.012	< 0.001	0.801
v*	0.567	< 0.001	0.993
GA	0.002	< 0.001	0.708
GGA	< 0.001	< 0.001	0.564
RGB$_{leaf}$ INDICES			
hue	0.412	< 0.001	0.817
a*	0.364	< 0.001	0.882
b*	0.580	< 0.001	0.999
u*	0.310	< 0.001	0.687
v*	0.787	< 0.001	0.999
GA	0.042	< 0.001	0.604
GGA	0.289	< 0.001	0.225
LCC	0.939	< 0.001	0.973
ANALYZED PARAMETERS			
Leaf %N	0.026	< 0.001	0.999
N/LA	0.347	< 0.001	0.972
C/N	0.120	< 0.001	1.000
δ^{15}N	0.375	< 0.001	1.000
δ^{13}C	< 0.001	< 0.001	0.937
SLA	0.822	0.004	0.89

As dependent variables, grain yield (GY), leaf nitrogen concentration (%N), NDVI at aerial and ground levels, RGB indices from canopy images and scanned leaves, leaf chlorophyll content (LCC), nitrogen per unit area (N/area), the stable carbon (δ^{13}C), and nitrogen (δ^{15}N) isotope composition, carbon nitrogen ratio (C/N) and specific leaf area (SLA).

(**Table 1**) with values increasing as N fertilization increased (**Table 2**). All the RGB$_{leaf}$ indices detected very significant differences between nitrogen treatments and genotypic differences were also found with GA (**Table 1**). At the same time, LCC also indicated highly significant differences between N fertilization levels but not genotypic differences.

All the analyzed leaf parameters (N, N/LA, SLA, δ^{15}N, δ^{13}C, C/N) were highly sensitive to variations in N fertilizer levels (**Table 1**). In contrast, apart from leaf N concentration, genotypic differences were only detected for δ^{13}C. Increasing N fertilization caused significant increases in leaf N, N/LA and SLA while δ^{15}N and the C/N ratio decreased (**Table 2**). Additionally, differences among N fertilization levels were also found in δ^{13}C but its trend was somewhat different: in the low-N levels δ^{13}C was quite steady and then it decreased at 80–160N, whereas leaf-N concentration increased.

Additionally, the effect of changing light in outdoor conditions was evaluated in RGB indices obtained from canopy images (Table S1). For this purpose, 57 plots were photographed twice in nearly consecutive days, firstly in a sunny day and secondly in a partly cloudy day. All indices were strongly correlated between replicates ($p < 0.001$), particularly the indices GA, GGA, u*, a* ($R^2 > 0.72$).

Grain Yield Assessment across Nitrogen Regimes and Genotypes

All vegetation indices (either ground and aerial NDVI, RGB$_{canopy}$, or LCC) were strongly correlated with GY variation across the whole set of plots of the two trials. The best results were obtained by using the RGB-indices GA and GGA at the canopy level, which showed an exponential regression model and explained 70–72% of GY variability (**Figure 2**). Meanwhile, the RGB$_{canopy}$ indices u* and a* evolved inversely with increasing GY and demonstrated lower accuracy ($R^2 = 0.326$ and $R^2 = 0.302$, respectively, data not shown). In contrast, LCC evolved linearly with increases in GY and explained 69% of GY variation (**Figure 2**). Finally, both NDVI approaches followed a power regression model and their determination coefficients were moderate and similar (NDVI$_{ground}$ at **Figure 2**; NDVI$_{aerial}$ $R^2 = 0.293$, data not shown).

Additionally, simple regression models from the P trial that explained GY across the different N fertilization levels were obtained by using the different VIs and validated for their accuracy in estimating the GY of the S trial (**Table 3**). The estimated GY from all VIs always fitted satisfactorily with the experimental GY for the entire trial. The determination coefficients increased further when six hybrids contrasting in their grain yield were selected, with three of them being high-yielding and the remaining ones low-yielding. Genotypic differences were found in the estimated GY from GGA, NDVI$_{aerial}$ and NDVI$_{ground}$ between the six selected hybrids and the experimental GY was also significantly different. Moreover, in all models, differences between N fertilization levels were always detected by the estimated GY.

TABLE 2 | Means of grain yield (GY) (Mg·ha^{-1}) from the two trials, leaf nitrogen concentration (%N), nitrogen per unit leaf area (N/LA), specific leaf area (SLA), the stable carbon (δ^{13}C), and nitrogen (δ^{15}N) isotope composition and the leaf C/N ratio according to the ten hybrids and the five nitrogen levels.

	GY (Mg·ha^{-1})	%N	N/LA (g/m^2)	SLA (m^2·kg)	C/N ratio	δ^{15}N (‰)	δ^{13}C (‰)	LCC (u)
GENOTYPE								
PAN 7M-81	5.03b	2.203ab	111.24ab	20.23a	21.07b	4.53b	−11.7ab	41.53a
TH11894	5.12b	2.196ab	101.75a	21.48a	20.77b	3.78ab	−11.49cd	42.02a
TH127591	5.14b	2.312abc	117.87ab	20.41a	19.75ab	4.04ab	−11.79a	43.06a
TH127053	5.7bc	2.32abc	118.29ab	19.56a	19.75ab	4.18ab	−11.55bcd	41.13a
SC635	3.52a	2.506bc	125.31ab	19.98a	18.38ab	3.66ab	−11.77a	41.84a
TH127618	5.38b	2.621c	130.12ab	20.54a	17.27a	3.64ab	−11.63abc	43.56a
TH13466	5.75bc	2.4abc	117.62ab	20.91a	18.96ab	3.76ab	−11.45*d*	42.57a
CZH1155	4.63b	2.526bc	133.84b	18.9a	18.48ab	3.19a	−11.64abc	42.28a
TH127004	6.49c	2.138a	113.71ab	19.05a	21.35b	3.87ab	−11.71ab	41.76a
SC537	4.67b	2.38abc	121.71ab	20.32a	19.65ab	3.79ab	−11.7ab	42.39a
NITROGEN LEVEL								
0	3.13a	1.887a	107.62a	17.82a	23.75d	5.612*e*	−11.584bc	32.5a
10	3.9b	1.881a	92.78a	20.6b	23.43d	4.613d	−11.59bc	33.4a
20	4.61c	2.089a	107.09a	19.91ab	21.01*c*	3.877c	−11.512cd	40.4b
80	6.48d	2.739b	138.68b	20.14b	16 b	2.898b	−11.652b	50.9c
160	7.59e	3.206c	149.56b	22.22b	13.54a	2.218a	−11.869a	53.8d

Letters are significantly different according to Duncan's multiple range test (P < 0.05).

Grain Yield Assessment across Genotypes within Each N Regime

To further assess the accuracy of these indices, the determination coefficients for GY prediction within each N-input level across genotypic means were performed (**Table 4**). GGA, GA, u*, and a* indices were correlated significantly to GY variation within all N levels, whereas both NDVI approaches were correlated significantly to GY only for some of the studied N levels. By contrast, LCC did not correlate with GY across plots within any of the N levels.

Leaf Nitrogen Assessment across N Regimes and Genotypes

LCC was the best predictor of leaf N concentration across the entire trial, explaining more than 80% of N variability, moderately surpassing the fitting accuracy of the RGB$_{leaf}$ indices (**Figure 3**). Thus, the RGB$_{leaf}$ index a* explained about 69% of leaf N variation across N fertilization treatments (**Figure 3**) and u*, b*, and v* were quite similar ($R^2 = 0.682$, $R^2 = 0.643$, and $R^2 = 0.621$, respectively, data not shown). For its part, NDVI$_{aerial}$ was also a good predictor of leaf N (**Figure 3**), whereas NDVI$_{ground}$was less accurate in its prediction ($R^2 = 0.116$, data not shown). Finally, the RGB index v* at the canopy level was more related to leaf N than it was to GY, and it was shown to be a reasonably good predictor of leaf N across the whole trial (**Figure 3**).

Leaf Nitrogen Assessment across Genotypes within Each N Regime

A table depicting the determination coefficient between the RGB$_{leaf}$ indices, NDVI$_{aerial}$, NDVI$_{ground}$, and LCC against leaf N across genotypic means within each of the N fertilization levels is presented (**Table 5**). In the low-N treatments (0N to 20N) the best determination coefficients were provided by the RGB$_{leaf}$ indices b*, v*, u*, and a*. In addition, most of the RGB$_{leaf}$ indices were also sensitive to leaf N variation at the 80N level but none of them related significantly at the 160N level. For its part, LCC showed a quite similar accuracy compared with the RGB$_{leaf}$ indices in their predictions of leaf N within the low-N levels, but it was not significantly correlated in the high-N fertilization levels (**Table 5**). Finally, NDVI$_{aerial}$ was especially sensitive to leaf N variations in the high-N and 0N treatments, whereas NDVI$_{ground}$ was generally unrelated to leaf N within each N treatment.

Leaf Parameters Performance and Relationships with VIs and Yield

Leaf N was strongly negatively correlated across N levels with δ^{15}N and the C/N ratio and to a lesser extent with δ^{13}C and SLA (Table S2). Correlations of these traits with GY were also negative but weaker, except for SLA which did not correlate.

Most of the RGB indices (both at the leaf and canopy scales), the LCC and the NDVI correlated with N/LA across N regimes, but always more moderately than they correlated with leaf N concentration (Table S2). The association of δ^{15}N with NDVI, LCC, and RGB indices (at the both scales) was highly significant and in some cases their correlation coefficients were higher than the respective coefficients between δ^{15}N and leaf N. Similarly, δ^{13}C was fairly well correlated with most of the RGB indices (especially at the leaf scale) and LCC. Regarding the C/N ratio, LCC was the best predictor but this correlation was smaller than with leaf N concentration. However, most of the RGB$_{leaf}$ indices (a*, b* u*, v*, GA), the RGB$_{canopy}$ indices (hue, u*, GA, GGA) as well as NDVI$_{ground}$ and NDVI$_{aerial}$ correlated more strongly with
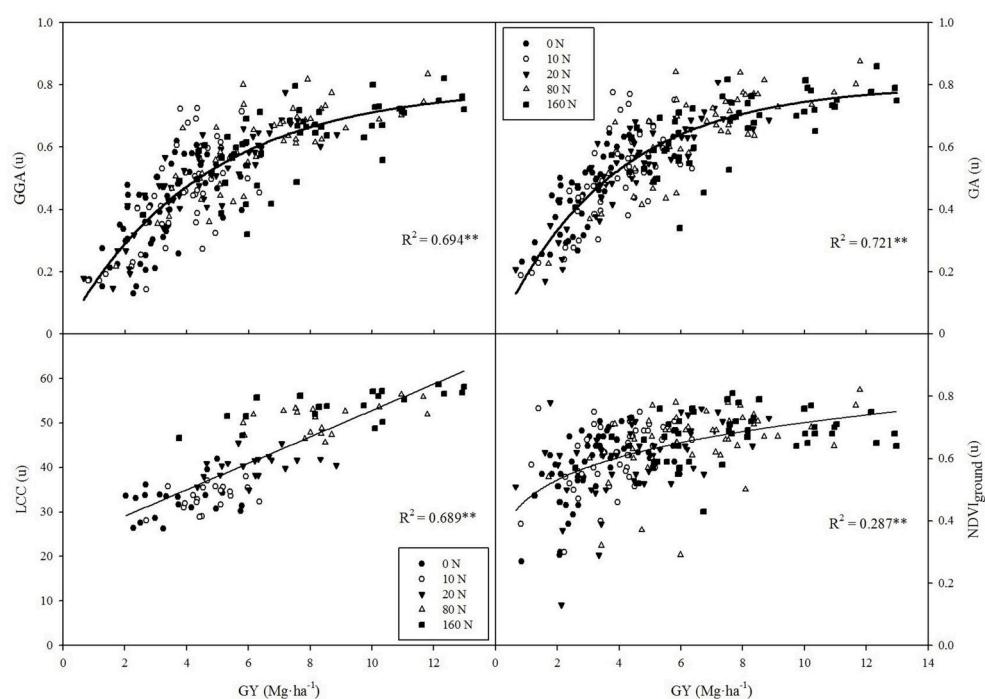
FIGURE 2 | Correlations between grain yield (GY) and leaf chlorophyll content (LCC), the RGB indices GA and GGA at the canopy level and NDVI at the ground level. R^2, determination coefficient; **$P < 0.001$.

the leaf C/N ratio than they did with leaf N (Table S1). Finally, SLA correlated strongly with the RGB$_{leaf}$ indices GA and GGA, and slightly with both NDVIs.

The relationships between leaf N, N/LA, C/N, δ^{13}C, δ^{15}N, and SLA with GY across genotypes within N fertilization treatments were almost all non-significant except for leaf N in the 160N treatment (**Table 4**). Regarding the genotypic correlations within each N fertilization level of these leaf traits with leaf N, only the leaf N derived parameters (C/N and N/LA) were significantly correlated (**Table 5**).

DISCUSSION

Crop Monitoring and Phenotyping Parameters for GY Estimation

As previously found in other studies in wheat grown under different stress conditions (Casadesus et al., 2007; Morgounov et al., 2014; Vergara-Diaz et al., 2015), the RGB$_{canopy}$ indices (from BreedPix software) measured at flowering were strongly correlated with GY. RGB-based indices may perform far better than NDVI for GY prediction, which has been recently described under water and biotic stresses in wheat (Elazab et al., 2015; Vergara-Diaz et al., 2015; Zhou et al., 2015). The lower accuracy of NDVI in comparison to digital-based RGB indices can be explained in several ways. On the one hand, graphs clearly highlight (**Figure 2**) that the variability in the canopy NDVI values at ground level is small, with more than 90% of values being in the range 0.5–0.8 and with the NDVI

values in the low N treatments being already relatively high (e.g., average of NDVI$_{ground}$ = 0.57 in the 0N treatment). Therefore, the NDVI values remained almost unchanged as GY increased from 4 to 13 Mg ha^{-1}. These results support the previously reported saturation of reflectance spectra in the red and near-infrared regions, such that increasing leaf area does not involve a parallel increase in NDVI values (Hobbs, 1995; Elazab et al., 2015). Thus, the relationship between NDVI and aerial biomass saturates as canopies become denser (i.e., LAI > 4) and as a consequence the relationship between the NDVI and GY also worsened as GY increased. Moreover near-infrared reflectance is sensitive to canopy architecture variations (Gitelson et al., 2002) which surely affected NDVI measurements in maize canopies. The use of multi-angular spectral data may solve these problems by capturing the scattering of sunlight by vegetation, which enables to assess three-dimensional vegetation structures (Hasegawa et al., 2010). Whereas this approach may improve the estimation of NDVI (and other spectral indices) for phenotyping, the increasing complexity (i.e., more time and resources needed) of the method makes it less feasible as low-cost alternative.

For its part, the range of variability in the RGB$_{canopy}$ index, GA, was much wider (only 63% of values were in the range of 0.5–0.8) and GA values in the low N treatments were somewhat smaller (average GA = 0.46 in the 0N treatment) than those of the NDVI, and in fact GY correlated much better with GA than with the NDVI. Even so, the RGB$_{canopy}$ indices also seemed to saturate for high GY but to a lesser extent than the NDVI because they mainly depend on changes in pigment concentration and the

TABLE 3 | Simple regression models obtained with different Vegetation Indices (the spectroradiometric indices NDVI$_{aerial}$, NDVI$_{ground}$, and the RGB$_{canopy}$ indices GA and GGA) in the P trial, explaining Grain Yield (GY) variation across nitrogen fertilization levels, were used for GY estimation in the S trial.

			GY exp. vs. GY est.			
	Predictors	Simple Regression models	R^2 for all trial	R^2 for 6 hybrids	Genotype	N-level
GY est.	GA	GY = e $^{((GA - 0.195)/0.24)}$	0.674**	0.711**	0.054	< 0.001
	GGA	GY = e $^{((GGA - 0.151)/0.248)}$	0.684**	0.719**	0.040	< 0.001
	NDVI$_{aerial}$	GY = e $^{((NDVIaerial - 0.192)/0.094)}$	0.452**	0.543**	0.002	< 0.001
	NDVI$_{ground}$	GY = e $^{((NDVIground - 0.461)/0.109)}$	0.231**	0.324**	0.001	< 0.001
GY exp.		GY 3-3 hybrids	–	–	0.044	< 0.001

The fit of the estimated Grain Yield (GY est.) to the experimental Grain Yield (GY exp.) was tested with the determination coefficient (R^2) for the entire trial and for six yield-contrasting hybrids. P-values were analyzed for all estimated GYs and for the experimental GY using the six selected hybrids. **P < 0.001.

TABLE 4 | Determination coefficients (R^2) of RGB-indices from canopy images (RGB$_{canopy}$), aerial NDVI, ground NDVI, leaf chlorophyll content (LCC), the leaf nitrogen concentration on a dry matter basis (Leaf %N), the nitrogen concentration on a leaf area basis (N/LA), the ratio of carbon to nitrogen concentration (C/N), the stable carbon (δ^{13}C) and nitrogen (δ^{15}N) isotope composition and the specific leaf area (SLA) predicting grain yield in the five N levels separately (0, 10, 20, 80, and 160 kg·ha^{-1} NH$_4$NO$_3$) following linear regression models.

	0N	10N	20N	80N	160N
SPECTRAL INDICES					
NDVI$_{aerial}$	0.019 ns	0.415*	0.092 ns	0.074 ns	0.621**
NDVI$_{ground}$	0.741**	0.381 ns	0.421*	0.212 ns	0.289 ns
RGB$_{canopy}$ INDICES					
hue	0.632**	0.653**	0.717**	0.364 ns	0.729**
a*	0.706**	0.634**	0.627**	0.524*	0.464*
b*	0.040 ns	0.160 ns	0.046 ns	0.056 ns	0.889**
u*	0.709**	0.608**	0.666**	0.491*	0.569*
v*	0.079 ns	0.239 ns	0.001 ns	0.120 ns	0.798 **
GA	0.771**	0.659**	0.704**	0.501**	0.764**
GGA	0.872**	0.664**	0.774**	0.555*	0.748**
LCC	0.148 ns	0.163 ns	0.076 ns	0.021 ns	0.004 ns
Leaf %N	0.059 ns	0.100 ns	0.014 ns	0.001 ns	0.504*
N/LA	0.037 ns	0.046 ns	0.006 ns	0.014 ns	0.189 ns
C/N	0.085 ns	0.069 ns	0.006 ns	0.009 ns	0.365 ns
δ^{13}C	0.002 ns	0.353 ns	0.019 ns	0.073 ns	0.001 ns
δ^{15}N	<0.001 ns	0.007 ns	0.005 ns	0.025 ns	0.065 ns
SLA	0.004 ns	0.002 ns	0.026 ns	0.008 ns	0.030 ns

*P < 0.05; **P < 0.001; ns, non-significant.

canopy LAI is less affected in the visible region than in the NIR region (Casadesus et al., 2007; Elazab et al., 2015).

In the case of the airborne NDVI data, the correlation with GY was also much lower than with GA taken on individual plots with GY. In fact, the images from the ADC multispectral camera have around four-fold less resolution than current digital camera technology (3.2 vs. 12 MP in our study, respectively). Although many ADC images were employed to obtain mosaics of the entire field trials, the resolution obtained at the flight altitude generated pixels which were mixed between pure vegetation, shadows and soil components. Such effects were successfully separated in the imagery collected at the near-canopy level with the RGB camera due to the higher resolution obtained. Altogether, the NDVIaerial provides a much lower amount of information than

the GA and other VIs derived from RGB images taken at the plot level.

In the case of the LCC, it correlated strongly and linearly with grain yield across fertilization levels (**Figure 2**). In fact the leaf chlorophyll meters calculate a spectral ratio of the leaf transmittance to the near-infrared and red bands and they were primarily developed to assess N fertilization levels (Fox et al., 1994; Markwell et al., 1995). LCC indirectly predicts GY when a wide range of N conditions are considered and this is probably due to the relationship between chlorophyll content, leaf N and yield (Argenta et al., 2004).

Concerning the applications in breeding, the determination coefficients within N levels across genotypic means (**Table 4**) support the strength of RGB$_{canopy}$ indices as phenotyping
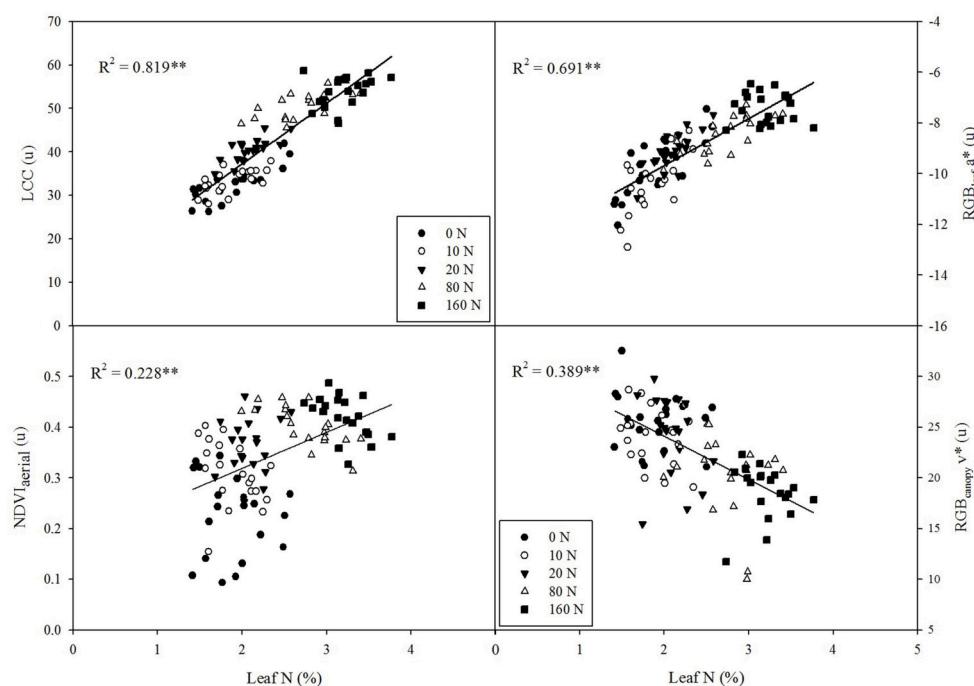
FIGURE 3 | Correlations between leaf nitrogen concentration (%N) and leaf chlorophyll content (LCC), the RGB index a* at the leaf level, the NDVI at the ground level, and the RGB index v* at the canopy level. R^2, determination coefficient; **P < 0.001.

parameters. Thus, these indices were able to indicate the most efficient genotypes in terms of grain yield within each N fertilization level, whereas the NDVI performed much worse as a phenotyping parameter. Although genetic variability in maize hybrids in response to low N doses is high (Wang et al, 1999; Zaman-Allah et al., 2015) it has been scarcely exploited by breeding programs since they mainly focus on breeding for maize performance under favorable conditions (Machado and Fernandes, 2001). In this sense, the proposed phenotyping parameters herein, based on the use of RGB images, can significantly contribute to selection of maize hybrids resilient to low N as well as being more responsive to increases in N fertilization. For its part, LCC was unrelated to genotypic GY variation at any of the N-levels tested (**Table 4**), and this is in agreement with previous reports in maize that have noted LCC as not always being significantly correlated with genotypic differences in GY (Gallais and Coque, 2005).

Crop Monitoring and Phenotyping Parameters for Leaf N Assessment

The importance of leaf N concentration for N management and breeding lies not only in its potential contribution to grain N (Gallais and Coque, 2005) but is also due to it being a component of the nitrogen uptake efficiency (Serret et al., 2008). Moreover, leaf N is an indicator of leaf photosynthetic capacity contributing to grain yield (Richards, 2000) as well as a key fodder trait (Van der Wal et al., 2000). Therefore, the estimation of leaf N concentration within a given N fertilization treatment may

provide valuable information about the genotypic efficiency for the uptake of N.

Our study highlights the potential of RGB indices for precise crop N management and for phenotyping genotypic performance under a wide-range of N conditions. As widely reported, LCC proved to be a very good indicator of leaf N concentration across nitrogen fertilization levels, therefore enabling monitoring of N application (Hirel et al., 2007). However, LCC failed to be effective as a phenotyping parameter, especially at high N-fertilization levels (**Table 5**). In contrast, the RGB_{leaf} indices demonstrated that they were the best genotypic predictors for leaf N concentration in the 0 to 80 kg·ha^{-1} N range. Thus, RGB indices at the leaf level have the potential to inform breeding programs about tolerance to N-deficiency stress in maize. This is a helpful insight because selection experiments have shown that the maximum genetic advance for low N is achieved when selecting in such N conditions (Gallais and Coque, 2005).

By contrast, in the highest N-fertilization level (160 kg ha^{-1}) the RGB_{leaf} indices and LCC were probably saturated because they did not correlate with variations in leaf N concentration. For its part, the $NDVI_{aerial}$ had an irregular trend as it was significantly correlated to changes in leaf N concentration at three of the five N fertilization levels (0, 80, and 160 kg ha^{-1}) and these correlations were especially strong in the high N levels. As discussed above, besides of some plot variability and soil exposure, the poorer performance of the $NDVI_{aerial}$ may be mainly explained by the relatively poor spectral resolution at the single plot level of the multispectral aerial images. Even

TABLE 5 | Determination coefficients (R^2) of RGB-indices from scanned leaves (RGB$_{leaf}$), leaf chlorophyll content (LCC), ground NDVI, aerial NDVI, the nitrogen concentration on a leaf area basis (N/LA), the ratio of carbon to nitrogen (C/N), leaf stable carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) isotopic composition and specific leaf area (SLA) predicting leaf nitrogen concentration on a dry matter basis separately in the five N fertilization levels (0, 10, 20, 80, and 160 kg·ha^{-1} NH$_4$NO$_3$).

	0N	10N	20N	80N	160N
SPECTRAL INDICES					
NDVI$_{aerial}$	0.545*	0.242 ns	0.352 ns	0.755**	0.650**
NDVI$_{ground}$	0.113 ns	0.223 ns	0.285 ns	0.403*	0.057 ns
RGB$_{leaf}$ INDICES					
hue	0.081 ns	0.155 ns	0.188 ns	0.094 ns	0.012 ns
a*	0.588*	0.261 ns	0.567*	0.500*	0.149 ns
b*	0.526*	0.489*	0.566*	0.437*	0.265 ns
u*	0.607**	0.171 ns	0.534*	0.507*	0.073 ns
v*	0.581*	0.546*	0.429*	0.316 ns	0.192 ns
GA	0.097 ns	0.013 ns	0.185 ns	0.517*	0.222 ns
GGA	0.261 ns	0.026 ns	0.229 ns	0.359 ns	0.026 ns
LEAF PARAMETERS					
LCC	0.587*	0.426*	0.401*	0.168 ns	0.005 ns
N/LA	0.810**	0.127 ns	0.643**	0.496**	0.552**
C/N	0.973**	0.947**	0.918**	0.980**	0.939**
$\delta^{13}C$	0.318 ns	0.083 ns	0.069 ns	0.158 ns	0.073 ns
$\delta^{15}N$	0.006 ns	0.004 ns	0.167 ns	0.183 ns	0.295 ns
SLA	0.167 ns	0.069 ns	0.219 ns	0.022 ns	0.121 ns

*$P < 0.05$; **$P < 0.001$; ns, non-significant.

so, according to our results this approach seems efficient for its implementation in aerial platforms.

Use of Leaf Analytical Parameters for Crop Management and Phenotyping

Besides the leaf N concentration discussed above, other leaf N parameters like the N concentration on an area basis (N/LA) and the C/N ratio were strongly associated with GY across N fertilization levels. In the case of the leaf $\delta^{15}N$, its value gradually decreased as the N application rate increased. This trend is due to the absorption of N from chemical fertilizers that are highly depleted in ^{15}N, whereas in the low N treatments plants absorb the N available in the soil, which is usually ^{15}N-enriched (Bateman et al., 2005; Masuka et al., 2012). However, the genotypic effect was not significant for $\delta^{15}N$, which does not support the use of this isotopic signature for maize phenotyping under low N stress. These results disagree with previous studies in wheat where genotypic differences were found under N stress conditions (Araus et al., 2013).

In agreement with previous studies (Dercon et al., 2006), low N induced higher $\delta^{13}C$ in maize, whereas it decreased in the high N fertilization treatments. This pattern of response appears related to the occurrence of some degree of water stress associated with a larger transpiring area due to nitrogen fertilization. In agreement with previous studies in maize, genotypic differences in leaf $\delta^{13}C$ may be attributed to differences in transpiration efficiency, but the variation in $\delta^{13}C$ was unrelated to GY within treatments (Cabrera-Bosquet et al., 2009).

Previous studies noted the relevance of SLA for the compositional and ecophysiological characterization of plants (Reich et al., 1998; Nautiyal et al., 2002). Several authors (Poorter and Evans, 1998; Meziane and Shipley, 2001) have reported a positive relationship between leaf N and SLA (**Table 2**). In turn, changes in SLA may be due to variations in leaf thickness and/or leaf density (Witkowski and Lamont, 1991). Increasing leaf density in low N conditions may be attributed to the increased synthesis of dense tissues such as sclerenchyma and vascular tissues that are rich in nitrogen-free substances (Garnier et al., 1997), whereas leaf thickness seems to have a minor role (Arendonk and Poorter, 1994). However, concerning its phenotyping use, SLA was shown to be homogeneous among the studied maize hybrids and unrelated to GY, as well as within a given N fertilization level, which excludes SLA as a phenotyping trait.

Regarding the relationship between VIs and the C/N ratio, most of the RGB indices (at the canopy and leaf levels) and both NDVI approaches were demonstrated as being even better correlated to the leaf C/N ratio than to leaf N concentration (Table S2). This finding may have considerable economic implications as the C/N ratio informs not only about the crop N status but also about the aerial biomass quality, including digestibility and nutritional quality (Van der Wal et al., 2000). Finally, all VIs and the LCC were better at capturing the differences in leaf N concentration than the amount of N concentration per unit leaf area (Table S2), thus avoiding the effect of leaf thickness or density. This evidence is enhanced by the weak relationship of the digital and spectral indices to SLA.

This finding is particularly interesting in the case of LCC (SPAD readings), which has been previously positively correlated with leaf thickness and negatively correlated with SLA in other species (Marenco et al., 2009).

Implications for Breeding and Crop Management

The tested vegetation indices based on RGB images and to a lesser extent the NDVI demonstrated a high-throughput for the accurate prediction of several traits that are highly valuable for maize breeders and agronomists such as grain yield, leaf N concentration and the ratio of carbon to nitrogen under a wide range of N fertilization levels. Proper N fertilization management may be assisted considerably by using these parameters as decision criteria controlling the expected production and the uptake of N by the above-ground biomass. Beyond this, maize breeding programs may benefit from these findings through their application during the characterization of genotypic performance within N fertilization levels. In this way the selection of the most efficient genotypes in terms of grain production and/or N uptake may respond to the needs of low N stress tolerant maize varieties.

Vegetation indices derived from RGB images proved to be broad-use because they were previously employed satisfactorily in other crops under biotic and water stress conditions (Casadesus et al., 2007; Vergara-Diaz et al., 2015). Therefore, since this technique has proven its efficiency for the evaluation of plant growth and leaf color, it may be probably applicable to a wide range of biotic and abiotic stresses and crop species. Moreover our study also supports the use of this technique to assess genotypic differences in grain yield under good agronomical conditions.

Although the performance of the RGB indices (obtained from JPEG images) worked well in this study, future research may address the possibility of further improve their accuracy by using input images saved in a lossless compression format as TIFF or PNG. Despite of storage inconvenient, their larger capability (16 bit per pixel instead of 8 bit) may maintain higher quality detail from the visible spectrum. Another important consideration is the effect of changing light conditions when making these outdoor measurements. Despite of the good strength and repeatability of the results (Table S1) fluctuating ambient lighting should be considered as a possible source of error. Further research should also be targeted toward implementation and evaluation of similar RGB phenotyping methods in remotely piloted aerial platforms (Elazab et al., 2016; Rasmussen et al., 2016).

REFERENCES

Andrade-Sanchez, P., Gore, M. A., Heun, J. T., Thorp, K. R., Carmo-Silva, A. E., French, A. N., et al. (2014). Development and evaluation of a field-based high-throughput phenotyping platform. *Funct. Plant Biol.* 41, 68–79. doi: 10.1071/FP13126

Araus, J. L., Cabrera-Bosquet, L., Serret, M. D., Bort, J., and Nieto-Taladriz, M. T. (2013). Comparative performance of δ13C, δ18O and δ15N for phenotyping durum wheat adaptation to a dryland environment. *Funct. Plant Biol.* 40, 595–608. doi: 10.1071/FP12254

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Araus, J. L., Sánchez, C., and Cabrera-Bosquet, L. (2010). Is heterosis in maize mediated through better water use? *New Phytol.* 187, 392–406. doi: 10.1111/j.1469-8137.2010.03276.x

Araus, J. L., Slafer, G. A., Royo, C., and Serret, M. D. (2008). Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant Sci.* 27, 377–412. doi: 10.1080/07352680802467736

Arendonk, J. J. C. M., and Poorter, H. (1994). The chemical composition and anatomical structure of leaves of grass species differing in relative growth rate. *Plant Cell Environ.* 17, 963–970. doi: 10.1111/j.1365-3040.1994.tb00325.x

Argenta, G., Silva, P. R. F. D., and Sangoi, L. (2004). Leaf relative chlorophyll content as an indicator parameter to predict nitrogen fertilization in maize. *Cienc. Rural.* 34, 1379–1387. doi: 10.1590/S0103-84782004000500009

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., et al. (2000). "Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data," in *Proceedings of the 5th International Conference on Precision Agriculture* (Bloomington, MN), 16–19.

Bateman, A. S., Kelly, S. D., and Jickells, T. D. (2005). Nitrogen isotope relationships between crops and fertilizer: implications for using nitrogen isotope analysis as an indicator of agricultural regime. *J. Agr Food Chem.* 53, 5760–5765. doi: 10.1021/jf050374h

Boegh, E., Søgaard, H., Broge, N., Hasager, C. B., Jensen, N. O., Schelde, K., et al. (2002). Airborne multispectral data for quantifying leaf

area index, nitrogen concentration, and photosynthetic efficiency in agriculture. *Remote Sens. Environ.* 81, 179–193. doi: 10.1016/S0034-4257(01)00342-X

Buerkert, A., Bationo, A., and Piepho, H. P. (2001). Efficient phosphorus application strategies for increased crop production in sub-Saharan West Africa. *Field Crop Res.* 72, 1–15. doi: 10.1016/s0378-4290(01)00166-6

Cabrera-Bosquet, L., Sánchez, C., and Araus, J. L. (2009). How yield relates to ash content, Δ13C and Δ18O in maize grown under different water regimes. *Ann. Bot.* 104, 1207–1216. doi: 10.1093/aob/mcp229

Cairns, J. E., Hellin, J., Sonder, K., Araus, J. L., MacRobert, J. F., Thierfelder, C., et al. (2013). Adapting maize production to climate change in sub-Saharan Africa. *Food Sec.* 5, 345–360. doi: 10.1007/s12571-013-0256-x

Cairns, J. E., Sonder, K., Zaidi, P. H., Verhulst, N., Mahuku, G., Babu, R., et al. (2012). Maize production in a changing climate: impacts, adaptation and mitigation strategies. *Adv. Agron.* 144, 1–58. doi: 10.1016/B978-0-12-394275-3.00006-7

Casadesus, J., Kaya, Y., Bort, J., Nachit, M. M., Araus, J. L., Amor, S., et al. (2007). Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water−limited environments. *Ann. Appl. Biol.* 150, 227–236. doi: 10.1111/j.1744-7348.2007.00116.x

Casadesús, J., and Villegas, D. (2014). Conventional digital cameras as a tool for assessing leaf area index and biomass for cereal breeding. *J. Integr. Plant Biol.* 56, 7–14. doi: 10.1111/jipb.12117

Chen, C., Pan, J., and Lam, S. K. (2014). A review of precision fertilization research. *Environ. Earth Sci.* 71, 4073–4080. doi: 10.1007/s12665-013-2792-2

Delgado, J. A., Follett, R. F., Buchleiter, G., Stuebe, A., Sparks, R. T., Dillon, M. A., et al. (2001). "Use of geospatial information for N management and conservation of underground water quality," in *Proceedings of the 3rd International Conference on Geospatial Information in Agriculture and Forestry* (Denver, CO), 5–7

Dercon, G., Clymans, E., Diels, J., Merckx, R., and Deckers, J. (2006). Differential ^{13}C isotopic discrimination in maize at varying water stress and at low to high nitrogen availability. *Plant Soil.* 282, 313–326. doi: 10.1007/s11104-006-0001-8

Elazab, A., Bort, J., Zhou, B., Serret, M. D., Nieto-Taladriz, M. T., and Araus, J. L. (2015). The combined use of vegetation indices and stable isotopes to predict durum wheat grain yield under contrasting water conditions. *Agr. Water Manage.* 158, 196–208. doi: 10.1016/j.agwat.2015.05.003

Elazab, A., Ordóñez, R. A., Savin, R., Slafer, G. A., and Araus, J. L. (2016). Detecting interactive effects of N fertilization and heat stress on maize productivity by remote sensing techniques. *Eur. J. Agron.* 73, 11–24. doi: 10.1016/j.eja.2015.11.010

Evans, R. D. (2001). Physiological mechanisms influencing plant nitrogen isotope composition. *Trends Plant Sci.* 6, 121–126. doi: 10.1016/S1360-1385(01)01889-1

FAO (2013). *Food and Agriculture Organisation of the United Nations; Statistic Division*. Available online at: http://faostat.fao.org/

Feng, Y. L., Fu, G. L., and Zheng, Y. L. (2008). Specific leaf area relates to the differences in leaf construction cost, photosynthesis, nitrogen allocation, and use efficiencies between invasive and noninvasive alien congeners. *Planta* 228, 383–390. doi: 10.1007/s00425-008-0732-2

Fischer, T., Byerlee, D., and Edmeades, G. (2014). *Crop Yields and Global Food Security*. Canberra, ACT: Australian Centre for International Agricultural Research, Monograph no. 158.

Fox, R. H., Piekielek, W. P., and Macneal, K. M. (1994). Using a chlorophyll meter to predict nitrogen fertilizer needs of winter wheat. *Commun. Soil Sci. Plan.* 25, 171–181. doi: 10.1080/00103629409369027

Furbank, R. T., and Tester, M. (2011). Phenomics–technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Gallais, A., and Coque, M. (2005). Genetic variation and selection for nitrogen use efficiency in maize: a synthesis. *Maydica* 50, 531–547.

Garnier, E., Cordonnier, P., Guillerm, J. L., and Sonié, L. (1997). Specific leaf area and leaf nitrogen concentration in annual and perennial grass species growing in Mediterranean old-fields. *Oecologia* 111, 490–498. doi: 10.1007/s004420050262

Gitelson, A. A., Kaufman, Y. J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80, 76–87. doi: 10.1016/S0034-4257(01)00289-9

Hasegawa, K., Matsuyama, H., Tsuzuki, H., and Sweda, T. (2010). Improving the estimation of leaf area index by using remotely sensed NDVI with BRDF signatures. *Remote Sens. Environ.* 114, 514–519. doi: 10.1016/j.rse.2009.10.005

Hatfield, J. L. (2000). *Precision Agriculture and Environmental Quality: Challenges for Research and Education*. Ames, IA: USDA National Resources Conservation Service.

Hergert, G., Ferguson, R., Gotway, C., and Peterson, T. (1996). "The Impact of VRT-N application on N use efficiency of furrow irrigated corn," in *Proceedings of the 3rd International Conference on Precision Agriculture*, eds P. C. Robert, R. H. Rust and W. E. Larson (Madison, WI), 389–397.

Hirel, B., Le Gouis, J., Ney, B., and Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *J. Exp. Bot.* 58, 2369–2387. doi: 10.1093/jxb/erm097

Hobbs, T. J. (1995). The use of NOAA-AVHRR NDVI data to assess herbage production in the arid rangelands of Central Australia. *Int. J. Remote Sens.* 16, 1289–1302. doi: 10.1080/01431169508954477

Liebisch, F., Kirchgessner, N., Schneider, D., Walter, A., and Hund, A. (2015). Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach. *Plant Methods* 11, 9. doi: 10.1186/s13007-015-0048-8

Machado, A. T., and Fernandes, M. S. (2001). Participatory maize breeding for low nitrogen tolerance. *Euphytica* 122, 567–573. doi: 10.1023/A:1017543426136

Marenco, R. A., Antezana-Vera, S. A., and Nascimento, H. C. S. (2009). Relationship between specific leaf area, leaf thickness, leaf water content and SPAD-502 readings in six Amazonian tree species. *Photosynthetica* 47, 184–190. doi: 10.1007/s11099-009-0031-6

Markwell, J., Osterman, J. C., and Mitchell, J. L. (1995). Calibration of the Minolta SPAD-502 leaf chlorophyll meter. *Photosynth. Res.* 46, 467–472. doi: 10.1007/BF00032301

Masuka, B., Araus, J. L., Das, B., Sonder, K., and Cairns, J. E. (2012). Phenotyping for abiotic stress tolerance in maize. *J. Integr. Plant Biol.* 54, 238–249. doi: 10.1111/j.1744-7909.2012.01118.x

Meziane, D., and Shipley, B. (2001). Direct and indirect relationships between specific leaf area, leaf nitrogen and leaf gas exchange. Effects of irradiance and nutrient supply. *Ann. Bot.* 88, 915–927. doi: 10.1006/anbo.2001.1536

Monneveux, P., Sheshshayee, M. S., Akhter, J., and Ribaut, J. M. (2007). Using carbon isotope discrimination to select maize (*Zea mays* L.) inbred lines and hybrids for drought tolerance. *Plant Sci.* 173, 390–396. doi: 10.1016/j.plantsci.2007.06.003

Morgounov, A., Gummadov, N., Belen, S., Kaya, Y., Keser, M., and Mursalova, J. (2014). Association of digital photo parameters and NDVI with winter wheat grain yield in variable environments. *Turk. J. Agric. For.* 38, 624–632. doi: 10.3906/tar-1312-90

Nautiyal, P. C., Rachaputi, N. R., and Joshi, Y. C. (2002). Moisture-deficit-induced changes in leaf-water content, leaf carbon exchange rate and biomass production in groundnut cultivars differing in specific leaf area. *Field Crop Res.* 74, 67–79. doi: 10.1016/S0378-4290(01)00199-X

Petropoulos, G. P., and Kalaitzidisz, C. (2012). Multispectral vegetation indices in remote sensing: an overview. *Ecol. Model* 2, 15–39.

Poorter, H., and Evans, J. R. (1998). Photosynthetic nitrogen-use efficiency of species that differ inherently in specific leaf area. *Oecologia* 116, 26–37. doi: 10.1007/s004420050560

Rasmussen, J., Ntakos, G., Nielsen, J., Svensgaard, J., Poulsen, R. N., and Christensen, S. (2016). Are vegetation indices derived from consumer-grade cameras mounted on UAVs sufficiently reliable for assessing experimental plots? *Eur. J. Agron.* 74, 75–92. doi: 10.1016/j.eja.2015.11.026

Reich, P. B., Ellsworth, D. S., and Walters, M. B. (1998). Leaf structure (specific leaf area) modulates photosynthesis–nitrogen relations: evidence from within and across species and functional groups. *Funct. Ecol.* 12, 948–958. doi: 10.1046/j.1365-2435.1998.00274.x

Richards, R. A. (2000). Selectable traits to increase crop photosynthesis and yield of grain crops. *J. Exp. Bot.* 51, 447–458. doi: 10.1093/jexbot/51.suppl_1.447

Roberts, R., English, B., and Mahajanashetti, S. (2001). "Environmental and economic effects of spatial variability and weather," in *Proceedings of the 3rd*

European Conference on Precision Agriculture, eds S. Blackmore and G. Grenier (Montpellier), 545–550.

Rorie, R. L., Purcell, L. C., Karcher, D. E., and King, C. A. (2011). The assessment of leaf nitrogen in corn from digital images. *Crop Sci.* 51, 2174–2180. doi: 10.2135/cropsci2010.12.0699

Serret, M. D., Ortiz Monasterio, I., Pardo, A., and Araus, J. L. (2008). The effects of urea fertilisation and genotype on yield, nitrogen use efficiency, $\delta^{15}N$ and $\delta^{13}C$ in wheat. *Ann. App. Biol.* 153, 243–257. doi: 10.1111/j.1744-7348.2008.00259.x

Seul, M., O'Gorman, L., and Sammon, M. J. (2000). *Practical Algorithms for Image Analysis. Description, Examples, and Code*. Cambridge, UK; Cambridge University Press

Stewart, W. M., Dibb, D. W., Johnston, A. E., and Smyth, T. J. (2005). The contribution of commercial fertilizer nutrients to food production. *Agron. J.* 97, 1–6. doi: 10.2134/agronj2005.0001

Svensgaard, J., Roitsch, T., and Christensen, S. (2014). Development of a mobile multispectral imaging platform for precise field phenotyping. *Agronomy* 4, 322–336. doi: 10.3390/agronomy4030322

Trussell, H. J., Vrhel, M., and Saber, E. (2005). Color image processing. *IEEE Signal Proc. Mag.* 22, 14–22. doi: 10.1109/MSP.2005.1407711

Van der Wal, R., Madan, N., Van Lieshout, S., Dormann, C., Langvatn, R., and Albon, S. D. (2000). Trading forage quality for quantity? Plant phenology and patch choice by Svalbard reindeer. *Oecologia* 123, 108-115. doi: 10.1007/s004420050995

Vergara-Diaz, O., Kefauver, S. C., Elazab, A., Nieto-Taladriz, M. T., and Araus, J. L. (2015). Grain yield losses in yellow-rusted durum wheat estimated using digital and conventional parameters under field conditions. *Crop J.* 3, 200–210. doi: 10.1016/j.cj.2015.03.003

Wang, F., Fraisse, C. W., Kitchen, N. R., and Sudduth, K. A. (2003). Site-specific evaluation of the CROPGRO-soybean model on Missouri claypan soils. *Agr. Syst.* 76, 985–1005. doi: 10.1016/S0308-521X(02)00029-X

Wang, R. L., Stec, A., Hey, J., Lukens, L., and Doebley, J., (1999). The limits of selection during maize domestication. *Nature* 398, 236–239. doi: 10.1038/18435

Wezel, A., Casagrande, M., Celette, F., Vian, J. F., Ferrer, A., and Peigné, J. (2014). Agroecological practices for sustainable agriculture. A review. *Agron. Sustain. Dev.* 34, 1–20. doi: 10.1007/s13593-013-0180-7

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crop Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003

Witkowski, E. T. F., and Lamont, B. B. (1991). Leaf specific mass confounds leaf density and thickness. *Oecologia* 88, 486–493. doi: 10.1007/BF00317710

Zaman-Allah, M., Vergara, O., Araus, J. L., Tarekegne, A., Magorokosho, C., Zarco-Tejada, P. J., et al. (2015). Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* 11:35. doi: 10.1186/s13007-015-0078-2

Zhou, B., Elazab, A., Bort, J., Vergara, O., Serret, M. D., and Araus, J. L. (2015). Low-cost assessment of wheat resistance to yellow rust through conventional RGB images. *Comput. Electron. Agr.* 116, 20–29. doi: 10.1016/j.compag.2015.05.017

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership