Current Plant Biology 42 (2025) 100464

Contents lists available at ScienceDirect

Current Plant Biology

journal homepage: www.elsevier.com/locate/cpb

Genome-wide association study (GWAS) of rice (*Oryza sativa* L.) panicle compactness

Masoumeh Kordi [a], Naser Farrokhi [a,1,*], Asadollah Ahmadikhah [a,2,*], Pär K. Ingvarsson [b,3,*], Abbas Saidi [a], Mehdi Jahanfar [a]

[a] Department of Cell & Molecular Biology, Faculty of Life Sciences & Biotechnology, Shahid Beheshti University, Tehran, Iran
[b] Department of Plant Biology, Swedish University of Agricultural Sciences, Uppsala, Sweden

ARTICLE INFO

Keywords:
Epistasis
Haplotype analysis
Linkage disequilibrium
P-TRAP
Protein-protein interaction
RNA-seq
Spikelet per panicle

ABSTRACT

The Green Revolution altered the rice panicle by increasing grain numbers per panicle. Here, we perform a genome-wide association study to understand the molecular mechanisms determining the number of grains in a panicle in rice. Panicle image analyses were performed on 158 genetically diverse rice accessions, and GWAS was run using the FarmCPU model with 34,072 single nucleotide polymorphisms to relate genotypic variation to the corresponding phenotypes. Flanking regions of candidate SNPs were separately defined for each chromosome based on LD decay distance to identify putative-associated genes. An RNA-seq data analysis was performed between stem and panicle to emphasize the role of candidate genes in panicle compactness. The results were further confirmed by a PPI network analysis using the putative candidate genes. In total 95 significant SNPs were identified; as close SNPs were considered a QTL that resulted in 56 QTLs across the 12 rice chromosomes. We identified novel candidate genes for panicle compactness traits, such as *cytochrome P450*, *polygalacturonase*, *glycosyltransferase*, *MADS-box, WRKY*, *YABBY*, *WUSCHEL-related homeobox, protein kinase*, *lipase, zinc finger transcription factor* and *protein phosphatase*. Haplotype analysis identified haplogroups *qNSSBB53*, *qNSSBB102*, *qNSSBU3*, *qNSSBU21*, and *qLS3* for three traits of NSSBB, NSSBU and LS. An analysis of epistatic interactions among candidate SNPs identified 91 significant SNP-SNP interactions.

1. Introduction

Rice is the most important food source for half of the world's population and, for example, it provides 76 % of the calorie intake of South East Asia [16,36,51,76]. Due to global warming, high localized population growth in some parts of the world, and demand for cheaper staple food sources, high-yielding rice cultivars are needed to meet future demand. To meet the predicted needs of the world's population, global rice production needs to increase by more than 1.2 % annually. In Asia, where rice is considered a staple crop, a 43 % increase in output over the next 30 years is required. Increasing rice yield per unit area can address parts of this increasing demand [40,57,71]. Parameters that determine yield include the number of panicles per square meter, the total number of spikelets per panicle (TNSP), the weight of thousand rice grains, and the percentage of filled and ripe grains [18,21,84,87].

In the grass family of plants, spikelets with flower(s) and their arrangements establish inflorescence architecture and yield. Thus, deciphering the molecular mechanisms involved in controlling spikelet development has been the subject of interest to breeders and biologists [90]. Some responsible genes and QTLs that are under the influence of environmental factors, controlling these traits and more specifically TNSP have been reported previously [100,47,79]. Genes controlling vascular development (*Ghd7*, *NAL1*, *qPE9–1/DEP1*, *PlA2/DEP3*), phytohormone pathways (*LOG*, *GNP1*, *PIN*, *CPB1, PMM1*, *FZP*), panicle architecture (*qPE9–1/DEP1*, *RGN1*, *sped1-D*), and branch differentiation (*GN1a*, *CKX2*, *PAP2*) have been highlighted as the genetic determinants of TNSP [45,53,70,78,81].

TNSP-related genes primarily affect panicle traits such as the number and length of branches and spikelet density [57,72,79]. *DEP3* encodes a Patatin-like phospholipase A2 (PLA2) with a role in vascular bundle

* Corresponding authors.
E-mail addresses: n_farrokhi@sbu.ac.ir (N. Farrokhi), a_ahmadikhah@sbu.ac.ir (A. Ahmadikhah), par.ingvarsson@slu.se (P.K. Ingvarsson).
[1] ORCID: 0000–0003-4676–1545
[2] ORCID: 0000–0001-5100–9740
[3] ORCID: 0000–0001-9225–7521

https://doi.org/10.1016/j.cpb.2025.100464
Received 7 January 2025; Received in revised form 5 March 2025; Accepted 5 March 2025
Available online 2 April 2025

formation that defines panicle length and grain number per panicle among other functions [70]. *APO1* (encodes an F-box) is a regulatory protein that controls panicle size and grain number [18,28]. *Gn1a*, identified in a genotype with significantly increased spikelet numbers [4,46], encodes a cytokinin (CK) oxidase/dehydrogenase (CKX) that catalyzes active CK. *DEP1* encodes a noncanonical Gγ subunit and underlies a major QTL for rice panicle architecture and grain size [102,29, 30]. *DEP1* positively regulates rice grain-filling by increasing the auxin and cytokinin contents [93].

GRAIN NUMBER 1.1 (*GN1.1*) has been shown to interact with *OsZAC* (an ADP-ribosylation factor-GTPase-activating protein) that together facilitates polar transport and elevation of panicle auxin content to impact panicle development [97]. *Panicle Morphology Mutant 1* (*PMM1*) is a cytochrome P450 involved in brassinosteroids biosynthesis, along with its paralogues, DWARF11 and DWARF4 in a redundant manner. This gene was reported as being involved in controlling the differentiation of spikelet primordia and panicle branching to define the inflorescence architecture in rice [45]. An aluminum-activated malate transporter (*OsALMT7*) was shown to maintain rice panicle size and grain yield via mutant analysis [26].

TAW1 [18], a member of ALOG gene family, encodes a nuclear protein and works upstream of MADS-box genes to regulate the phase transition by promoting the activity of inflorescence meristems [2,89]. Two other ALOG gene family members, (*OsG1L1* and *OsG1L2*), with similar expressions and mutant phenotypes to TAW1 have been reported with a proposed function in inflorescence development. In the same study, a gene regulatory network (GRN) with *OsG1L2* has established the possible involvement of a homeodomain-leucine zipper TF (*OsHOX14*) in the same pathway as further confirmed via phenotypic analysis of CRISPR loss-of-function mutants [6]. *OsMADS34*, a *Sepallata-like* gene, was shown to control the development of rice spikelets and inflorescence [19]. Taking a combinatory approach of phenotypic QTL and expression QTL analysis followed by a thorough functional analysis of the candidate genes, the roles of *OsMADS18*, a modulatory TF partially determining primary rachis average length, and *OsFLT1* with pleiotropic effects on several panicle traits have been demonstrated [80]. *OsSPL14*, encoding a SQUAMOSA promoter binding TF, is responsible for the negative regulation of phase transition to spikelet meristem and it has a definitive role in controlling plant height, panicle length, spikelet number, grain number, and thousand-grain weight [35, 41,55,60,7]. In another study, *OsSPL9* was shown to be involved in controlling secondary branch formation and grain number per panicle via activation of *RCN1* (*RICE TERMINAL FLOWER 1/CENTRORADIALIS* homolog, [27]). *SCM3* encodes a TCP family TF that acts downstream of strigolactone signaling to inhibit outgrowth of the axillary buds in rice [12]. *SCM3* negatively regulates tiller number per plant and enhances the thickness of the culm wall [88]. *qSSP7* controls the number of spikelets per panicle [85]. Fan et al. [17] reported a *DLT3-MOC3-MOC1* protein complex that binds to the promoter of *FLORAL ORGAN NUMBER 1* (*FON1*) via DLT3, a Dof protein, to positively regulate tiller and panicle numbers.

Epigenetic modification is also involved in controlling panicle development as shown by the study of the *COMPASS*-like complex, a complex of proteins containing WD40 protein (*OsWDR5a*), *TRITHORAX*-like protein (*OsTrx1*) and SET domain group protein 723 (*SDG723*), with histone 3 Lys 4 trimethylation (H3K4me3) activity [33].

In this study, we measure 16 traits related to panicle density (Table 1) and the number of spikelets per panicle in a population of 158 rice genotypes and dissect the genetic basis of these traits using GWAS (Fig. 1). RNA-seq data analysis was used to assess the expression patterns of the candidate genes between panicle and stem. Finally, we used post-GWAS analyses to define the role of associated SNP markers and the underlying candidate genes in controlling panicle density.

Table 1
Trait designation, description and units.

Trait name	Symbol	Trait description	Measurement Unit
Total Number of Spikelets in Panicle	TNSP	total number of spikelets in a panicle from the base to the tip	No./panicle
Lateral Spikelet	LS	total number of spikelets in a panicle from the base to the tip except for terminal spikelet	No./panicle
Terminal Spikelet	TS	total number of spikelets in a panicle from the base to the tip in terminal branches	No./panicle
Total Number of Spikelets in Panicle/ Panicle Length	TNSP/PL	the ratio of the number of spikelets in panicle/ panicle length	No./cm
Total Number of Spikelets in Panicle/ Total Primary Branch Length	TNSP /TPBL	ratio of the number of spikelets in panicle/ primary branch length	No./cm
Total Number of Spikelets in Panicle/ Panicle length + Total Primary Branch Length	TNSP/ PL+ TPBL	the ratio of the number of spikelets in panicle/ panicle length + primary branches length	No./cm
Number of Spikelet Primary Rachis Branches in Up of Panicle	NSPBU	number of spikelets on primary rachis branches in the upper third of the panicle	No./panicle
Number of Spikelet Primary Rachis Branches in Middle of Panicle	NSPBM	number of spikelets on primary rachis branches in the middle third of the panicle	No./panicle
Number of Spikelet Primary Rachis Branches in Bottom of Panicle	NSPBB	number of spikelets on primary rachis branches in the bottom third of the panicle	No./panicle
Number of Spikelet Secondary Rachis Branches in Up of Panicle	NSSBU	number of spikelets on secondary rachis branches in the upper third of the panicle	No./panicle
Number of Spikelet Secondary Rachis Branches in Middle of Panicle	NSSBM	number of spikelets on secondary rachis branches in the middle third of the panicle	No./panicle
Number of Spikelet Secondary Rachis Branch in Bottom of Panicle	NSSBB	number of spikelets on secondary rachis branches in the bottom third of the panicle	No./panicle
Number of Spikelet Primary Branches in Upper Half of Panicle	NSPBUH	number of spikelets on primary rachis branches in the upper half of the panicle	No./panicle
Number of Spikelet Primary Branches in Bottom Half of Panicle	NSPBBH	number of spikelets on primary rachis branches in the bottom half of the panicle	No./panicle
Number of Spikelet Secondary Branches in Upper Half of Panicle	NSSBUH	number of spikelets on secondary rachis branches in the upper half of the panicle	No./panicle
Number of Spikelet Secondary Branches in Bottom Half of Panicle	NSSBBH	number of spikelets on secondary rachis branches in the bottom half of the panicle	No./panicle

2. Material and methods

2.1. Plant material and phenotyping

In this study, 282 rice accessions from 82 countries were obtained from the International Rice Research Institute of the Philippines. The seeds were planted at the Ahwaz Agricultural Research Station, Iran

Fig. 1. Panicle phenotypic diversity in the sampled rice genotypes. The panicles were attached to white cardboard with colorless glue and photographed with a camera fixed at the same distance from the panicles. To obtain phenotypic data, the images were called in P-TRAP software. The phenotypic data were arranged in an Excel sheet and used for subsequent GWAS analyses.

(31°50′N, 48°27′E) during 2022–2023 in a randomized complete design with three replicates (25 × 25 cm spacing in plots of 1 × 2 m size), using standard field management practices throughout the growing season. From this collection, 158 rice accessions (Table S1) produced panicles and these plants were evaluated at the stage of panicle maturity. The panicle-producing accessions contained a mixture of landraces, breeding lines, varieties, and cultivars belonging to the TEJ (temperate japonica), IND (indica), AUS (aus), ARO (aromatic), TRJ (tropical japonica), and ADMIX subpopulations [99]. Ten rice panicles from each accession were collected from the main tillers in the field and transported to the laboratory. Panicle-related traits were measured on three randomly selected panicles of each accession at maturity. Analysis of variance and mean comparisons (Duncan's multiple tests) were carried out for all traits using the Agricolae package at R. For reliable and high-throughput accurate measurements of rice panicle compactness, panicle-related traits were obtained with P-TRAP (Panicle TRAit Phenotyping; [3]) imaging software. The input to P-TRAP was an RGB image of a fixed rice panicle taped onto a white cardboard (Fig. 1). The 16 traits included in this study were selected from the review of similar published articles and were included due to their importance of determining panicle spikelet density (Table 1).

2.2. SNP genotyping data and genome-wide association studies (GWAS)

The development of an SNP array hybridization platform for rice have previously been described by Zhao et al. [99]. Single nucleotide polymorphisms (SNPs) information for all used accessions were downloaded from the Gramene portal (http://gramene.org). The 44 K SNP array was used and ultimately 33,484 high-quality SNP markers were included in our GWAS analyses after filtering on a minor allele frequency greater than 0.05. We assessed the underlying population structure in the population using PCA calculated using rMVP (https://cran.r-project.org/web/packages/rMVP/index.html). GWAS was performed using three different models, MLM, FarmCPU, and GLM, all implemented in rMVP. Based on a comparison of the quantile–quantile (Q–Q) plots for the three statistical models used in the GWAS analyses (MLM, GLM, FarmCPU), we ultimately used the Fixed and Random Model Circulating Probability Unification (FarmCPU) with three principal components (PCs) to account for the underlying population structure in the data set [50,64]. The results of the other two models, *i.e.*, MLM and GLM are not presented here. We generated Manhattan plots for all traits using the $-\log_{10}(p)$ values in the rMVP package in R [64]. SNP markers with $-\log10(p) > 3.5$ and $-\log10(p) > 4$ were considered significant SNPs for different traits based on visual inspection of the quantile-quantile (Q–Q) plots. The threshold number for each trait is based on the location of the break in the QQ plot line for each trait. A phylogenetic tree and a heatmap of the kinship (k) matrix values were created from the numeric SNP data using the Popkin package in R [65]. Linkage disequilibrium decay was estimated among all 158 accessions in the association panel. The resolution of LD was evaluated by performing pairwise calculations of LD using r^2 in a sliding window of 50 markers using TASSEL (https://www.softpedia.com/get/Science-CAD/TASSEL.shtml). The LD decay with physical distance between the SNPs was visualized using ggplot2 in R [69,82]. Candidate genes were displayed on chromosomes using Mapchart (https://www.wur.nl/en/show/mapchart.htm).

2.3. Gene identification

We identified flanking regions surrounding significant SNPs from the

GWAS analyses based on the distance where the r^2 value decayed to half its original value. These flanking regions were then to identify putative candidate genes surrounding the significant SNPs [100]. A detailed literature survey was performed for all the candidate genes that were found to be associated with panicle compactness. Among the candidate genes, some genes showed pleiotropic effects on multiple traits. The broad-sense heritability (H^2b) of each trait was estimated using the heritability package in R based on data from three replicates. The repeatability of these traits was examined using the Gaussian model implemented in rptR package in R (https://cran.r-project.org/web/packages/rpart/index.html).

2.4. Gene expression analysis

The average expression values of candidate genes between reproductive (panicle with 50 biological replicates) and vegetative (stem with 7 biological replicates) tissues were obtained from the rice RNA-seq database (http://ipf.sustech.edu.cn/pub/ricerna/). Expression data of all putative candidate genes were obtained and calculated using FPKM values. The resulting expression data, after batch-effect correction and normalization, was used to identify DEGs using the limma, edgeR, and umap packages in R, with a significance threshold of "adjusted *p*-values < 0.05 and $|\log_2$ fold change$| > 1$. Volcano diagrams of expression data were drawn with tidyverse and ggrepel packages. A heatmap of the expression data from all candidate genes showing significant expression differences was drawn with gplots package. A gene ontology (GO) analysis was performed using PANTHER (https://www.pantherdb.org/geneListAnalysis.do) with the ID list of all significant DEGs. Charts were drawn for the molecular function (MF), biological process (BP), cellular component (CC), and protein class (PC). A protein-protein interaction network was estimated for all DEGs in STRING (https://string-db.org/) using a minimum required interaction score with medium confidence 0.4. The network was confirmed with Cytoscape (V, 3.10.2, algorithm, BottleNeck). Hub genes were colored yellow, and a protein-protein interaction network for the genes with $|\log_2$ FC$= > 3$ was generated in string (https://string-db.org/) and visualized via Cytoscape.

2.5. Haplotype and epistatic analyses

Haplotype analysis was performed for all gene blocks using the GeneHapR package in R. Haplogroups consisting of less than five accessions were excluded. To determine potential genic interactions between different QTLs influencing panicle compactness, epistatic interaction analysis was performed using functional regression with FRGEpistasis [91].

3. Results

3.1. Population structure

Distribution graphs for all traits are shown in Figure S1 and the maximum, minimum, and average phenotypic values of each trait for all accessions are reported in Table S2. The population could be subdivided into six subpopulations *viz*. ADMIX, AROMATIC, AUS, IND, TEJ, TRJ according to Zhao et al. [99]. However, a principal component analysis (Figure S2) suggested that the data could be subdivided into four subpopulations. On the other hand, the kinship matrix, which summarize the distribution of pairwise relative relationship coefficients among accessions based on SNP information, included six subpopulations consisting of temperate *japonica*, *indica*, Aus, aromatic, admix and tropical *japonica* (AROMATIC, AUS, IND, TEJ, TRJ, ADMIX, Figure S3). The pattern of LD decay was different across chromosomes, with the mean LD, based on r^2 values, ranging from 200.1 kbp on chromosome 11–700.2 kbp on chromosome 5 (Table S3). LD decay plots for all chromosomes are depicted in Figure S4. The distributions of SNP markers on 12 chromosomes are shown in Figure S5. SNPs on chromosomes 1 and 7 were the densest and the most scattered, respectively. On average, one SNP in every 10,954 base pairs was scored across the rice genome (Table S3).

3.2. Genome-wide association study results

The total SNP data consisted of 34,072 SNPs and these were used to assess genotype-phenotype relationships with 16 different panicle-related traits. Based on the Manhattan plots (Fig. 2a-e) we identified 95 SNPs with $-\log_{10}(p)$ values exceeding 3.5 and ≥ 4, according to Q-Q plots of each trait (Table 2). The highest marker-trait associations were obtained by FarmCPU (MLM and GLM models were also used - data not shown). The candidate genes, peak markers, $-\log_{10}(p)$ and QTLs are presented in Table 2 and S5. Significant SNPs (95) were distributed across all 12 rice chromosomes (Fig. 2a-e). The functions of candidate genes in the close vicinity of detected SNPs are presented in Table S6 and Fig. 2f.

The SNPs that are shared among traits, indicating possible pleiotropy, are listed in Table S8. Pleiotropy occurs when one gene affects more than one trait and contributes to the genetic correlation between traits [5]. Apart from pleiotropy, linkage disequilibrium due to physical linkage or population structure can also contribute to genetic correlations between traits [8]. The genes with pleiotropic effects, and hence with potential functions in several traits, are reported in Table S9. The genes were *serine/threonine protein phosphatase*, *NF-Y-A*, *peptidase*, *cytochrome P450*, *Yabby*, *NB-ARC*, *pentatricopeptide*, *protein kinase*, *tetratricopeptide*, *DUFs*, *bHLH* transcription factor, *lipase*, and *GDSL*. Repeatability (0.68–0.99) and heritabilities (0.49–0.99) of all traits are reported in Table S10.

3.3. Differentially expressed candidate genes between panicle and stem

A PCA was performed for all RNA-seq data related to the candidate genes identified in the GWAS for two tissue types: *i.e.*, panicle and stem. The data before and after normalization are shown in Figure S5. Gene expression comparisons of the normalized data were performed between the two tissues, panicle and stem, and differentially expressed genes (DEGs) were identified (Fig. 3a, in total 246 DEGs; including 117 up- and 129 down-regulated genes, all DEGs are reported in Table S11) by comparing normalized expression values of the flanking genes between the two tissues (Figure S5). A heatmap was drawn for the expression values of flanking genes (Fig. 3b). Genes with higher fold change ($\log_2$FC $\geq +3$) in the panicle in comparison to stem were cellulose synthase (*Os10g0467800*), TB2/DP1, HVA22 related protein family protein (*Os03g0250200*), thaumatin-like protein (1−4)-β-mannan endohydrolase-like protein (*Os01g0663300*), peptidase S8 subtilisin-related domain containing protein (*Os03g0119300*), microtubule-associated protein TORTIFOLIA1 (*Os02g0739900*), glycoside hydrolase family N-terminal protein (*Os06g0179000*), protein of unknown function DUF plant family protein (*Os03g0246500*), trehalose-6-phosphate synthase (*Os08g0414700*), receptor-like protein kinase 3 (*Os07g0134200*), H0525G02.7 protein (*Os03g0659800*), and peptide transporter (*Os03g0235700*).

Genes down-regulated genes ($\log_2$FC ≤ -3) in panicles were 40S ribosomal protein S13 (*Os07g0572900*), protein of unknown function DUF26 domain-containing protein (*Os02g0734800*), β-glucanase family protein (*Os03g0117300*), Copper chaperone homolog CCH (*Os08g0405700*), chemocyanin precursor (basic blue protein) (Plantacyanin) (*Os02g0731400*). Some genes are novel candidates with reported functional roles in inflorescence growth and development (Table S6).

The functional annotation of GO terms was performed based on molecular function (MF), biological process (BP), cellular component (CC), and protein class (PC) to make sense of the biological and systematic features of 246 DEGs. Figure S6A summarizes the nine

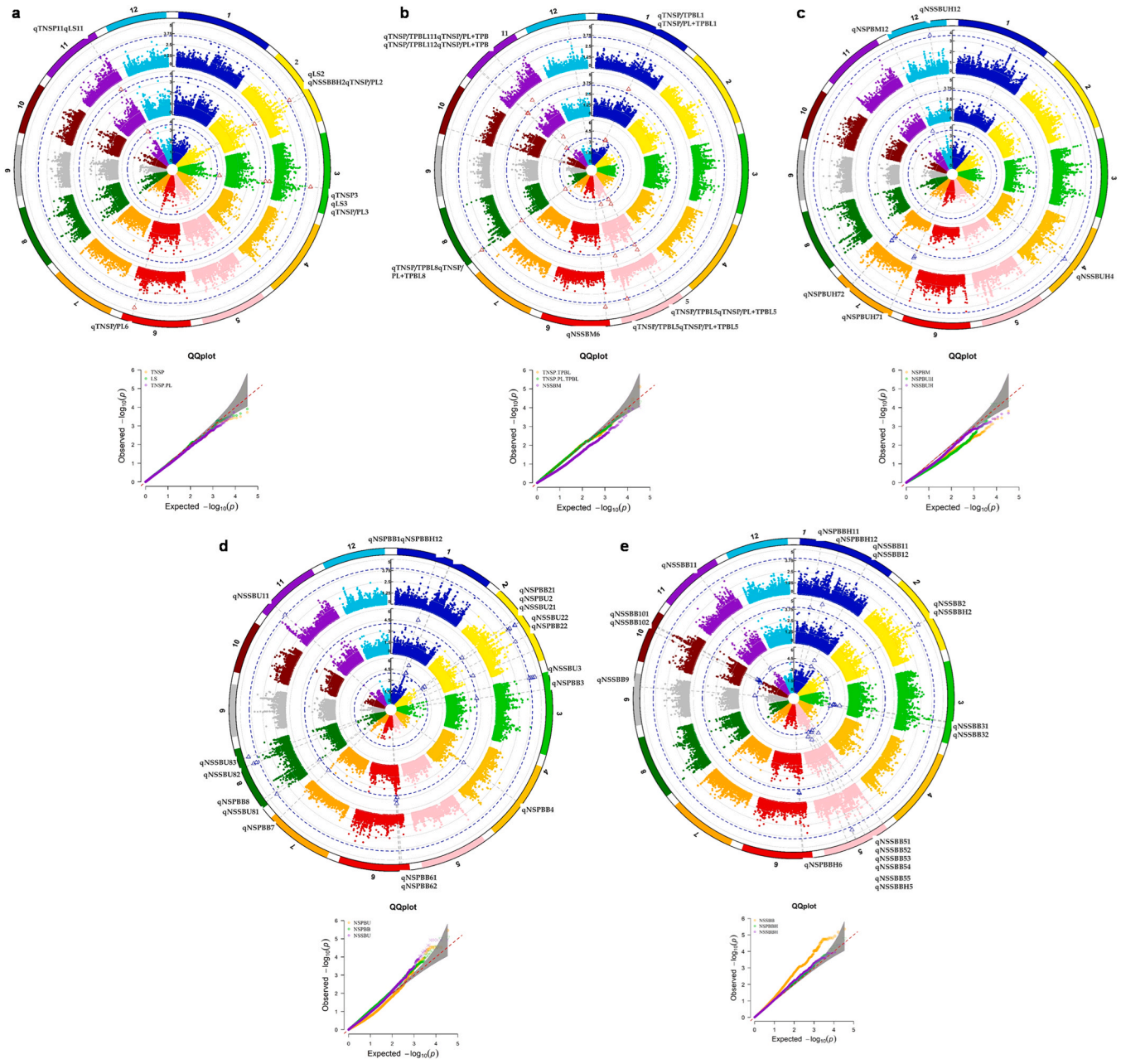

Fig. 2. Multiple Circular-Manhattan plots and Q-Q plots for panicle compactness GWAS and the location of novel candidate genes on 12 rice chromosomes. Significant SNPs are shown as hollow triangles, indicated by grey dashed lines within the radius of each circle. The names of the QTL associated with that SNP with a Koch cut are written on the edge of the circle. The threshold value was set at 3.5 for a, b & c. and 4 for d & e. The traits are shown in circles:(a) from inside to outside include TNSP, LS, TNSP/PL and (b) TNSP/TPBL, TNSP/PL+TPBL, NSSBM and (c) NSPBM, NSPBUH, NSSBUH and (d) NSPBU, NSPBB, NSSBU and (e) NSSBB, NSPBBH, NSSBBH respectively. (f) The hollow boxes represent the rice chromosomes, and the names and locations of the genes on each chromosome are indicated. We have colored the transcription factors [pink (*Zinc finger*); red (*MADS-box*); green (*WRKY*); dark blue (*YABBY*)]. The novel genes include *cytochrome P450*s, *polygalacturonases*, *glycosyltransferases*, *MADS-box*es, *WRKY*s, *YABBY*s, *WUSCHEL-related homeobox*es, *protein kinases, lipases, zinc finger transcription factors, protein phosphatases*.

molecular functions to which all the activities belong. Unclassified genes made up 66.9 % of the data. The other main molecular functions were catalytic activity (Go: 000003824, 15.5 %), antioxidant activity (GO: 0016209, 0.4 %), molecular function regulator activity (GO: 0098772, 0.8 %), molecular transducer activity (GO: 0060089, 0.4 %), structural molecule activity (GO: 0005198, 3.3 %), transcription regulator activity (GO: 0140110, 1.3 %), ATP-dependent activity (GO:0140657, 1.7 %), transporter activity (GO:0005215, 4.6 %) and binding (GO: 0005488, 13.8 %). The overrepresented GO terms for biological processes were unclassified (67.8 %), cellular process (GO: 0009987, 25.5 %), metabolic process (GO: 0008152, 17.2 %), detoxification (GO: 0098754, 0.4 %), biological regulation (GO: 0065007, 4.6 %), localization (GO:0051179, 7.9 %), reproduction (GO:0000003, 0.4 %), reproductive process (GO:0022414, 0.4 %), interspecies interaction between organisms (GO: 0044419, 0.4 %) and response to stimulus (GO: 0050896, 5.9 %) (Figure S6B). The main cellular component terms in which these targets perform their function was categorized as 59.8 % unclassified (Figure S6C), cellular anatomical entity (GO: 0110165, 38.9 %), and

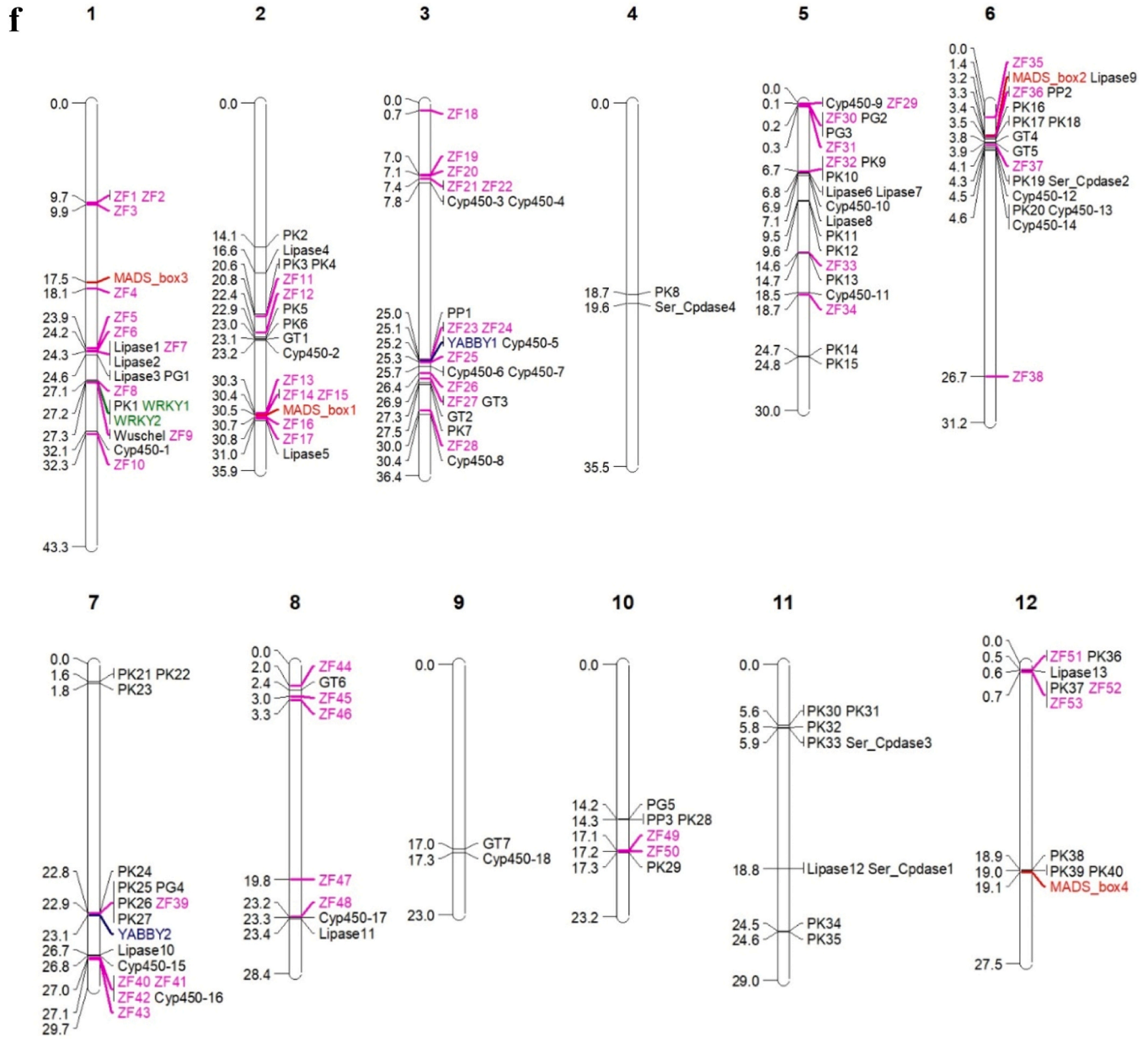

Fig. 2. (*continued*).

protein-containing complex (GO: 0032991, 12.1 %). For protein class terms (Figure S6D), 49.4 % were unclassified. The other main players belong to metabolite interconversion enzyme (PC00262, 15.1 %), DNA metabolism protein (PC00009, 1.3 %), calcium-binding protein (PC00060, 0.4 %), chaperone (PC00072, 2.1 %), chromatin/chromatin binding or regulatory protein (PC00077, 0.8 %), cytoskeletal protein (PC00085, 2.1 %), defense/immunity protein (PC00090, 0.4 %), gene-specific transcriptional regulator (PC00264, 3.8 %), translational protein (PC00263, 5.4 %), RNA metabolism protein (PC00031, 2.9 %), protein binding activity modulator (PC00095, 1.7 %), scaffold/adaptor protein (PC00226, 0.8 %), transfer/carrier protein (PC00219, 0.4 %), translational protein (PC00263, 5.4 %), transmembrane signal receptor (PC00197, 1.3 %), transporter (PC00227, 4.2 %), protein modifying enzymes (PC00260, 5 %), metabolite interconversion enzyme (PC00262, 15.1 %), and membrane traffic protein (PC00150, 2.9 %).

The PPI network of all DEGs (246) is depicted in Fig. 5a-b. A protein-protein interaction network was constructed for DEGs with $\log_2$FC $\geq +3$. By PPI analysis, the hub proteins with the highest number of edges could be classified into eight protein groups according to their function. Interestingly, they were mostly involved in RNA processing (nuclear ribonucleoproteins), DNA replication (DEAD-box ATP-dependent RNA helicase), protein synthesis (40S ribosomal proteins, 60S ribosomal proteins, Elongation Factor Tu, translation initiation factor) and turnover (Ubiquitin-60S ribosomal protein, proteasome), genome stability (Histone H4), energy management (ATP synthase), cell cycle regulation (AMEIOTIC1), intracellular compartment pH control (Vacuolar-type proton ATPase) and some unknown proteins (Fig. 5a). Another PPI network was generated for DEGs with $\log_2$FC $\leq$ -3. The proteins in this network were β-glucanase family protein, similar to 40S ribosomal protein S13, similar to microtubule-associated protein TORTIFOLIA1, and similar to (1−4)-β-mannan endohydrolase-like protein (Fig. 5b).

3.4. Haplotype analysis

Haplotype analysis of NSSBB (Number of Spikelet Secondary Rachis

Table 2
Peak marker list.

Trait	Chromosome	Peak marker	Position (bp)	QTL
TNSP	3	id3010981	25182779	*qTNSP3*
	11	id11009673	24485844	*qTNSP11qLS11*
NSSBM	6	id6000857	1174430	*qNSSBM6*
NSSBB	1	id1019517	32026514	*qNSSBB12*
	1	id1019604	32108071	*qNSSBB11*
	2	id2007913	20377685	*qNSSBB2*
	3	id3011335, id3011351	26946603, 26964732	*qNSSBB31*
		id3011358, id3011500	26967947, 27292413	
		id3011502, id3011506	27294316, 27304002	
		id3011493, id3011260	27278957, 26830010	
		id3011485	27274022	
	3	id3014260	30198305	*qNSSBB32*
	5	id5000043, id5000094	101950,186388	*qNSSBB51*
		id5000097	186597	
	5	id5003459	6826167	*qNSSBB55*
	5	id5004778	9736974	*qNSSBB52*
	5	id5005867, id5005869	14240248, 14240906	*qNSSBB53*
		id5005876, id5005883	14298814, 14355479	
		id5005937, id5005959	14440907, 14450796	
		wd5002107	14547125	
	5	id5007371	18435915	*qNSSBB54*
	9	id9005493	17148241	*qNSSBB9*
	10	id10003825, id10003776	14487961, 14418751	*qNSSBB101*
		id10003777	14418925	
	10	id10004679, id10004767	16751682, 16982012	*qNSSBB102*
		id10004974, id10004985	17342968, 17377688	
	11	id11007060	18773506	*qNSSBB11*
NSPBU	1	id1014144, id1014344	24055495, 24316976	*qNSPBU1*
		id1014414, id1014425	24373975, 24376323	
	2	id2006499, id2006513	16100817, 16125288	*qNSPBU2*
		id2006538, ud2000899	16201982, 16251432	
		id2006577, id2006552	16313676, 16259296	
NSPBM	12	id12000148	435265	*qNSPBM12*
NSPBB	1	id1010751	17791344	*qNSPBB1 qNSPBBH12*
	2	wd2001147	13900471	*qNSPBB21*
	2	id2013634	30626029	*qNSPBB22*
	3	id3004022	7428049	*qNSPBB3*
	4	id4005406	18572034	*qNSPBB4*
	6	id6002713	3319644	*qNSPBB61*
	6	ud6000141, id6003218	3959748, 4367518	*qNSPBB62*
		id6002713	3319644	
	7	id7005370	27081531	*qNSPBB7*
	8	id8000944	3047925	*qNSPBB8*
NSPBUH	7	ud7000123, id7000308	1690767, 1627513	*qNSPBUH71*
	7	id7003993, id7004048	22932237, 23096454	*qNSPBUH72*
		id7004061	23109129	
NSPBBH	1	id1007155, id1007156	9633974, 9634128	*qNSPBBH11*
	1	id1010751	17791344	*qNSPBB1 qNSPBBH12*
	6	ud6000141, id6003218	3959748, 4367518	*qNSPBBH6*
NSSBBH	2	id2007949	20587834	*qNSSBBH2 qTNSP/PL2*
	5	id5005867	14240248	*qNSSBBH5*
LS	2	id2008980	22597869	*qLS2*
	3	id3010981, id3010993	25182779, 25248551	*qLS3*
		id3011000	25253030	
	11	id11009673	24485844	*qTNSP11qLS11*
TNSP/PL	2	id2007949	20587834	*qNSSBBH2 qTNSP/PL2*
	3	id3010993, id3011000	25248551, 25253030	*qTNSP/PL3*
	6	id6015076	26857644	*qTNSP/PL6*
TNSP/ TPBL	1	id1015770, id1015767	27083820, 27082764	*qTNSP/ TPBL1*
	5	id5003529, id5003459	6994407, 6826167	*qTNSP/ TPBL5 qTNSP/PL+TPBL51*
	5	id5011416	24581277	*qTNSP/ TPBL5 qTNSP/PL+TPBL52*
	8	id8000663	2154333	*qTNSP/ TPBL8 qTNSP/PL+TPBL8*
	11	id11002180	5340175	*qTNSP/ TPBL111qTNSP/PL+TPBL111*
	11	id11005231	15338678	*qTNSP/ TPBL112 qTNSP/PL+TPBL112*
TNSP/PL+TPBL	1	id1015770	27083820	*qTNSP/PL+TPBL1*
	5	id5003529, id5003459	6994407, 6826167	*qTNSP/ TPBL5 qTNSP/PL+TPBL51*
	5	id5011416	24581277	*qTNSP/ TPBL5 qTNSP/PL+TPBL52*
	8	id8000663	2154333	*qTNSP/ TPBL8 qTNSP/PL+TPBL8*
	11	id11002180, id11002573	5340175, 5890369	*qTNSP/ TPBL111qTNSP/PL+TPBL111*
	11	id11005231	15338678	*qTNSP/ TPBL112 qTNSP/PL+TPBL112*
NSSBUH	4	id4005826	19715224	*qNSSBUH4*
	12	id12006366	19116715	*qNSSBUH12*
NSSBU	2	id2006552, id2006523	16259296,16161784	*qNSSBU21*
		id2006499, id2006513	16100817, 16125288	

(*continued on next page*)

Table 2 (*continued*)

Trait	Chromosome	Peak marker	Position (bp)	QTL
		id2006538, ud2000899	16201982, 16251432	
		id2006577	16313676	
	2	id2009201	23099368	*qNSSBU22*
	3	ud3000042, id3000536	906506, 909867	*qNSSBU3*
		id3000527, id3000506	906918, 877707	
	8	id8000370	1289777	*qNSSBU81*
	8	id8005260, id8005219	19642854, 19600803	*qNSSBU82*
		id8005235, id8005256	19609178, 19639358	
	8	id8006704	23402228	*qNSSBU83*
	11	id11002387	5675109	*qNSSBU11*

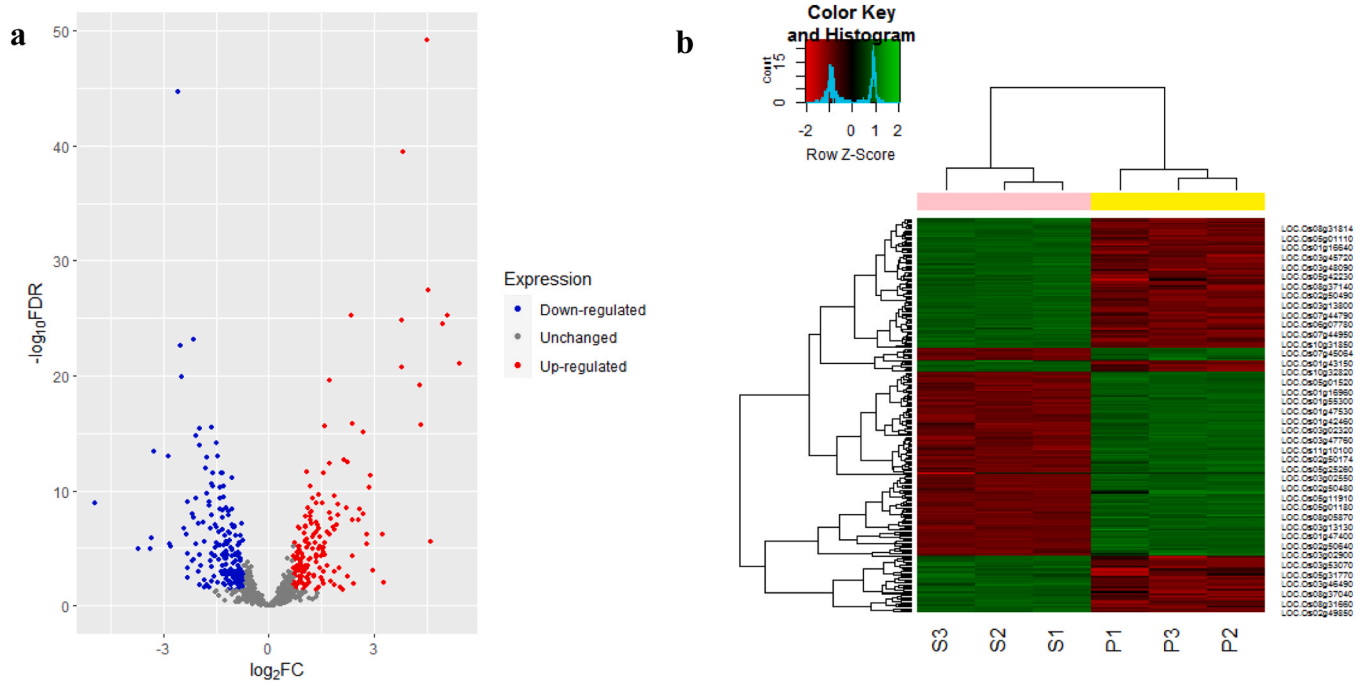

Fig. 3. (a) Volcano plot showing the significantly up- and down-regulated genes between panicle and root. The logarithm of fold change (log_2) for each transcript is presented on the x-axis, and the log10 of the *p*-values (false discovery rate) is displayed on the y-axis. The number of transcripts with increased expression was 117, the number of transcripts with decreased expression was 129, and the number of transcripts with non-significant expression difference was 719. (b) Heatmap of transcripts between panicle (3) and root samples (11). P11, P12 and P13 are the panicle samples (yellow bar at the top of the graph) and S1, S2 and S3 are stem samples (pink bar). Red represents lower expression, green represents high expression, columns represent individual experiments, and rows represent candidate genes with up- and down-regulations. Two-dimensional hierarchical clustering classifies 246 differentially expressed profiles.

Branch in Bottom) showed two haplotypes of *qNSSBB53* on chromosome 5 [H001 (present in 107 accessions with wide distribution around the world) and H002 (present in 48 accessions originating from Asia and West Africa)] (Figure S7) with significant difference between the two as illustrated by a violin plot. H001 mostly contains TEJ, TRJ and ADMIX accessions, while H002 mostly contains accessions from IND, AUS and ADMIX (Figure S7). Furthermore, two significantly different haplotypes of *qNSSBB102* were found on chromosome 10 with H001 present in 130 accessions and H002 present in 22 accessions (Figure S8). According to geographic distribution, H002 was found mostly in accessions belonging to AUS, Aromatic, IND and ADMIX that originated in Asia or West Africa, while H001 mostly contained TEJ, TRJ, IND and ADMIX accessions that were distributed worldwide (Figure S8).

Haplotype analysis of NSSBU (Number of Spikelet Secondary Rachis Branch in Upper) showed two significant haplotypes in the genomic region *qNSSBU21* on rice chromosome 2 with H001 present in 99 accessions and H002 present in 57 accessions (Fig. 4g). Haplotype H002 mostly contains accessions belonging to IND, AUS, ADMIX and Aromatic that originate in Asia, west or South Africa while H001 mostly contains accessions belonging to TEJ, TRJ and ADMIX that have a worldwide distribution (Fig. 4e). Furthermore, three haplotypes, H001 present in 57 accessions, H002 present in 57 accessions, and H003 present in 37 accessions were identified on chromosome 3 for *qNSSBU3* (Figure S9g). Haplotypes H002 and H003 did not show any significant differences in trait values. Haplotype H002 mostly contains accessions belonging to IND, AUS, ADMIX and Aromatic with an origin in Asia and Africa while H001 mostly contains accessions from TRJ, TEJ and ADMIX and H003 mostly contains accessions from TEJ and ADMIX (Figure S9).

A haplotype analysis of LS (Lateral Spikelet) identified three haplotypes, H001 present in 129 accessions, H002 present in 16 accessions and H003 present in 13 accessions, in the region qLS3 on chromosome 3 (Figure S10b). Here, the two haplogroups H001 and H003 showed significant differences in trait values while H001 and H002 were not significantly different. Individuals carrying haplotype H002 (AUS and TRJ) belong to the accessions that originated in Asia, while individuals carrying haplotype H003 contain accessions from ADMIX, TRJ, TEJ and IND that are predominantly found in Asia, west Africa, and Central America, while H001 carrying individuals belong to TEJ, TRJ, IND and ADMIX which are found everywhere (Figure S10).

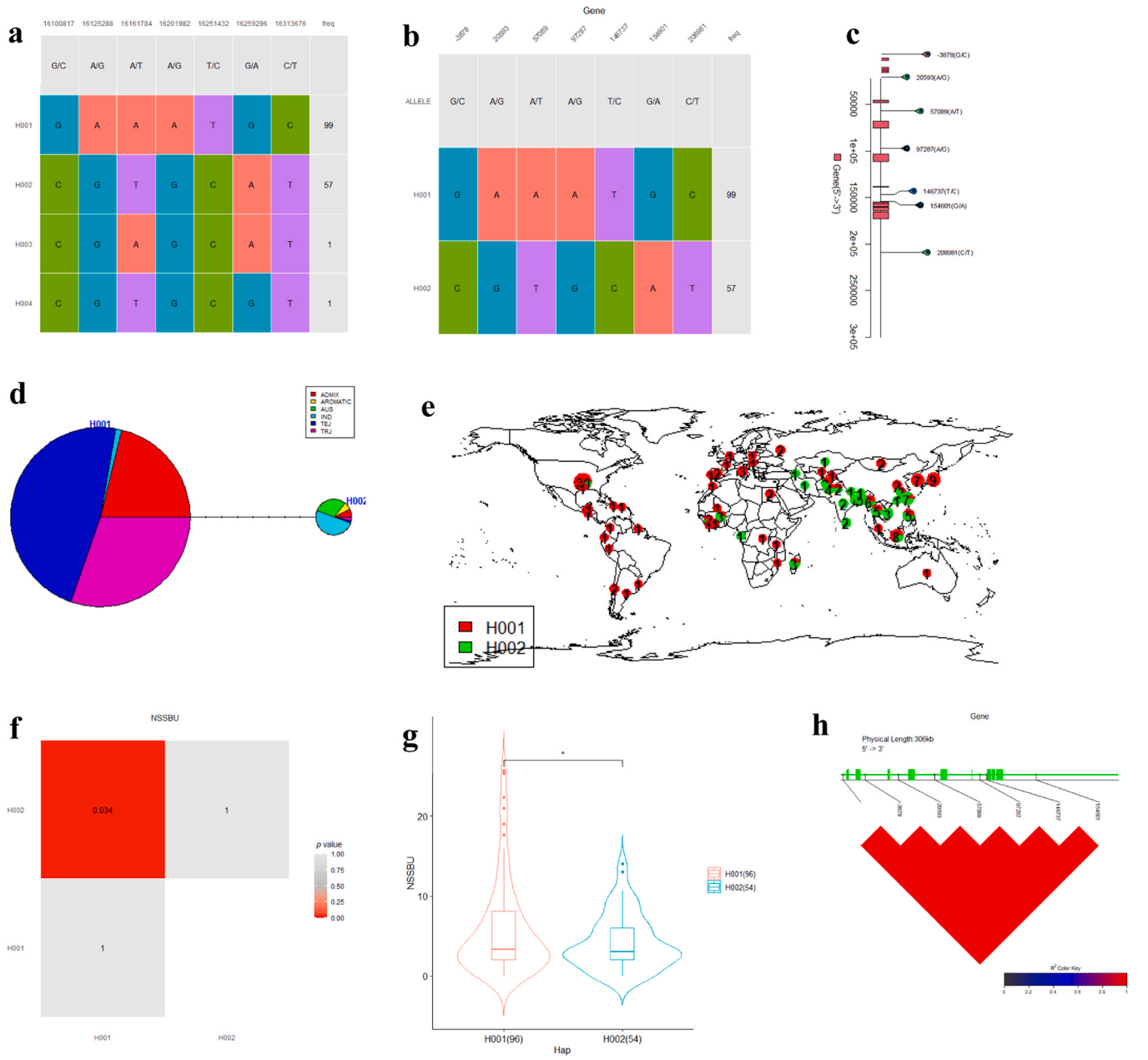

Fig. 4. Haplotype analysis of qNSSBU21 region for NSSBU (Number of Spikelet Secondary Rachis Branch in Upper) located on Chromosome 2. a) four haplogroups were observed in the 158 accessions based on nine SNPs. Colored boxes are haplotype variants with the corresponding frequency in the last column. * *significant at $p \leq 0.01$. b) Two haplogroups were represented by more than five accessions and the SNP positions are given in the upper box c) The gene models associated with qNSSBU21. Visualization of variant positions is given above the gene models with black lines representing the chromosome and rectangles representing upstream and downstream variants. (d) The haplotype network of qNSSBU21. Each circle represents a haplotype, and the size indicates the number of accessions carrying that haplotype. Different colored pie charts represent the origin of accessions in each haplotype. The symbol on the line between haplotypes represents the number of variants differentiating the two haplotypes. (e) Geographic distribution of the major haplotypes at qNSSBU21 where circle sizes represent accession counts and the pie charts in different colors represent the ratios of classified haplotype categories for relevant accessions derived from different geographic regions. H002 is found mostly in Asia and West Africa, H003 is found in Asia and West Africa, while H001 is found almost everywhere. f) The heatmap for phenotypic differences between haplogroups is shown in panel d. g) Violin plot showing the comparison of LS among accessions carrying the three different haplotypes at qNSSBU21 (h) An LD block visualization of each site in haplotype region. The LD block indicates the close relationship between the SNPs of this haplotype block. The gene model is presented at the top of the plot, the line represents the genome and the rectangle represents the upstream and the downstream gene variants.

3.5. Epistatic interaction analysis

The epistasis analyses between our candidate genes identified 91 significant SNP-SNP interactions (Zinc finger, peptidase, serine/threonine protein phosphatase, pentatricopeptide, ubiquitin, MADS-box, threonine endopeptidase, heat shock protein, Armadillo, FAR1, Protein kinase, Leucine-rich repeat plant specific containing protein, F-box domain cyclin-like domain-containing protein and others) for the LS, NSPBB, NSPBUH, NSPBU, NSSBU, TNSP/PL traits (Table S12). A circos plot visualizing the epistatic interactions is shown in Fig. 5c.

4. Discussion

One of the most important goals of rice breeding programs is to

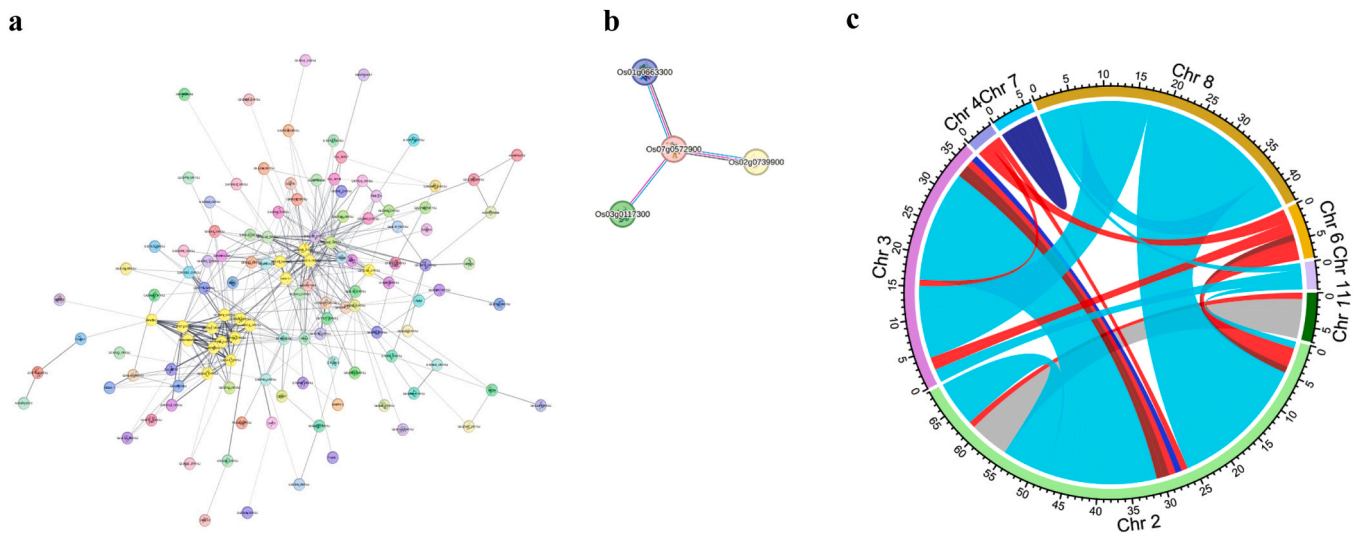

Fig. 5. Protein-protein interaction (PPI) network analysis of all differentially expressed genes associated with panicle formation in rice (a). Hub genes (a) include 40S ribosomal proteins (*Os02g0478600, Os02g0478700, Os03g0249400, Os03g0667800, Os04g0399000, Os07g0572900, Os10g0411800*), 60S ribosomal proteins (*Os02g0591700, Os01g0276000, Os05g0207300, Os06g0181566, Os10g0466700*), nuclear ribonucleoproteins (*Os05g0314100, Os02g0586500, Os03g0240800, Os03g0241200*), Histone H4 (*Os10g0418000, Os03g0119900*), Elongation factor (*Os02g0595700, Os05g0104800*), Ubiquitin-60S ribosomal protein (*Os03g0234200, Os06g0167600*), cDNA clone (*Os05g0498400, Os08g0156800, Os10g0411700, Os10g0465800, Os02g0736600, Os08g0412200, Os08g0416100, Os03g0650400*), ATP synthase (*Os05g0106100, Os08g0478200*). With minimum required interaction score, medium confidence 0.4 in STRING. Protein-protein interaction (PPI) network analysis of DEGs with logFC > 3 (b) list of all genes with logFC > 3 reported in Table S12b accessions are included β-glucanase family protein (*Os03g0117300*), similar to 40S ribosomal protein S13 (*Os07g0572900*), similar to microtubule-associated protein TORTIFOLIA1 (*Os02g0739900*), similar to (1−4)-β-mannan endohydrolase-like protein (*Os01g0663300*). (c) Circus plot of epistatic interactions between chromosomes. Colored crescents around the circle represent different chromosomes (Dark green, light green, light purple, dark purple, dark blue, turquoise, light mustard, dark mustard). The internal area of the circus plot depicts significant epistatic interactions between SNPs affecting panicle compactness. The light blue line is the interaction between two SNPs related to the LS, and dark blue lines are the interactions between three SNP pairs related to NSPBUH. Turquoise lines are the interactions between 65 SNP pairs related to NSSBU. Dark red lines are the interactions between three SNP pairs related to TNSP/PL. Dark grey lines are the interaction between six SNP pairs related to NSPBU. Red lines are the interactions between 11 SNP pairs related to NSPBB (Full details are given in Table S12).

increase the number of spikelets in a panicle by producing a more compact panicle [14,39]. A molecular understanding of the traits in producing compact panicles helps to improve rice yield. In this study, we used GWAS to identify candidate genes for controlling panicle architecture in a collection of 158 diverse rice accessions. To perform GWAS analyses for the 16 panicle-related traits, we used the FarmCPU method due to its high statistical power [101]. The LD decay patterns differed among 12 rice chromosomes [54]. LD decay valuesoften differ between studies due to differences in population size, genetic structure and size, and the number and density of SNP markers [25,73,96]. All panicle compactness traits had high heritability values, similar to what has been observed in earlier studies [1,21]. In total, across the 16 traits, we found 95 significant SNPs that could be grouped into 56 unique QTLs. The lists of all accessions and candidate genes are reported in Table S13. By analysis of highly expressed genes in panicle and looking at PPI networks, it can be seen that the dynamic of growth and development is at a high state. Accordingly, all the mechanisms involved in cell cycle regulation (RNA processing, DNA replication, protein synthesis, ATP synthesis) are active, and the corresponding proteins can be considered as the hub proteins. On the other hand, genes that are involved in cell wall hydrolysis are being down-regulate to possibly allow better cell division and proliferation.

4.1. Candidate genes

Previously reported genes that have been implicated in influencing rice panicle compactness and that were found to be associated with traits in our study include a *serine/threonine-protein kinase* [43,74], a *tetratricopeptide* [62], a *pentatericopeptide* [34], a *glutamine synthase* [63], and transcription factors from the *GATA* family [92], *helix-loop-helix* [56] and *NF-Y* ([15,59] (Table 3). *Serine/threonine-protein kinase*, *tetratricopeptide*, *glutamine synthase*, and *NF-Y* genes were also found in our genes showing differential expression (DEGs) between panicle and stem.

4.2. Novel candidate genes

In our GWAS on rice panicle compactness, we identified several genes that have not previously been reported to be associated with the traits of the study. These genes were considered novel genes affecting panicle-related traits and include a *cytochrome P450* [24,45], *polygalacturonase* [32], *glycosyltransferase* [95,61], *MADS-box* [38,52,67,75], *WRKY transcription factor* [42,83], *YABBY* [77], *WUSCHEL-related homeobox* [10,11], *protein kinase* [58,66], *lipase* [48,56], *Zinc finger protein* [103,86], *protein phosphatase* [68] (Table 4). Of these, *cytochrome P450 CYP94E4, WRKY transcription factor 16, Receptor-like protein kinase 3, Zinc finger CCHC retroviral-type domain-containing protein* and protein phosphatase 2 C family protein genes were differentially expressed in panicle compared to stem.

Cytochrome P450 (found in flanking SNPs associated with NSSBU, NSSBB, NSSBBH, NSPBU, NSPBBH, TNSP, TNSP/PL, TNSP/TPBL, TNSP/PL+TPBL and LS) has been implicated in BRs biosynthesis with role in spikelet primordia differentiation and panicle branch patterning [45], rice panicle development, rachis growth, and plant height (Table S6; [24]). *Polygalacturonase*, found in the flanking region of the SNPs associated with NSSBB, NSPBU and NSPBUH, has previously been shown to regulate floral organ numbers in rice (Table S6; [32]). *Glycosyltransferase*, found in the flanking region of the SNPs associated with NSPBBH, NSSBU, NSSBB and TNSP/PL+TPBL, plays critical roles in rice height and spikelet fertility [95] and pollen wall formation (Table S6; [61]).

MADS-box genes are involved in rice floral organ development [52], palea development, spikelet morphology [67], panicle length and spikelet fertility [52], embryogenesis and flower development, spikelet development [75] and formation (Table S6; [38]). We identified several

Table 3
Gene list of candidate genes in our study, that reported in previously research.

Gene name	ref	Traits in other research	Trait in our research
Serine/threonine-protein kinase	Sekhar et al., [43,74]	inflorescence architecture, grain filling, compact-panicle rice	NSBU, NSSBU, NSSBB, NSPBU, NSPBB, NSPBM, NSSBBH, NSPBUH, TNSP/TPBL, TNSP/PL, TNSP/PL+TPBL and LS
Tetratricopeptide	Ni et al., [62]	spikelet number	NSSBB, NSPBB, NSPBBH, NSPBU, NSSBBH, NSSBU, TNSP/PL and LS
Pentatericopeptide	Jiang et al., [34]	development of inflorescence branches, number of spikelets	NSSBM, NSSBB, NSPBU, NSPBM, NSPBB, NSPBB, NSPBBH, NSSBU, NSSBU, TNSP/TPBL, TNSP/PL+TPBL and LS
GATA transcription factor	Zhang et al., [92]	grain number per panicle, heading date	NSSBB, TNSP/TPBL and TNSP/PL+TPBL
Helix-loop-helix	Luo et al., [56]	number of grain, grain size, awn development,	NSSBB, NSPBM, NSPBB, NSSBBH, NSSBU, TNSP/PL, TNSP/TPBL and TNSP/PL+TPBL
Glutamine synthase	Obara et al., [63]	panicle weight, number of spikelets in the rice panicle	NSSBB and NSPBB
NF-Y family of TFs	Das et al., [15,59]	panicle development, seed size, grain width and length, number of leaves, plant height, compactness of the panicle structure, panicle branches	NSSBB, TNSP, TNSP/PL and LS

Table 4
Gene list of novel genes in our study for panicle traits.

Gene name	ref	Traits in other research	Trait in our research
Cytochrome P450	Guo et al., [24,45]	differentiation of spikelet primordia, patterns of panicle branches, panicle development, panicle growth, main rachis growth, plant height	NSSBU, NSSBB, NSSBBH, NSPBU, NSPBBH, TNSP, TNSP/PL, TNSP/TPBL, TNSP/PL+TPBL and LS
Polygalacturonase	Jang et al., [32]	floral organ number	NSSBB, NSPBU and NSPBUH
Glycosyltransferase	Moon et al., [95,61]	plant height, spikelet fertility, pollen wall formation	NSPBBH, NSSBU, NSSBB and TNSP/PL+TPBL
MADS box	Prasad et al., [38,52,67,75]	floral organ development, panicle length and spikelet fertility, palea development, spikelet morphological	NSSBUH and NSPBB
WRKY transcription factor	Xiang et al., [42,83]	panicle development, seed size, stem elongation	TNSP/TPBL and TNSP/PL+TPBL
YABBY	Tanaka et al., [77]	lateral organ development, spikelet development	NSPBUH, TNSP, TNSP/PL and LS
WUSCHEL-related homeobox	Cho et al., [10,11]	organ development, lemma and palea morphogenesis in spikelets	TNSP/PL and TNSP/PL+TPBL
Serine carboxypeptidase	Li et al., [98,44]	cell division, grain size and the number of secondary branches	NSSBU, NSSBB, NSPBB, NSPBBH and NSSBUH
Protein kinase	Manimaran et al., [58,66]	development of panicle apical spikelets, grain filling	NSSBB, NSPBU, NSPBM, NSPBB, NSPBUH, NSPBBH, NSSBBH, NSSBUH, NSSBU, TNSP, TNSP/TPBL, TNSP/PL+TPBL and LS
Lipase	Li et al., [48,56]	tillering, plant height, and spikelet fertility, expansion and cell division, branch and spikelet fertility	NSSBB, NSPBU, NSPBM, NSSBU, TNSP/TPBL and TNSP/PL+TPBL
Zinc finger	Xu et al., [103,86]	spikelet development	NSSBU, NSSBB, NSPBU, NSPBM, NSPBB, NSPBBH, NSSBBH, TNSP/PL, TNSP/TPBL, TNSP/PL+TPBL and LS.
Protein phosphatase	Qi et al., [68]	cell proliferation and grain size determination by affecting cell cycle proteins such as Cyclin-T1	NSSBB, NSPBU, NSPBM, NSPBUH, NSPBBH, NSSBU, TNSP/PL

MADS-box genes in the flanking region of SNP associated with NSSBUH and NSPBB. *WRKY transcription factor*, found in the flanking region of SNPs associated with TNSP/TPBL and TNSP/PL+TPBL, has a reported role in panicle development (Table S6; [83]). *YABBY*, found in the flanking region of SNP associated with NSPBUH, TNSP, TNSP/PL and LS, plays a regulatory role in flower development in *Arabidopsis*. In rice, the genes have been reported to be involved in lateral organ development and the maintenance of meristem organization in spikelets (Table S6; [77]). *WUSCHEL-related homeobox* genes, found in the flanking region of SNPs associated with TNSP/PL and TNSP/PL+TPBL, are involved in reproductive organ development in rice [10] and lemma and palea morphogenesis in spikelets (Table S6; [11]). There are reports on the role of zinc finger proteins in regulating spikelet development in rice (Table S6; [86,103]). This *zinc finger protein* gene was found in the flanking region of SNPs associated with NSSBU, NSSBB, NSPBU, NSPBM, NSPBB, NSPBBH, NSSBBH, TNSP/PL, TNSP/TPBL, TNSP/PL+TPBL and LS.

Serine carboxypeptidase, found in the flanking region of SNP associated with NSSBU, NSSBB, NSPBB, NSPBBH and NSSBUH has been shown to have roles in cell division and grain size [44], and the number of secondary branches (Table S6; [98]). *Protein kinase* genes, found in the flanking region of the SNPs associated with NSSBB, NSPBU, NSPBM, NSPBB, NSPBUH, NSPBBH, NSSBBH, NSSBUH, NSSBU, TNSP, TNSP/TPBL, TNSP/PL+TPBL and LS, have reported roles in apical spikelet development [66] and rice grain filling (Table S6; [58]). *Serine/threonine protein phosphatase* gene likely plays a role in cell proliferation and grain size determination by affecting cell cycle proteins such as Cyclin-T1 (Table S6; [68]). Isoforms of lipase, found in the flanking region of SNPs associated with NSSBB, NSPBU, NSPBM, NSSBU, TNSP/TPBL and TNSP/PL+TPBL, have previously been demonstrated to regulate spikelet development [48], tillering, plant height, and spikelet fertility in rice (Table S6; [56]).

4.3. Haplotype and epistatic analysis

A haplotype is a collection of alleles, SNPs or other types of variation such as indels, that are inherited together and are physically located in the same region of a chromosome [13]. Having knowledge of haplotypes is useful when concurrent breeding is performed on more than one trait and developing markers that can differentiate haplotypes will therefore

help increase the speed of breeding during a selection period. Considering the fact that the high spikelet density at the top of the rice panicle was the selected trait in our study, it is interesting to note that four of the haplotype analyses that showed significant differences among haplotypes refer to traits affecting the density spikelet numbers at top or bottom of the panicle. From these haplotype studies, it is clear that haplotypes H002 in the *qNSSBB53* and *qNSSBB102* genomic regions (traits that serve as indicators of the number of spikelets at the bottom of the panicle) contained the most desirable rice accessions with the desirable compact panicle form. Between *qNSSBU21* and *qNSSBU3* (indicators of the number of spikelets at the up of the panicle) haplotypes, haplotypes H001 and H003 contained accessions with preferable trait values. For *qLS*, an indicator of panicles with high numbers of pikelet per panicle, the H003 haplotype is preferable due to its greater performance for the trait in question.

In our epistasis analyses, several genes co-expressed, highlighting possible interactions that together help define panicle compactness. Both *Helix-loop-helix* and *Spl* genes were detected in the epistasis analyses. According to previous studies, these two genes are responsible for lateral spikelet production on secondary branches in rice [49]. In rice, *OsMPK6* was shown to directly interact with zinc finger protein by phosphorylating it to modulate phytohormone homeostasis and influencing inflorescence shape [22]. The LAX1 protein, encoded by a helix-loop-helix TF, interacts with *OsMADS1*, *OsMADS6* and *OsMADS7* and plays a role in axillary meristem initiation, panicle length, and spikelet fertility [52]. The interaction between YABBY proteins and tetratricopeptide repeat (TPR) domain-containing proteins significantly influences the number of spikelets per panicle in rice. *OsYAB1* has been implicated in floral organ development, where its ectopic expression leads to the formation of extra stamens and carpels, indicating its role in meristem maintenance and organ development [31]. On the other hand, the TPR domain protein *DES4* is crucial for determining spikelet numbers, as mutations in *DES4* resulted in reduced spikelet numbers and altered expression of other spikelet-related genes [62]. These findings suggest that both *OsYAB1* and *DES4* are essential for proper spikelet development, with *DES4* potentially acting upstream of other genes affecting spikelet number. The interplay between these proteins may regulate the balance between spikelet formation and floral organ development, ultimately affecting rice yield [31,62]. The interaction between FAR1-related proteins and peptidases significantly influences the number of spikelets per panicle in rice. The *THERMO-SENSITIVE BARREN PANICLE* (*TAP*) gene, which encodes a FAR1-related protein regulates panicle and spikelet development, particularly under high-temperature conditions. *TAP* promotes the expression of *OsYABBY* transcription factors, which are essential for spikelet initiation and panicle architecture [94]. Additionally, the manipulation of ligand-receptor pairs, such as the *EPIDERMAL PATTERNING FACTOR* (*EPF*) family interacting with the *OsER1* receptor, has been shown to negatively regulate spikelet number, indicating a complex signaling network that balances spikelet development and fertility [23]. Furthermore, genes like *TAW1* and *APO1* directly increase spikelet numbers, highlighting the genetic basis for enhancing yield. The roles of WD40 proteins and serine/threonine protein phosphatases in determining the number of spikelets per panicle in rice are significant. WD40 proteins, such as *GORI* and *OsKRN2*, regulate pollen tube growth and grain yield, respectively. *GORI* enhances pollen germination and tube elongation, indirectly affecting reproductive success and spikelet formation [37]. Meanwhile, *OsKRN2*, a WD40 protein, negatively regulates grain number by controlling secondary panicle branches; as knocking out this gene led to an increase in grain yield by enhancing spikelet numbers [9]. On the other hand, the *serine/threonine protein phosphatase* found in the qGL3/OsPPKL1 genomic region is crucial for brassinosteroid signaling, which affects grain length and indirectly influences spikelet development. It interacts with WD40 proteins to modulate its stability and localization, impacting growth responses [20]. Together, these findings highlight the intricate interplay between WD40 proteins and protein phosphatases in optimizing spikelet number and overall yield in rice.

5. Conclusion

We have dissected the genetic basis of panicle compactness traits using 158 diverse rice accessions. GWAS and post-GWAS analyses revealed several novel candidate genes associated with panicle traits that affect its compactness. The identified genes need to be functionally characterized and used in future rice breeding programs.

Funding

The research received no funding. Shahid Beheshti University provided the infrastructure for research.

Declaration of Competing Interest

All authors confirm that the article is the author's original work. The authors declare that they do not have any conflict of interest. This manuscript has not been submitted for publication elsewhere. All authors listed have contributed notably, read the manuscript, and agreed to its submission to Plant Science. On behalf of co-authors, the corresponding authors takes full responsibility for the submission.

Acknowledgments

We are grateful for the support of Shahid Beheshti University for providing the infrastructure and Ahwaz Agricultural Research Center for planting the rice accessions.

Author contributions

Masoumeh Kordi conducted the research as her Ph.D. dissertation, curated and analyzed the data, and wrote the original draft. Naser Farrokhi proposed the research idea, supervised the research, and reviewed and edited the manuscript. Asadollah Ahmadikhah supervised the research, helped with data analysis, and edited the manuscript. Abbas Saidi, Mehdi Jahanfar and Pär Ingvarsson reviewed and edited the text.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cpb.2025.100464.

Data availability

Data will be made available on request.

References

[1] H. Adam, A. Gutiérrez, M. Couderc, F. Sabot, F. Ntakirutimana, J. Serret, J. Orjuela, J. Tregear, S. Jouannic, M. Lorieux, Genomic introgressions from African rice (*Oryza glaberrima*) in Asian rice (*O. sativa*) lead to the identification of key QTLs for panicle architecture, BMC Genom. 24 (1) (2023) 587.
[2] Z. Ali, Q. Raza, R.M. Atif, U. Aslam, M. Ajmal, G. Chung, Genetic and molecular control of floral organ identity in cereals, Int. J. Mol. Sci. 20 (11) (2019) 2743.
[3] F. Al-Tam, H. Adam, A. d Anjos, M. Lorieux, P. Larmande, A. Ghesquière, S. Jouannic, H.R. Shahbazkia, P-TRAP: a panicle trait phenotyping tool, BMC Plant Biol. 13 (1) (2013) 1–14.
[4] M. Ashikari, H. Sakakibara, S. Lin, T. Yamamoto, T. Takashi, A. Nishimura, E. R. Angeles, Q. Qian, H. Kitano, M. Matsuoka, Cytokinin oxidase regulates rice grain production, Science 309 (5735) (2005) 741–745.
[5] G.A. Auge, S. Penfield, K. Donohue, Pleiotropy in developmental regulation by flowering-pathway genes: is it an evolutionary constraint? N. Phytol. 224 (1) (2019) 55–70.
[6] V.M. Beretta, E. Franchini, I.U. Din, E. Lacchini, L. Van den Broeck, R. Sozzani, G. Orozco-Arroyo, E. Caporali, H. Adam, S. JJouannic, V. Gregis, M.M. Kater, The

ALOG family members OsG1L1 and OsG1L2 regulate inflorescence branching in rice, Plant J. 115 (2) (2023) 351–368.
[7] J. Cao, K. Liu, W. Song, J. Zhang, Y. Yao, M. Xin, Z. Hu, H. Peng, Z. Ni, Q. Sun, Pleiotropic function of the SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE gene TaSPL14 in wheat plant architecture, Planta 253 (2021) 1–12.
[8] J. Chebib, F. Guillaume, Pleiotropy or linkage? Their relative contributions to the genetic correlation of quantitative traits and detection by multitrait GWA studies, Genetics 219 (4) (2021) iyab159.
[9] W. Chen, L. Chen, X. Zhang, N. Yang, J. Guo, M. Wang, S. Ji, X. Zhao, P. Yin, L. Cai, Convergent selection of a WD40 protein that enhances grain yield in maize and rice, Science 375 (6587) (2022) eabg7985.
[10] S. Cheng, Y. Huang, N. Zhu, Y. Zhao, The rice WUSCHEL-related homeobox genes are involved in reproductive organ development, hormone signaling and abiotic stress response, Gene 549 (2) (2014) 266–274.
[11] S.H. Cho, S.C. Yoo, H. Zhang, D. Pandeya, H.J. Koh, J.Y. Hwang, G.T. Kim, N. C. Paek, The rice narrow leaf2 and narrow leaf3 loci encode WUSCHEL-related homeobox 3 A (OsWOX3 A) and function in leaf, spikelet, tiller and lateral root development, N. Phytol. 198 (4) (2013) 1071–1084.
[12] M.-S. Choi, M.-O. Woo, E.-B. Koh, J. Lee, T.-H. Ham, H.S. Seo, H.-J. Koh, Teosinte Branched 1 modulates tillering in rice plants, Plant Cell Rep. 31 (2012) 57–65.
[13] M.J. Christoffers, Chapter 7 - Weed Population Genetics, in: R.L. Zimdahl (Ed.), Fundamentals of Weed Science (Fifth Edition), Academic Press, 2018, pp. 179–208, https://doi.org/10.1016/B978-0-12-811143-7.00007-X.
[14] K. Das, B.B. Panda, S. Sekhar, E. Kariali, P.K. Mohapatra, B.P. Shaw, Comparative proteomics of the superior and inferior spikelets at the early grain filling stage in rice cultivars contrast for panicle compactness and ethylene evolution, J. Plant Physiol. 202 (2016) 65–74.
[15] S. Das, S.K. Parida, P. Agarwal, A.K. Tyagi, Transcription factor *OsNF-YB9* regulates reproductive growth and development in rice, Planta 250 (2019) 1849–1865.
[16] S. De Nardi, C. Carnevale, S. Raccagni, L. Sangiorgi, Data-driven models to forecast the impact of temperature anomalies on rice production in Southeast Asia, Forecasting 6 (1) (2024) 100–114.
[17] Y. Fan, H. Chen, B. Wang, D. Li, R. Zhou, W. Lian, G. Shao, X. Wei, W. Wu, Q. Liuu, L. Sun, X. Zhan, S. Cheng, Y. Zhang, L. Cao, DWARF AND LESS TILLERS ON CHROMOSOME 3 promotes tillering in rice by sustaining *FLORAL ORGAN NUMBER 1* expression, Plant Physiol. 196 (2) (2024) 1064–1079.
[18] A. Fukushima, H. Ohta, N. Yokogami, N. Tsuda, A. Yoshida, J. Kyozuka, M. Maekawa, Effects of genes increasing the number of spikelets per panicle, *TAW1* and *APO1*, on yield and yield-related traits in rice, Plant Prod. Sci. 20 (4) (2017) 485–489.
[19] X. Gao, W. Liang, C. Yin, S. Ji, H. Wang, X. Su, C. Guo, H. Kong, H. Xue, D. Zhang, The *SEPALLATA-like* gene *OsMADS34* is required for rice inflorescence and spikelet development, Plant Physiol. 153 (2) (2010) 728–740.
[20] S. Gasparis, M.M. Miłoszewski, Genetic basis of grain size and weight in rice, wheat, and barley, Int. J. Mol. Sci. 24 (23) (2023) 16921.
[21] A. Gunasekaran, G. Seshadri, S. Ramasamy, R. Muthurajan, K.S. Karuppasamy, Identification of newer stable genetic sources for high grain number per panicle and understanding the gene action for important panicle traits in rice, Plants 12 (2) (2023) 250.
[22] T. Guo, Z.-Q. Lu, J.-X. Shan, W.-W. Ye, N.-Q. Dong, H.-X. Lin, *ERECTA1* acts upstream of the *OsMKKK10-OsMKK4-OsMPK6* cascade to control spikelet number by regulating cytokinin metabolism in rice, Plant Cell 32 (9) (2020) 2763–2779.
[23] T. Guo, Z.-Q. Lu, Y. Xiong, J.-X. Shan, W.-W. Ye, N.-Q. Dong, Y. Kan, Y.-B. Yang, H.-Y. Zhao, H.-X. Yu, Optimization of rice panicle architecture by specifically suppressing ligand–receptor pairs, Nat. Commun. 14 (1) (2023) 1640.
[24] M. Guo, Y.-H. Yang, M. Liu, Q.-C. Meng, X.-H. Zeng, L.-X. Dong, S.-Z. Tang, M.-H. Gu, C.-J. Yan, Clustered spikelets 4, encoding a putative cytochrome P450 protein *CYP724B1*, is essential for rice panicle development, Chin. Sci. Bull. 59 (2014) 4050–4059.
[25] S. Hao, H. Lou, H. Wang, J. Shi, D. Liu, Baogerile, J. Tao, S. Miao, Q. Pei, L. Yu, M. Wu, M. Gao, N. Zhao, J. Dong, M. You, M. Xin, Genome-wide association study reveals the genetic basis of five quality traits in Chinese wheat, Front. Plant Sci. 13 (2022) 835306.
[26] Y. Heng, C. Wu, Y. Long, S. Luo, J. Ma, J. Chen, J. Liu, H. Zhang, Y. Ren, M. Wang, J. Tan, S. Zhu, J. Wang, C. Lei, X. Zhang, H. Guo, Z. Cheng, J. Wan, OsALMT7 maintains panicle size and grain yield in rice by mediating malate transport, Plant Cell 30 (4) (2018) 889–906.
[27] L. Hu, W. Chen, W. Yang, X. Li, C. Zhang, X. Zhang, L. Zheng, X. Zhu, P. Yin, Y. Wang, B. Ma, S. Li, H. Yuan, B. Tu, *OsSPL9* regulates grain number and grain yield in rice, Front Plant Sci. 12 (2021) 682018.
[28] L. Huang, K. Hua, R. Xu, D. Zeng, R. Wang, G. Dong, G. Zhang, X. Lu, N. Fang, D. Wang, The *LARGE2-APO1/APO2* regulatory module controls panicle size and grain number in rice, Plant Cell 33 (4) (2021) 1212–1228.
[29] L.-y Huang, X.-x Li, Y.-b Zhang, S. Fahad, W. Fei, dep1 improves rice grain yield and nitrogen use efficiency simultaneously by enhancing nitrogen and dry matter translocation, J. Integr. Agric. 21 (11) (2022) 3185–3198.
[30] X. Huang, Q. Qian, Z. Liu, H. Sun, S. He, D. Luo, G. Xia, C. Chu, J. Li, X. Fu, Natural variation at the *DEP1* locus enhances grain yield in rice, Nat. Genet. 41 (4) (2009) 494–497.
[31] S. Jang, J. Hur, S.-J. Kim, M.-J. Han, S.-R. Kim, G. An, Ectopic expression of *OsYAB1* causes extra stamens and carpels in rice, Plant Mol. Biol. 56 (2004) 133–143.
[32] S. Jang, B. Lee, C. Kim, S.-J. Kim, J. Yim, J.-J. Han, S. Lee, S.-R. Kim, G. An, The *OsFOR1* gene encodes a polygalacturonase-inhibiting protein (*PGIP*) that regulates floral organ number in rice, Plant Mol. Biol. 53 (2003) 357–372.
[33] P. Jiang, S. Wang, H. Jiang, B. Cheng, K. Wu, Y. Ding, The COMPASS-Like complex promotes flowering and panicle branching in rice, Plant Physiol. 176 (4) (2018) 2761–2771.
[34] G. Jiang, Y. Xiang, J. Zhao, D. Yin, X. Zhao, L. Zhu, W. Zhai, Regulation of inflorescence branch development in rice through a novel pathway involving the pentatricopeptide repeat protein *sped1-D*, Genetics 197 (4) (2014) 1395–1407.
[35] Y. Jiao, Y. Wang, D. Xue, J. Wang, M. Yan, G. Liu, G. Dong, D. Zeng, Z. Lu, X. Zhu, Regulation of *OsSPL14* by OsmiR156 defines ideal plant architecture in rice, Nat. Genet. 42 (6) (2010) 541–544.
[36] N.R. Kabange, G.D. Dzorkpe, D.-S. Park, Y. Kwon, S.-B. Lee, S.-M. Lee, J.-W. Kang, S.-G. Jang, K.-W. Oh, J.-H. Lee, Rice (*Oryza sativa* L.) Grain Size, Shape, and Weight-Related QTLs identified using GWAS with multiple GAPIT models and high-density SNP chip DNA markers, Plants 12 (23) (2023) 4044.
[37] Y.-J. Kim, K.-H. Jung, WD40-domain protein *GORI* is an integrative scaffold that is required for pollen tube growth in rice, Plant Signal. Behav. 18 (1) (2023) 2082678.
[38] K. Kobayashi, M. Maekawa, A. Miyao, H. Hirochika, J. Kyozuka, PANICLE PHYTOMER2 (*PAP2*), encoding a SEPALLATA subfamily MADS-box protein, positively controls spikelet meristem identity in rice, Plant Cell Physiol. 51 (1) (2010) 47–57.
[39] M. Kordi, N. Farrokhi, A. Ahmadikhah, P.K. Ingvarsson, A. Saidi, M. Jahanfar, Genome-wide association study of rice (*Oryza sativa* L.) inflorescence architecture, Plant Sci. (2025) 112382.
[40] M. Kordi, N. Farrokhi, M.I. Pech-Canul, A. Ahmadikhah, Rice husk at a glance: from agro-industrial to modern applications, Rice Sci. 31 (1) (2024) 14–32.
[41] J. Kyozuka, Grass inflorescence: basic structure and diversity, in: Advances in botanical research, 72, Elsevier, 2014, pp. 191–219.
[42] J. Lan, Q. Lin, C. Zhou, Y. Ren, X. Liu, R. Miao, R. Jing, C. Mou, T. Nguyen, X. Zhu, Small grain and semi-dwarf 3, a *WRKY* transcription factor, negatively regulates plant height and grain size by stabilizing SLR1 expression in rice, Plant Mol. Biol. 104 (2020) 429–450.
[43] N. Li, L. Chen, Y. Li, Control of grain size and number by MAPK signaling in rice, Seed Biol. 2 (1) (2023).
[44] Y. Li, C. Fan, Y. Xing, Y. Jiang, L. Luo, L. Sun, D. Shao, C. Xu, X. Li, J. Xiao, Natural variation in *GS5* plays an important role in regulating grain size and yield in rice, Nat. Genet. 43 (12) (2011) 1266–1269.
[45] Y. Li, X. Li, D. Fu, C. Wu, Panicle Morphology Mutant 1 (*PMM1*) determines the inflorescence architecture of rice by controlling brassinosteroid biosynthesis, BMC Plant Biol. 18 (1) (2018) 1–13.
[46] M. Li, X. Pan, H. Li, Pyramiding of *gn1a*, *gs3*, and *ipa1* exhibits complementary and additive effects on rice yield, Int. J. Mol. Sci. 23 (20) (2022) 12478.
[47] Y. Li, W. Shi, S. Zhu, L. Yu, Y. Cai, Z. Wang, J. Liu, Y. Wu, C. Pan, Z. Chen, Genome-wide association study reveals candidate genes controlling tillering and effective panicle number in rice (*Oryza sativa* L.), Plant Breed. (2024). Doi: z10.1111/pbr.13215.
[48] H. Li, D. Xue, Z. Gao, M. Yan, W. Xu, Z. Xing, D. Huang, Q. Qian, Y. Xue, A putative lipase gene *EXTRA GLUME1* regulates both empty-glume fate and spikelet development in rice, Plant J. 57 (4) (2009) 593–605.
[49] G. Li, H. Zhang, J. Li, Z. Zhang, Z. Li, Genetic control of panicle architecture in rice, Crop J. 9 (3) (2021) 590–597.
[50] X. Liu, M. Huang, B. Fan, E.S. Buckler, Z. Zhang, Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies, PLoS Genet. 12 (2) (2016) e1005767.
[51] J. Liu, Q. Yi, G. Dong, Y. Chen, L. Guo, Z. Gao, L. Zhu, D. Ren, Q. Zhang, Q. Li, Improving rice quality by regulating the heading dates of rice varieties without yield penalties, Plants 13 (16) (2024) 2221.
[52] E. Liu, S. Zhu, M. Du, H. Lyu, S. Zeng, Q. Liu, G. Wu, J. Jiang, X. Dang, Z. Dong, *LAX1*, functioning with MADS-box genes, determines normal palea development in rice, Gene 883 (2023) 147635.
[53] Y. Lu, M. Chuan, H. Wang, R. Chen, T. Tao, Y. Zhou, Y. Xu, P. Li, Y. Yao, C. Xu, Genetic and molecular factors in determining grain number per panicle of rice, Front. Plant Sci. 13 (2022) 964246.
[54] Q. Lu, M. Zhang, X. Niu, S. Wang, Q. Xu, Y. Feng, C. Wang, H. Deng, X. Yuan, H. Yu, Genetic variation and association mapping for 12 agronomic traits in indica rice, BMC Genom. 16 (2015) 1–17.
[55] L. Luo, W. Li, K. Miura, M. Ashikari, J. Kyozuka, Control of tiller growth of rice by *OsSPL14* and strigolactones, which work in two independent pathways, Plant Cell Physiol. 53 (10) (2012) 1793–1801.
[56] J. Luo, H. Liu, T. Zhou, B. Gu, X. Huang, Y. Shangguan, J. Zhu, Y. Li, Y. Zhao, Y. Wang, *An-1* encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice, Plant Cell 25 (9) (2013) 3360–3376.
[57] W. Mai, B. Abliz, X. Xue, Increased number of spikelets per panicle is the main factor in higher yield of transplanted vs. Direct-seeded rice, Agronomy 11 (12) (2021) 2479.
[58] P. Manimaran, S.K. Mangrauthia, R. Sundaram, S. Balachandran, Constitutive expression and silencing of a novel seed specific calcium dependent protein kinase gene in rice reveals its role in grain filling, J. Plant Physiol. 174 (2015) 41–48.
[59] L. Ming, D. Fu, Z. Wu, H. Zhao, X. Xu, T. Xu, X. Xiong, M. Li, Y. Zheng, G. Li, Transcriptome-wide association analyses reveal the impact of regulatory variants on rice panicle architecture and causal gene regulatory networks, Nat. Commun. 14 (1) (2023) 7501.
[60] K. Miura, M. Ikeda, A. Matsubara, X.-J. Song, M. Ito, K. Asano, M. Matsuoka, H. Kitano, M. Ashikari, *OsSPL14* promotes panicle branching and higher grain productivity in rice, Nat. Genet. 42 (6) (2010) 545–549.

[61] S. Moon, S.-R. Kim, G. Zhao, J. Yi, Y. Yoo, P. Jin, S.-W. Lee, K.-h Jung, D. Zhang, G. An, Rice glycosyltransferase1 encodes a glycosyltransferase essential for pollen wall formation, Plant Physiol. 161 (2) (2013) 663–675.

[62] S. Ni, Z. Li, J. Ying, J. Zhang, H. Chen, Decreased Spikelets 4 encoding a novel tetratricopeptide repeat domain-containing protein is involved in DNA repair and spikelet number determination in rice, Genes 10 (3) (2019) 214.

[63] M. Obara, T. Sato, S. Sasaki, K. Kashiba, A. Nagano, I. Nakamura, T. Ebitani, M. Yano, T. Yamaya, Identification and characterization of a QTL on chromosome 2 for cytosolic glutamine synthetase content and panicle number in rice, Theor. Appl. Genet. 110 (2004) 1–11.

[64] R. Panahabadi, A. Ahmadikhah, L.S. McKee, P.K. Ingvarsson, N. Farrokhi, Genome-wide association mapping of mixed linkage (1, 3; 1, 4)-β-glucan and starch contents in rice whole grain, Front. Plant Sci. 12 (2021) 665745.

[65] R. Panahabadi, A. Ahmadikhah, L.S. McKee, P.K. Ingvarsson, N. Farrokhi, Genome-wide association study for lignocellulosic compounds and fermentable sugar in rice straw, Plant Genome (2021), https://doi.org/10.1002/tpg2.20174.

[66] Y. Peng, F. Hou, Q. Bai, P. Xu, Y. Liao, H. Zhang, C. Gu, X. Deng, T. Wu, X. Chen, Rice calcineurin B-like protein-interacting protein kinase 31 (*OsCIPK31*) is involved in the development of panicle apical spikelets, Front. Plant Sci. 9 (2018) 1661.

[67] K. Prasad, S. Parameswaran, U. Vijayraghavan, *OsMADS1*, a rice MADS-box factor, controls differentiation of specific cell types in the lemma and palea and is an early-acting regulator of inner floral organs, Plant J. 43 (6) (2005) 915–928.

[68] P. Qi, Y.-S. Lin, X.-J. Song, J.-B. Shen, W. Huang, J.-X. Shan, M.-Z. Zhu, L. Jiang, J.-P. Gao, H.-X. Lin, The novel quantitative trait locus GL3. 1 controls rice grain size and yield by regulating Cyclin-T1; 3, Cell Res. 22 (12) (2012) 1666–1680.

[69] W. Qian, C. Wu, Y. Fu, G. Hu, Z. He, W. Liu, Novel rice mutants overexpressing the brassinosteroid catabolic gene *CYP734A4*, Plant Mol. Biol. 93 (2017) 197–208.

[70] Y. Qiao, R. Piao, J. Shi, S.-I. Lee, W. Jiang, B.-K. Kim, J. Lee, L. Han, W. Ma, H.-J. Koh, Fine mapping and candidate gene analysis of dense and erect panicle 3, DEP3, which confers high grain yield in rice (*Oryza sativa* L.), Theor. Appl. Genet 122 (7) (2011) 1439–1449.

[71] G. Quero, L. Gutiérrez, E. Monteverde, P. Blanco, F. Perez de Vida, J. Rosas, S. Fernández, S. Garaycochea, S. McCouch, N. Berberian, Genome-wide association study using historical breeding populations discovers genomic regions involved in high-quality rice, Plant Genome 11 (3) (2018) 170076.

[72] D. Ren, Y. Li, G. He, Q. Qian, Multifloret spikelet improves rice yield, N. Phytol. 225 (6) (2020) 2301–2306.

[73] A. Saidi, Z. Hajibarat, Application of next generation sequencing, GWAS, RNA seq, WGRS, for genetic improvement of potato (*Solanum tuberosum* L.) under drought stress, Biocatal. Agric. Biotechnol. 29 (2020) 101801.

[74] S. Sekhar, J. Kumar, S. Mohanty, N. Mohanty, R.S. Panda, S. Das, B.P. Shaw, L. Behera, Identification of novel QTLs for grain fertility and associated traits to decipher poor grain filling of basal spikelets in dense panicle rice, Sci. Rep. 11 (1) (2021) 13617.

[75] N. Sentoku, H. Kato, H. Kitano, R. Imai, *OsMADS22*, an STMADS11-like MADS-box gene of rice, is expressed in non-vegetative tissues and its ectopic expression induces spikelet meristem indeterminacy, Mol. Genet. Genom. 273 (2005) 1–9.

[76] P.K. Subudhi, 2023, Molecular Research in Rice. In (Vol. 24, pp. 10063): MDPI.

[77] W. Tanaka, T. Toriba, Y. Ohmori, A. Yoshida, A. Kawai, T. Mayama-Tsuchida, H. Ichikawa, N. Mitsuda, M. Ohme-Takagi, H.-Y. Hirano, The YABBY gene TONGARI-BOUSHI1 is involved in lateral organ development and maintenance of meristem organization in the rice spikelet, Plant Cell 24 (1) (2012) 80–95.

[78] W. Wang, W. Chen, J. Wang, FRIZZLE PANICLE (FZP) regulates rice spikelets development through modulating cytokinin metabolism, BMC Plant Biol. 23 (1) (2023) 650.

[79] P. Wang, L. Ma, D. Li, B. Zhang, T. Zhou, X. Zhou, Y. Xing, Fine mapping of the panicle length QTL *qPL5* in rice, Mol. Breed. 44 (2) (2024) 6.

[80] F. Wang, K. Yano, S. Nagamatsu, M. Inari-Ikeda, E. Koketsu, K. Hirano, K. Aya, M. Matsuoka, Genome-wide expression quantitative trait locus studies facilitate isolation of causal genes controlling panicle structure, Plant J. 103 (1) (2020) 266–278.

[81] X. Weng, L. Wang, J. Wang, Y. Hu, H. Du, C. Xu, Y. Xing, X. Li, J. Xiao, Q. Zhang, Grain number, plant height, and heading date7 is a central regulator of growth, development, and stress response, Plant Physiol. 164 (2) (2014) 735–747.

[82] H. Xia, Z. Luo, J. Xiong, X. Ma, Q. Lou, H. Wei, J. Qiu, H. Yang, G. Liu, L. Fan, Bi-directional selection in upland rice leads to its adaptive differentiation from lowland rice in drought resistance and productivity, Mol. Plant 12 (2) (2019) 170–184.

[83] J. Xiang, S. Tang, H. Zhi, G. Jia, H. Wang, X. Diao, Loose Panicle1 encoding a novel *WRKY* transcription factor, regulates panicle development, stem elongation, and seed size in foxtail millet [*Setaria italica* (L.) P. Beauv.], PLoS One 12 (6) (2017) e0178730.

[84] J. Xie, F. Li, N.U. Khan, X. Zhu, X. Wang, Z. Zhang, X. Ma, Y. Zhao, Q. Zhang, S. Zhang, Identifying natural genotypes of grain number per panicle in rice (*Oryza sativa* L.) by association mapping, Genes Genom. 41 (2019) 283–295.

[85] Y. Xing, W. Tang, W. Xue, C. Xu, Q. Zhang, Fine mapping of a major quantitative trait loci, *qSSP7*, controlling the number of spikelets per panicle as a single Mendelian factor in rice, Theor. Appl. Genet. 116 (2008) 789–796.

[86] Q. Xu, H. Yu, S. Xia, Y. Cui, X. Yu, H. Liu, D. Zeng, J. Hu, Q. Zhang, Z. Gao, The C2H2 zinc-finger protein LACKING RUDIMENTARY GLUME 1 regulates spikelet development in rice, Sci. Bull. 65 (9) (2020) 753–764.

[87] G. Yang, Q. Wang, G. Yang, G. Zhang, H. Chen, X. Wang, P. Ma, Y. Hu, Application of open panicle traits in improving the filling characteristics at the base of Indica rice panicles, Plants 13 (15) (2024).

[88] K. Yano, T. Ookawa, K. Aya, Y. Ochiai, T. Hirasawa, T. Ebitani, T. Takarada, M. Yano, T. Yamamoto, S. Fukuoka, Isolation of a novel lodging resistance QTL gene involved in strigolactone signaling and its pyramiding with a QTL gene involved in another mechanism, Mol. Plant 8 (2) (2015) 303–314.

[89] A. Yoshida, M. Sasao, N. Yasuno, K. Takagi, Y. Daimon, R. Chen, R. Yamazaki, H. Tokunaga, Y. Kitaguchi, Y. Sato, *TAWAWA1*, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition, Proc. Natl. Acad. Sci. 110 (2) (2013) 767–772.

[90] Z. Yuan, S. Persson, D. Zhang, Molecular and genetic pathways for optimizing spikelet development and grain yield, aBiotech 1 (4) (2020) 276–292.

[91] F. Zhang, E. Boerwinkle, M. Xiong, Epistasis analysis for quantitative traits by functional regression model, Genome Res. 24 (6) (2014) 989–998.

[92] S. Zhang, L. Deng, L. Zhao, C. Wu, Genome-wide binding analysis of transcription factor Rice Indeterminate 1 reveals a complex network controlling rice floral transition, J. Integr. Plant Biol. 64 (9) (2022) 1690–1705.

[93] D. Zhang, M. Zhang, Y. Zhou, Y. Wang, J. Shen, H. Chen, L. Zhang, B. Lü, G. Liang, J. Liang, The rice G protein γ subunit *DEP1/qPE9–1* positively regulates grain-filling process by increasing auxin and cytokinin content in rice grains, Rice 12 (2019) 1–11.

[94] P. Zhang, W. Zhu, Y. He, J. Fan, J. Shi, R. Fu, J. Hu, L. Li, D. Zhang, W. Liang, THERMOSENSITIVE BARREN PANICLE (TAP) is required for rice panicle and spikelet development at high ambient temperature, New Phytol. 237 (3) (2023) 855–869.

[95] L. Zhao, H. Liu, K. Peng, X. Huang, Cold-upregulated glycosyltransferase gene 1 (*OsCUGT1*) plays important roles in rice height and spikelet fertility, J. Plant Res. 136 (3) (2023) 383–396.

[96] J. Zhao, M.B. Mantilla Perez, J. Hu, M.G. Salas Fernandez, Genome-wide association study for nine plant architecture traits in Sorghum, Plant Genome 9 (2) (2016), 2015.2006.0044.

[97] H.-Y. Zhao, J.-X. Shan, W.-w Ye, N.-Q. Dong, Y. Kan, Y.-B. Yang, H.-X. Yu, Z.-Q. Lu, S.-Q. Guo, J.-J. Lei, B. Liao, H.-X. Lin, A QTL GN1.1, encoding FT-L1, regulates grain number and yield by modulating polar auxin transport in rice, J. Integr. Plant Biol. 66 (10) (2024) 2158–2174.

[98] L. Zhao, L. Tan, Z. Zhu, L. Xiao, D. Xie, C. Sun, *PAY 1* improves plant architecture and enhances grain yield in rice, Plant J. 83 (3) (2015) 528–536.

[99] K. Zhao, C.-W. Tung, G.C. Eizenga, M.H. Wright, M.L. Ali, A.H. Price, G.J. Norton, M.R. Islam, A. Reynolds, J. Mezey, Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa, Nat. Commun. 2 (1) (2011) 467.

[100] H. Zheng, S. Sun, L. Bai, S. Jiang, G. Ding, T. Wang, H. Zhao, J. Wang, H. Liu, L. Yang, Identification of candidate genes for panicle length in *Oryza sativa* L. ssp. japonica via genome-wide association study and linkage mapping, Euphytica 218 (2) (2022) 16.

[101] H. Zhong, S. Liu, X. Meng, T. Sun, Y. Deng, W. Kong, Z. Peng, Y.J.B. LiUncovering the genetic mechanisms regulating panicle architecture in rice with GPWAS and GWAS. 22, 1-13.2021, Uncoverin Genet. Mech. Regul. panicle Archit. rice GPWAS GWAS113.

[102] Y. Zhou, J. Zhu, Z. Li, C. Yi, J. Liu, H. Zhang, S. Tang, M. Gu, G. Liang, Deletion in a quantitative trait gene *qPE9-1* associated with panicle erectness improves plant architecture during rice domestication, Genetics 183 (1) (2009) 315–324.

[103] H. Zhuang, H.-L. Wang, T. Zhang, X.-Q. Zeng, H. Chen, Z.-W. Wang, J. Zhang, H. Zheng, J. Tang, Y.-H. Ling, *NONSTOP GLUMES1* encodes a C2H2 zinc finger protein that regulates spikelet development in rice, Plant Cell 32 (2) (2020) 392–413.