



## OPEN ACCESS

## EDITED BY

Rohini Garg,  
Shiv Nadar University, India

## REVIEWED BY

Nkulu Kabange Rolly,  
Kyungpook National University,  
Republic of Korea  
Xueqiang Wang,  
Zhejiang University, China

## \*CORRESPONDENCE

Naser Farrokhi

✉ n\_farrokhi@sbu.ac.ir

Asadollah Ahmadikhah

✉ ahmadikhaha@gmail.com

Pär K. Ingvarsson

✉ par.ingvarsson@slu.se

RECEIVED 08 May 2024

ACCEPTED 03 September 2024

PUBLISHED 08 October 2024

## CITATION

Kazemzadeh S, Farrokhi N, Ahmadikhah A,  
Tabar Heydar K, Gilani A, Askari H and  
Ingvarsson PK (2024) Genome-wide  
association study and genotypic  
variation for the major tocopherol  
content in rice grain.  
*Front. Plant Sci.* 15:1426321.  
doi: 10.3389/fpls.2024.1426321

## COPYRIGHT

© 2024 Kazemzadeh, Farrokhi, Ahmadikhah,  
Tabar Heydar, Gilani, Askari and Ingvarsson.  
This is an open-access article distributed under  
the terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Genome-wide association study and genotypic variation for the major tocopherol content in rice grain

Sara Kazemzadeh<sup>1</sup>, Naser Farrokhi <sup>1\*</sup>,  
Asadollah Ahmadikhah <sup>1\*</sup>, Kourosh Tabar Heydar<sup>2</sup>,  
Abdolali Gilani<sup>3</sup>, Hossein Askari<sup>1</sup> and Pär K. Ingvarsson <sup>4\*</sup>

<sup>1</sup>Department of Cell and Molecular Biology, Faculty of Life Sciences & Biotechnology, Shahid Beheshti University, Tehran, Iran, <sup>2</sup>Chemistry and Chemical Engineering Research Center of Iran, Tehran, Iran, <sup>3</sup>Agricultural and Natural Resources Research Institute of Khuzestan, Ahwaz, Iran, <sup>4</sup>Department of Plant Biology, Swedish University of Agricultural Sciences, Uppsala, Sweden

Rice tocopherols, vitamin E compounds with antioxidant activity, play essential roles in human health. Even though the key genes involved in vitamin E biosynthetic pathways have been identified in plants, the genetic architecture of vitamin E content in rice grain remains unclear. A genome-wide association study (GWAS) on 179 genotypically diverse rice accessions with 34,323 SNP markers was conducted to detect QTLs that define total and  $\alpha$ -tocopherol contents in rice grains. Total and  $\alpha$ -tocopherol contents had a strong positive correlation and varied greatly across the accessions, ranging from 0.230–31.76 and 0.011–30.83 ( $\mu\text{g/g}$ ), respectively. A total of 13 QTLs were identified, which were spread across five of the rice chromosomes. Among the 13 QTLs, 11 were considered major with phenotypic variation explained (PVE) greater than 10%. Twelve transcription factor (TF) genes, one microprotein (miP), and a transposon were found to be associated with the QTLs with putative roles in controlling tocopherol contents. Moreover, intracellular transport proteins, ABC transporters, nonspanins, and SNARE, were identified as associated genes on chromosomes 1 and 8. In the vicinity of seven QTLs, protein kinases were identified as key signaling factors. Haplotype analysis revealed the QTLs *qAlph1.1*, *qTot1.1*, *qAlph2.1*, *qAlph6.1*, *qTot6.1*, and *qTot8.3* to have significant haplogroups. Quantitative RT-PCR validated the expression direction and magnitude of *WRKY39* (*Os02g0265200*), *PIP5Ks* (*Os08g0450800*), and *MADS59* (*Os06g0347700*) in defining the major tocopherol contents. This study provides insights for ongoing biofortification efforts to breed and/or engineer vitamin E and antioxidant levels in rice and other cereals.

## KEYWORDS

kinases, microProtein, quantitative trait loci (QTLs), transcription factor, transporter

## Introduction

The nutritional quality of rice is important for the intake of essential nutrients, *esp.* in countries where rice is considered a staple (Wang et al., 2018). Rice grain contains a wide range of bioactive compounds such as vitamin E (mostly found in bran),  $\gamma$ -oryzanol, phenolic acids, flavonoids, and phytic acid (Shammugasamy et al., 2015). Vitamin E, also known as tocopherol, an essential nutrient in humans, is being synthesized by some cyanobacteria and all plant tissues, including leaves and grains (Sattler et al., 2004; Mène-Saffrané et al., 2010). The vitamin E biosynthetic pathway (Figure 1A) is highly conserved in the plant kingdom (DellaPenna and Mène-Saffrané, 2011). Common features of all tocopherols are a hydrophobic polyprenyl side chain with anchoring properties into bio-membranes and an antioxidant chromanol ring derived from homogentisate (Schuy et al., 2019). According to the various types of side chains, vitamin E can be defined as tocopherol ( $\alpha$ -,  $\beta$ -,  $\gamma$ -, and  $\delta$ ), tocotrienol, plastochromanol-8 (PC-8), and tocomonoenol. All four tocopherol isoprenoid side chains are being produced by GGDP synthases (Muñoz and Munne-Bosch, 2019; Figures 1B, C). Among the four types of tocopherols,  $\alpha$ -tocopherol shows the highest antioxidant activity and is preferentially recognized by human cells (Grusak and DellaPenna, 1999; Havaux et al., 2005; Traber and Atkinson, 2007).

As a lipid-soluble antioxidant, vitamin E scavenges free radicals (Zingg, 2007). Vitamin E deficiency exacerbates the onset of many diseases, such as cancers, Alzheimer's disease, and cardiovascular ailments (Sen et al., 2007; Falk and Munné-Bosch, 2010; Sozen et al., 2019). Vitamin E is essential in human diets and is also required for plant-environment adaptations (Niu et al., 2022). Vitamin E is most abundant in seeds, where they are essential for protecting membrane lipids, especially during seed desiccation, storage, and germination (Albert et al., 2022), effectively scavenging lipid peroxy radicals, thus prolonging seed longevity (Lee et al., 2020). Elevated vitamin E levels in livestock feed results in meat with longer storage life, more desirable appearance, and greater consumer preference, presumably by inhibiting lipid peroxidation (Lipka et al., 2013). The probable role of vitamin E as a signaling molecule in plant cells has been recently reported (Ali et al., 2022).

Tocopherol biosynthesis has been most extensively investigated in *Arabidopsis thaliana*, where genes for the synthesis of homogentisic acid (*HGA*), phytyl diphosphate (*PDP*), and the core vitamin E pathway itself (*VTE1* through *VTE6*) were initially identified. GWAS of seed tocopherols in 814 *Arabidopsis* lines (part of the 1001 *Arabidopsis* Genome Panel, Alonso-Blanco et al., 2016) proposed the involvement of a novel seed-specific  $\alpha/\beta$  hydrolase (*AtVTE7*). *AtVTE7* is targeted to the chloroplast envelope and accounts for 55% of the total seed tocopherols. Consistent with the results in *Arabidopsis*, the maize orthologous gene *ZmVTE7* controls 38% and 49% of total tocopherols in kernel and leaf, respectively (Albert et al., 2022). Of note is that *AtVTE7* was only detected in seeds with no control over the tocopherol content of the leaf. Several chromosomal loci, identified via genome-wide association studies (GWAS), have been reported to be associated with natural variation in the content and composition of vitamin E in different grains. For instance, Haddadi et al. (2012) identified

*Vte4*, *HPPD*, *GST*, and *Droug1* to be related to tocopherol content in sunflower. Lipka et al. (2013) found an SNP (#591822) in maize grain to be significantly associated with *ZmVTE4* and *ZmVTE1*. Others have reported strong associations between *Vte4* and  $\alpha$ -tocopherol content in maize grain (Lipka et al., 2013; Wang et al., 2018; Baseggio et al., 2019). Graebner et al. (2015) identified 13 SNPs associated with vitamin E biosynthesis in barley. In another study, two genes *homogentisate phytyltransferase* (*HPT-7H*) and *homogentisate geranylgeranyltransferase* (*HGGT*), were identified on barley chromosome 7H. These genes code for key enzymes controlling the accumulation of tocopherols in leaves and tocotrienols in grains, respectively (Schuy et al., 2019; Figure 1A). Wang et al. (2015) performed GWAS using 1.44 million high-quality SNPs acquired from the re-sequencing of 137 accessions of a diverse rice core collection. This led to the identification of 13 candidate genes, of which  *$\gamma$ -tocopherol methyltransferase* (*O $\gamma$ TMT*) was identified as the major gene controlling  $\alpha$ -tocopherol content.

Although much information about tocopherol and biosynthesizing genes has accumulated, the knowledge concerning the molecular mechanisms of transport, turnover, signaling, and regulation of tocopherol biosynthesis remains fragmentary. Despite the impact of some transcription factors (TFs) in the methylerythritol phosphate (MEP) and shikimate (SK) pathways (Figure 1A), the regulatory mechanism in the tocopherol pathway remains unclear. Synthetic zinc finger transcription factors (ZFP-TFs) were designated to upregulate the expression of the endogenous *Arabidopsis*  *$\gamma$ -tocopherol methyltransferase* gene. This gene encodes the enzyme responsible for the conversion of  $\gamma$ -tocopherol to  $\alpha$ -tocopherol. Overexpression of ZFP-TF under the control of the seed-specific promoter caused a 20-fold increase in  $\alpha$ -tocopherol compared to control in *Arabidopsis* seeds (Van Eenennaam et al., 2004). In tomato, the overexpression of a MYB, *ODORANT1*, led to a significant increase in phenylpropanoid content concomitantly with transcriptional upregulation of several SK pathway genes (Dal Cin et al., 2011). Downregulation of *DET1*, which is involved in suppressing light responses in the absence of light in tomato fruit, resulted in increased contents of carotenoids, flavonoids, and tocopherol (Enfissi et al., 2010).

In addition to TFs, microproteins (miPs) are important regulatory factors consisting of small (5-20 kilodaltons) proteins known to regulate gene expression at the post-translational level through protein-protein interactions (Straub and Wenkel, 2017). The first miP family reported in plants was the LITTLE ZIPPER (ZPR) family of proteins (ZPR1-4). ZPR has a leucine zipper domain that interacts and post-translationally regulates other leucine zipper-containing proteins to modulate plant development (Wenkel et al., 2007). As newly discovered genes, very little is currently known about the structure and function of microproteins, making these genes interesting targets for discovering new biological mechanisms (Rathore et al., 2018). Transposable elements (TEs) are other genomic elements that have also been described as controlling elements, due to their ability to influence gene expression reversibly (Pachamuthu and Borges, 2023). Only a few TEs with regulatory functions have been discovered in rice to date, and their activities are mostly induced by artificial treatments (*e.g.*, tissue culture and hybridization). *Tos17*, belonging to the

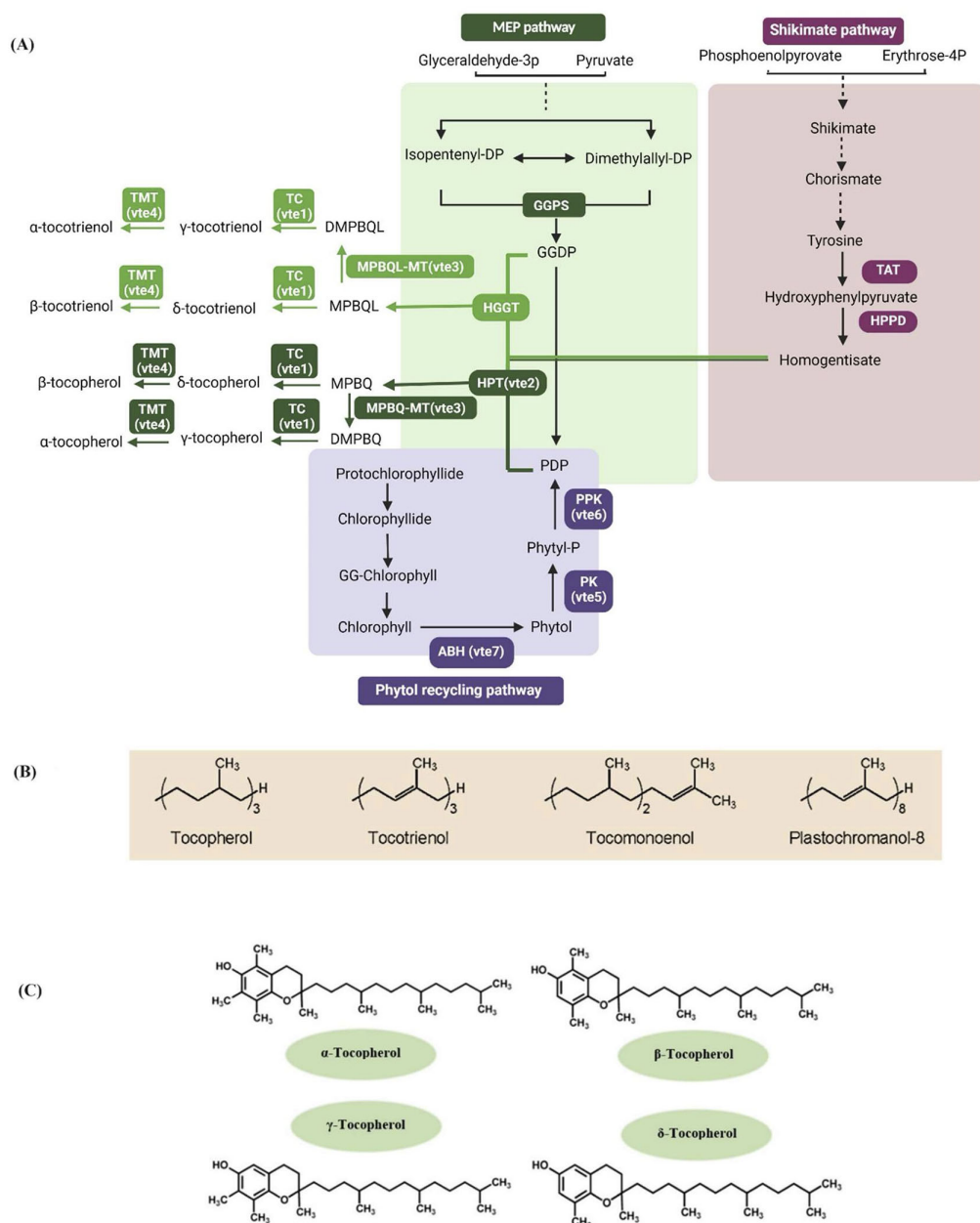


FIGURE 1

(A) Biosynthetic pathway of tocopherols and tocotrienols are formed from the combination of MEP (methylerythritol phosphate) and shikimate pathways. The shikimate pathway gives rise to a chromanol ring from HGA (homogentisate). MEP pathway provides the prenol tail from GGDP (geranylgeranyl diphosphate) and PDP (phytyl diphosphate) for the synthesis of tocotrienol and tocopherol, respectively (Muñoz and Munne-Bosch, 2019). VTE2 (a homogentisate phytyltransferase) condenses PDP and HGA to produce MPBQ (methylphytyl benzoquinol) (Sattler et al., 2004). In monocot lineage, HGA can also be condensed with GGDP by HGGT1 (homogentisate geranylgeranyl transferase) to generate MGGBQ (methyl geranylgeranyl benzoquinol), the committed step for tocotrienol synthesis. MPBQ and MGGBQ are substrates for a series of methylations by MPBQ/ MGGBQ methyltransferase (VTE3) and  $\gamma$ -tocopherol methyltransferase (VTE4) and cyclization by tocopherol cyclase (VTE1) whose sequence and numbers of reactions generate the  $\alpha$ ,  $\beta$ ,  $\delta$ , and  $\gamma$  isoforms of tocopherols and tocotrienols (Muñoz and Munne-Bosch, 2019). At the same time, GGDP for tocotrienol synthesis comes directly from the isoprenoid pathway. The generation of PDP for tocopherol synthesis is more complex and still not completely resolved. Differences in leaf tocopherol synthesis exist between monocots and dicots; tocopherol synthesized in seeds requires chlorophyll biosynthesis (Diepenbrock et al., 2017; Zhan et al., 2019), and the activity of VTE7, an alpha/beta hydrolase that interfaces with chlorophyll synthesis to release phytol (Albert et al., 2022). Phytol is then sequentially phosphorylated to PDP through the action of phytol kinase (VTE5) and phytol phosphate kinase (VTE6) (Valentin et al., 2006; Vom Dorp et al., 2015) (B) Tocopherol contains fully saturated aliphatic side chain, while the side chain of tocotrienol contains three extra trans double bonds. PC-8 has a similar unsaturated as tocotrienol but a longer side chain, whereas tocomonoenol has only one double bond on its side chain (Tian et al., 2007; Sadre et al., 2010; Stacey et al., 2016). (C) Tocopherol chemical structures feature a chromanol head and a prenol tail. DMGGBQ, dimethylgeranylgeranylbenzoquinol; DMPBQ, dimethylphytylbenzoquinol; GGDP, geranylgeranyl diphosphate reductase; GGPS, geranylgeranyl diphosphate synthase; HPPD, hydroxyphenylpyruvate dioxygenase; HPT, homogentisate phytyl transferase; MPBQ-MT, MPBQ methyltransferase; TAT, tyrosine aminotransferase; TC, tocopherol cyclase; TMT, tocopherol methyltransferase; ABH, alpha/beta hydrolase; PK, Phytol kinase; PPK, Phytol phosphate kinase. The figure was created at BioRender.com.

Copia superfamily, is the first active retrotransposon isolated from rice (Hirochika et al., 1996), which exhibits a high level of transposition activity under tissue culture and has been used to induce rice mutants for functional genomics studies (Hirochika, 2001). Although tocopherols have been mostly related to photosynthetic tissues, the accumulation of tocopherols occurs in several non-photosynthetic tissues. The recent discovery of an  $\alpha$ -tocopherol-binding protein (TBP, Muñoz and Munne-Bosch, 2019) has altered our view on tocopherol transport in plants, showing intraorganellar vitamin E transport within chloroplasts. TBPs may facilitate the rapid exchange of tocopherols between intraorganellar membranes in chloroplasts (Muñoz and Munne-Bosch, 2019).

Here, we investigated the regulatory genes (TFs, miPs, and transposons) and genes involved in the transport and signaling of tocopherol biosynthesis using a natural rice population. Rice is a model plant with great resources of genetic information. Collectively, our findings provide genetic information for increasing/modulating tocopherol content in rice, other cereals and crops through molecular breeding strategies. Moreover, the candidate genes in this study can be used in the metabolic engineering of microorganisms for the industrial production of tocopherols which currently are synthesized chemically for pharma.

## Materials and methods

### Plant material

A collection of 282 accessions of rice (*Oryza sativa*) from 82 countries was obtained from the International Rice Research Institute (IRRI), Philippines (Supplementary Table S1), in addition to three local cultivars included as check varieties (Sadri, Sang Tarom, Dom Siah Kalat). The experiment was conducted at the Agricultural Research Station, Shavoor, Ahwaz, Iran (48°27' E, 31°50' N) during the 2021-2022 cultivation season. The genotypes were planted in an augmented block design (Federer and Raghavarao, 1975) with three check varieties in every block (3 replicates). The land was fertilized using superphosphate triple (150 kg/ha) and potash (150 kg/ha) at the plowing stage. All seeds were harvested at maturity and stored in dry conditions at 4°C. A total of 179 accessions, belonging to *indica* (42), *japonica* (79), *aus* (21), aromatic (4), and admix (33), in addition to local cultivars, completed the growing season and used in our subsequent studies.

### SNP genotyping data

The development and sequencing of an SNP hybridization array for the rice population have previously been described by Zhao et al. (2011). Data for the rice 44.1 K SNPs array was downloaded from the Gramene portal (<http://gramene.org>) for all the rice accessions included in the study. SNP loci were further filtered to produce a high-quality set with minimum allele frequencies (MAF) greater than 0.05 in TASSEL v5.0 (Bradbury et al., 2007). After excluding low-quality and monomorphic loci, 34,323 SNP loci were retained.

## Analysis of rice grain tocopherol content

Rice grains (2 g) were crushed for 1 min and vortex-mixed in *n*-hexane (3:1 solvent: powder) for 1 min (Hu et al., 1996). The mixture was wrapped in aluminium foil to reduce light exposure and shaken in a rotary shaker at 220 rpm, 26°C for 16 h. The next day, samples were centrifuged at 8000  $\times$ g for 10 min at 4°C. Supernatant was poured into a fresh tube and placed at 22°C for solvent evaporation. The remaining oil (500  $\mu$ l) was passed through a PTFE filter (0.45  $\mu$ m) and mixed with acetonitrile in a ratio of 1:1.

Tocopherol contents were determined by high-performance liquid chromatography (LC pump K-1001, KNAUER, Germany) coupled with a fluorescence detector (Shimadzu RF - 10AxL, Kyoto, Japan; excitation of 295 nm and an emission of 330 nm). The mobile phase was 100% acetonitrile at a flow rate of 1 ml.min<sup>-1</sup> at a column temperature of 40°C. The oil injection volume was 20  $\mu$ l. A C18 reverse phase column (GALAK, 150 mm length, 5  $\mu$ m particle sizes, 100  $\text{Å}$ , 330 m<sup>2</sup>. g<sup>-1</sup>) was used to identify tocopherols. Amounts of tocopherol were calculated based on the established standard curve under the same HPLC conditions.

### GC- MS analysis

The fatty acid profile of grains of selected rice accessions was determined using a Gas Chromatogram (Agilent 6890 series, Palo Alto, CA, USA) attached to a mass spectrometer (Agilent 5973 Network mass selective detector, Delaware, USA). The FAMES (*fatty acid methyl esters*) of the oils were separated using a CP-SIL 8CB-MS FS 50  $\times$  2.25 column. The column oven temperature was initially maintained at 60°C for 5 min, increased at 5°C.min<sup>-1</sup> to 280°C, and kept at 280°C for 15 min. The carrier gas was helium with a flow rate of 1.5 ml.min<sup>-1</sup>. Rice grain oil (0.5 ml) was dissolved in 1 ml hexane, reaching 5 ml with 2 M methanolic KOH. The mixture was stirred for 1 h at 22°C and centrifuged at 4,000 rpm for 10 min. The upper layer of *n*-hexane containing FAME was pipetted from the solution, placed in a 2 ml glass vial, and stored below -4°C until GC-MS. The composition was reported as a relative percentage of the total peak area. The FAMES and essential oil constituents were identified by matching their mass spectra with the peaks in the library (NIST) database (Hagr and Adam, 2020).

### Population structure analysis

All high-quality SNPs were sampled to calculate population structure (Q) and kinship (K). Q and K matrices were used as covariates in the models for trait-marker association analysis. Principal component analysis (PCA) was conducted to summarize the genetic structure and variation present in the collection. The analyses to obtain principal components matrices of the accessions were performed using rMVP in R (Yin et al., 2021). Kinship analysis was also performed using 34,323 enumerated SNP data markers from 179 genotypes. The input file was prepared using TASSEL v5.0, and the kinship matrix was obtained using the



Popkin package in R. A heat map of kinship relationships was generated using the gplots package in R.

The normality of frequency distribution for the phenotypic data was tested using skewness and kurtosis tests in R. The ANOVA was carried out using the Augmented Block Design using R agricolae package 1.4.0. The broad-sense heritability ( $H^2b$ ) and narrow-sense heritability ( $H^2n$ ) of the alpha and total tocopherol were estimated using the heritability package in R (<https://cran.r-project.org/web/packages/heritability/index.html>) by the following equation:  $H^2b = V_G/V_P$ ,  $H^2n = V_A/V_P$  where  $V_G$  is the genotypic variance,  $V_P$  is the phenotypic variance, and  $V_A$  is the additive variance.

## Linkage disequilibrium estimation

The filtered SNP data was used to calculate linkage disequilibrium (LD) between SNPs using  $r^2$  in a sliding window of 50 markers using TASSEL v5.0 (Bradbury et al., 2007). Graphs depicting the decay of LD with the physical distance between SNPs were created by aod, dplyr, and stringr and visualized using ggplot2 in R. The distance across the chromosome where  $r^2$  dropped to half its maximum value was called the LD decay distance (Huang et al., 2010).

## GWAS analysis

Genotypic data such as single nucleotide polymorphisms (SNPs), phenotypic data, and covariates (Q and K matrices) were collected to perform GWAS using rMVP (Yin et al., 2021). To ensure accuracy and to minimize the false-positive rate, two models were utilized independently: i) Mixed linear model (MLM; Price et al., 2006) and ii) Fixed and Random Model Circulating Probability Unification (FarmCPU; Liu et al., 2016). MLMs are models that account for both fixed and random effects in a dataset, while FarmCPU is a method that employs MLMs to identify genetic associations between traits and multiple markers in genomic data. This model uses a compressed MLM to improve computational efficiency and allows for the analysis of large datasets. Manhattan plots with suggested threshold lines were produced using rMVP, considering the very low LD decay distance in rice (Mather et al., 2007). The QTL span was defined as the location of significantly associated SNPs  $\pm$  LD decay distance. Significant QTLs with  $-\log_{10}(p) \geq 4$  for  $\alpha$ T and total tocopherol contents were determined based on Q-Q plot (Galwey, 2023). To reduce QTL redundancy, if overlaps were observed between two or more QTL flanking areas, they were combined into one QTL (Chen et al., 2014; Wang et al., 2015; Guo et al., 2018). Genes annotated as either involved in vitamin E biosynthesis or the closest gene near each QTL peak were proposed to be the most likely candidate genes.

## Bioinformatics and gene identification

To identify genes underlying the QTLs of vitamin E content that overlapped with the genomic regions (*i.e.*, their associated SNPs), the genes were checked at IRGSP-1.0 (<https://rapdb.dna.affrc.go.jp/>).

The flanking region of each SNP marker was chosen based on the distance of LD decay (Mather et al., 2007), and checked for map order uncertainty and LD. To ascertain whether candidate genes underlying the detected QTLs, which were previously cloned and functionally annotated, a thorough literature search was done for each gene. The online Plantpan3 tool (<http://plantpan.itps.ncku.edu.tw/>, Chen et al., 2019) was used to check each corresponding promoter and co-expression analysis of the main genes (Mène-Saffrané, 2017) in tocopherol biosynthesis pathway. The STRING tool (Search Tool for the Retrieval of Interacting Genes - <https://string-db.org/>) was employed to construct the protein association networks between candidate gene and main genes involved in tocopherol biosynthesis, with a confidence parameter set to a threshold of 0.40 and a maximum number of interactors in the first and second shells. The Cytoscape v3.9.1 software was used for network visualization. The miPfinder algorithm was used to identify the microproteins among the genes near detected QTLs. The effect of each associated SNP was annotated using SnpEff: rice genome annotation information from the RAP database (RAP-DB, <https://rapdb.dna.affrc.go.jp/>) was used in the SnpEff analysis. MapMan software (ver. 3.6.0RC1) was used to map transcriptomic data. Osa\_RAPDB\_mapping files were downloaded from the Map-Man store server (<http://mapman.gabipd.org/web/guest/mapmanstore>).

## Haplotyping

Haplotype analyses were performed, and the outputs were visualized using geneHapR (Zhang et al., 2023). The phenotypic value of each haplotype was assessed by calculating the average phenotypic value across the germplasm with each type of SNP locus linked to a specific target trait. In this study, favorable haplogroups were defined as the haplotype that showed the highest average values over other haplotypes. Haplotypes with heterozygous or missing data were removed.

## Epistatic interaction (QQ) analysis

Epistatic interactions between significant QTLs obtained via GWAS ( $-\log_{10}(p) > 4.0$  in FarmCPU model) for total and  $\alpha$ -tocopherol were analyzed. For QQ analysis, the snpPairInteraction function in the FRGEpistasis of R package (Zhang et al., 2014) was used and determined to be significant at a  $P$ -value  $\leq 0.05$ . The Circlize package was used to visualize the significant epistatic interactions (Gu et al., 2014).

## Tissue-specific expression of candidate genes

The expression profiles of the candidate genes in various tissues were obtained from the Rice Genome Annotation Project database (RGAP, <http://rice.plantbiology.msu.edu/>), including shoots (library name in NCBI SRA: SRR042529),

leaves-20 days (SRX100741), seed-5 DAP (days after pollination; SRX100749) and seed-10 DAP (SRX100755). A heatmap was generated to visualize the gene expression patterns across the different tissues. Functional enrichment analyses of candidate genes were depicted using MapMan 3.0.0 software based on the functional annotated file *osa\_RAPDB\_mapping* (<https://MapMan.gabipd.org/MapManstore>).

## Quantitative reverse transcription polymerase chain reaction

Two selected genotypes, #g95 (low tocopherol) and #g217 (high tocopherol), were grown in the spring of 2023, and fertilized three times with complete fertilizer. The genotypes were sampled at milky (7 days post pollination (DPP)), doughy (14 DPP), and mature developing seed (28 DPP) stages. Sampled seeds were wrapped in foil and placed in liquid nitrogen. Total RNA was isolated from seeds in different developmental stages using GenUP™ Total RNA Mini Kit (biotechrabbit GmbH, Berlin, Germany). First-strand cDNA was synthesized from total RNA using Thermo Scientific RevertAid RT Kit (#K1691, Thermo Fisher Scientific). qRT-PCR was performed using SYBR Green (FP205, Tiangen) reagents on a Real-Time PCR machine (T100 Thermal Cycler, Bio-Rad). All qRT-PCRs were performed in three independent replicates, and the relative expression levels were calculated using the  $2^{-\Delta\Delta Ct}$  method (Livak and Schmittgen, 2001). Statistical significance of Ct differences between low (#g95) and high (#g217) tocopherol genotypes was calculated with a two-way ANOVA, employing Tukey's post-test comparisons, using the GraphPad Prism 9.5.1 software.

The *OsActin* (*Os03g0718100*) was used as an internal control for expression normalization. Primer pairs used for the rice cDNA amplification were designed using Primer3Plus (<https://www.primer3plus.com>, Supplementary Table S8).

## Gene ontology enrichment and KEGG pathway determination

ShinyGO v0.75 (<http://bioinformatics.sdstate.edu/go/>) served as a web-based tool for exploring GO term enrichment and KEGG analysis for candidate genes (Ge et al., 2020). The search was conducted against the *Oryza sativa japonica* group.

## Results

### Phenotypic variation

In general, alpha and total tocopherol phenotypes were skewed towards lower values for the panel of 179 rice accessions (Supplementary Figure S1). Alpha-tocopherol content was significantly correlated with total tocopherol content ( $r^2 = 0.97$ ). The correlation was consistent with the fact that alpha is the major tocopherol in rice grain and contributes greatly to the total tocopherol content in rice (Supplementary Table S2). The total

tocopherol content in the accessions varied widely from 0.230 to 31.76 ( $\mu\text{g/g}$ ), with an average value of 3.70 ( $\mu\text{g/g}$ ) and a coefficient of variation (C.V.) of 6.7%. alpha-tocopherol contents in the seeds of the 179 accessions varied from 0.011 to 30.83 ( $\mu\text{g/g}$ ), with a mean value of 3.30 ( $\mu\text{g/g}$ ) and a C.V. of 7.3%. The alpha/total tocopherol ratio of accessions in each rice subpopulation was calculated. The *indica* genotypes had a higher alpha/total tocopherol ratio than other accessions. Among the 179 rice accessions, the top five accessions with the highest alpha and total tocopherol contents were g253, g252, g120, g217, and g134. Estimates of  $H^2b$  and  $H^2n$  were 0.86 and 0.76 for alpha-tocopherol and 0.87 and 0.79 for total tocopherol content, respectively. Estimated genotypic variance ( $V_G$ ) and additive variance ( $V_A$ ) accounted for the alpha-tocopherol were 29.75 and 26.64, and for total tocopherol were 30.71 and 27.88, respectively (Table 1). Relatively high heritability values suggest that phenotypic-based selection for tocopherol content can be used effectively. These results showed that tocopherol content is mainly influenced by genetic factors with limited effect from environmental factors (Yu et al., 2022). Therefore, the tocopherol content of rice in this study was appropriate for GWAS.

### Population stratification

Rice accessions in our panel have previously been classified into five subpopulations: *Japonica*, *Indica*, *AUS*, *Aromatic*, and *Admix*, according to (Zhao et al., 2011). The diversity level and stratification of the accessions were examined before performing GWAS. A kinship matrix was used to summarize the distribution of pairwise relative relationship coefficients among the accessions in the association panel based on SNP information. A phylogenetic tree and heat map of the values in the kinship matrix were constructed from the SNPs, both of which showed relatedness among the accessions (Figure 2A).

### Markers statistics, LD decay, and GWAS

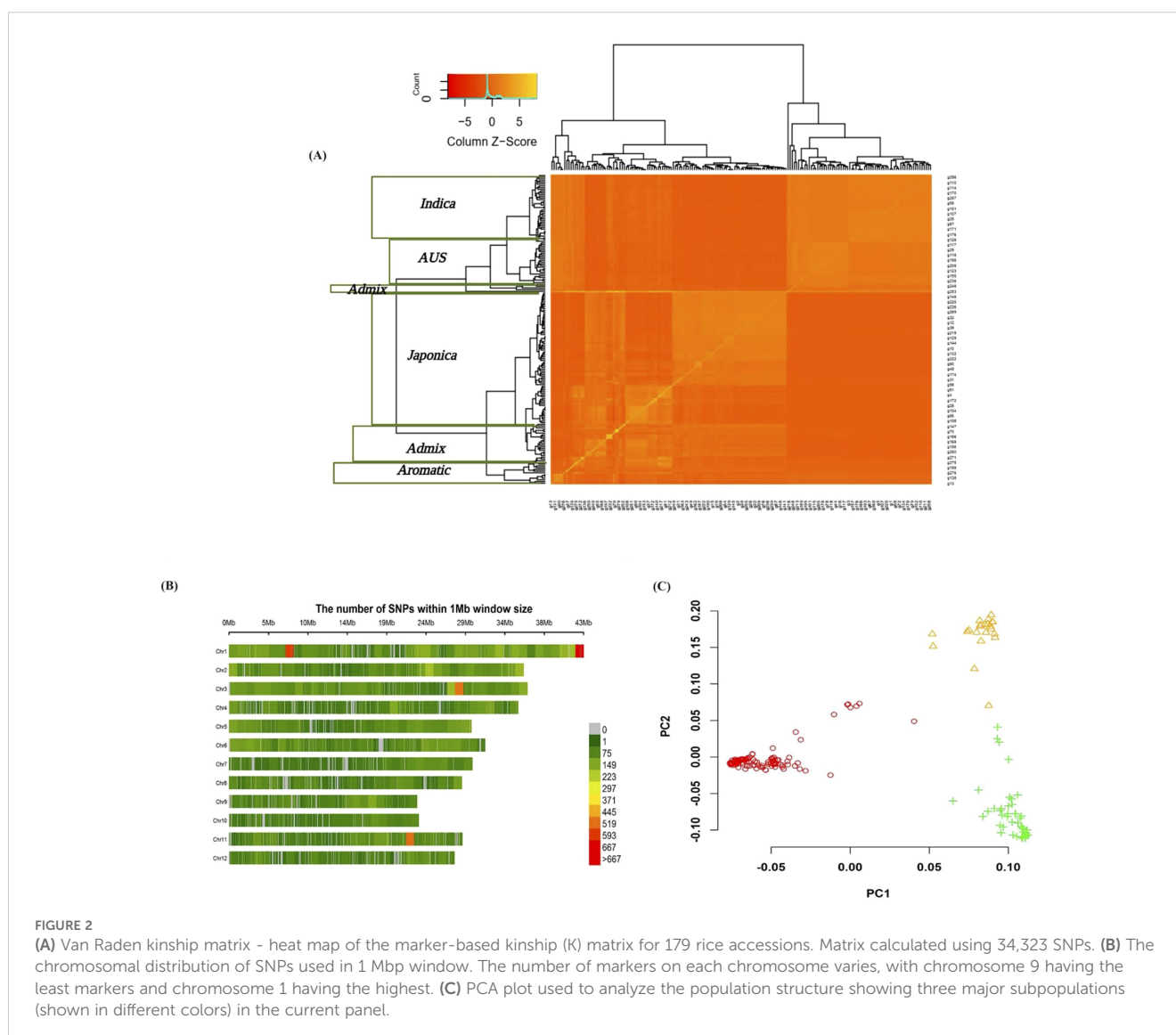
A total of 34,323 SNPs distributed across 12 chromosomes were examined. The number of markers on each chromosome varied, with chromosome 9 having the fewest markers and chromosome 1 having the largest number (Figure 2B). According to the PCA plot (Figure 2C), three major subpopulations were evident in the current association panel. The LD dropped to half of its maximum value at a distance of 300-1300 kb across the 12 chromosomes (Figure 3A). Multi-locus association analyses were carried out using MLM and FarmCPU models (Supplementary Figure S2). Quantile-quantile (Q-Q) plots suggested that the FarmCPU model better fits the expected distribution. Accordingly, FarmCPU was used for all subsequent analyses to generate the corresponding Manhattan plots for alpha and total tocopherols (Figure 3B). A total of 74 ( $\alpha\text{T}$ : 27, total tocopherol: 47) SNPs were significantly associated with tocopherol contents. These significant SNPs were located on chromosomes 1, 2, 3, 6, and 8 (Supplementary Table S3). Thirteen significant QTLs with  $-\log_{10}(p) \geq 4$  for the contents of alpha and total tocopherols were detected. The strongest QTL for  $\alpha\text{T}$  and total

TABLE 1 Statistical and variation analysis of tocopherol (toc) content in the tested rice population (n = 179).

Traits	Mean ( $\mu\text{g/g}$ )	Range ( $\mu\text{g/g}$ )	C.V%	$H^{2b}$	$H^{2n}$	VA	VG
Total toc content	3.70	0.230 - 31.76	6.7	0.87	0.79	27.88	30.71
$\alpha$ -Toc content	3.30	0.011 - 30.83	7.3	0.86	0.76	26.64	29.75

tocopherol was located on chromosome 3 (Supplementary Table S3, Figure 3B–a, b). The flanking regions (SNP position  $\pm$  LD decay) of the 13 QTLs were checked for the presence of the marker-associated genes. For each QTL, the Phenotypic Variation Explained (PVE) was calculated. The PVE for all QTLs with  $-\log_{10}(p) \geq 4$  was greater than 10%, except for the QTLs on chromosome 2 (*qTol2.1*, *qAlph2.1*). Therefore, of the 13 QTLs, 11 were considered major and two (*qTol2.1*, *qAlph2.1*) were minor (Supplementary Table S3). Various regulatory (TFs, miPs, and transposons) and functional genes with roles in cellular transport and signaling were found near the mapped QTLs. Among these genes, 51 primary TFs ( $\alpha T = 22$  and total tocopherol = 29) were identified using PlantTFDB5.0 (<http://planttfdb.gao-lab.org/>; Supplementary Table S4). A

microprotein (miP; RGP-2) on chromosome 1 and a transposon (*Os03g0822100*) on chromosome 3 were also associated with tocopherol content. RGP-2 with a length of 116 amino acids, was for the first time identified in our study. Transporter genes (9) from three families of ABC transporters, nonspanins, and SNARE proteins on chromosomes 1 and 8 were also associated with tocopherol content (Table 2). Protein kinases were identified to be associated with seven QTLs as the probable key signaling factors and belonging to the receptor-like kinases (RLKs), S-domain subfamily of receptor-like kinases (SDRLKs), calcium-dependent protein kinase (CDPK), NAD kinase, cyclin-dependent kinase (CDK), nucleoside monophosphate (NMP) kinase and lipid kinase family (Table 2).



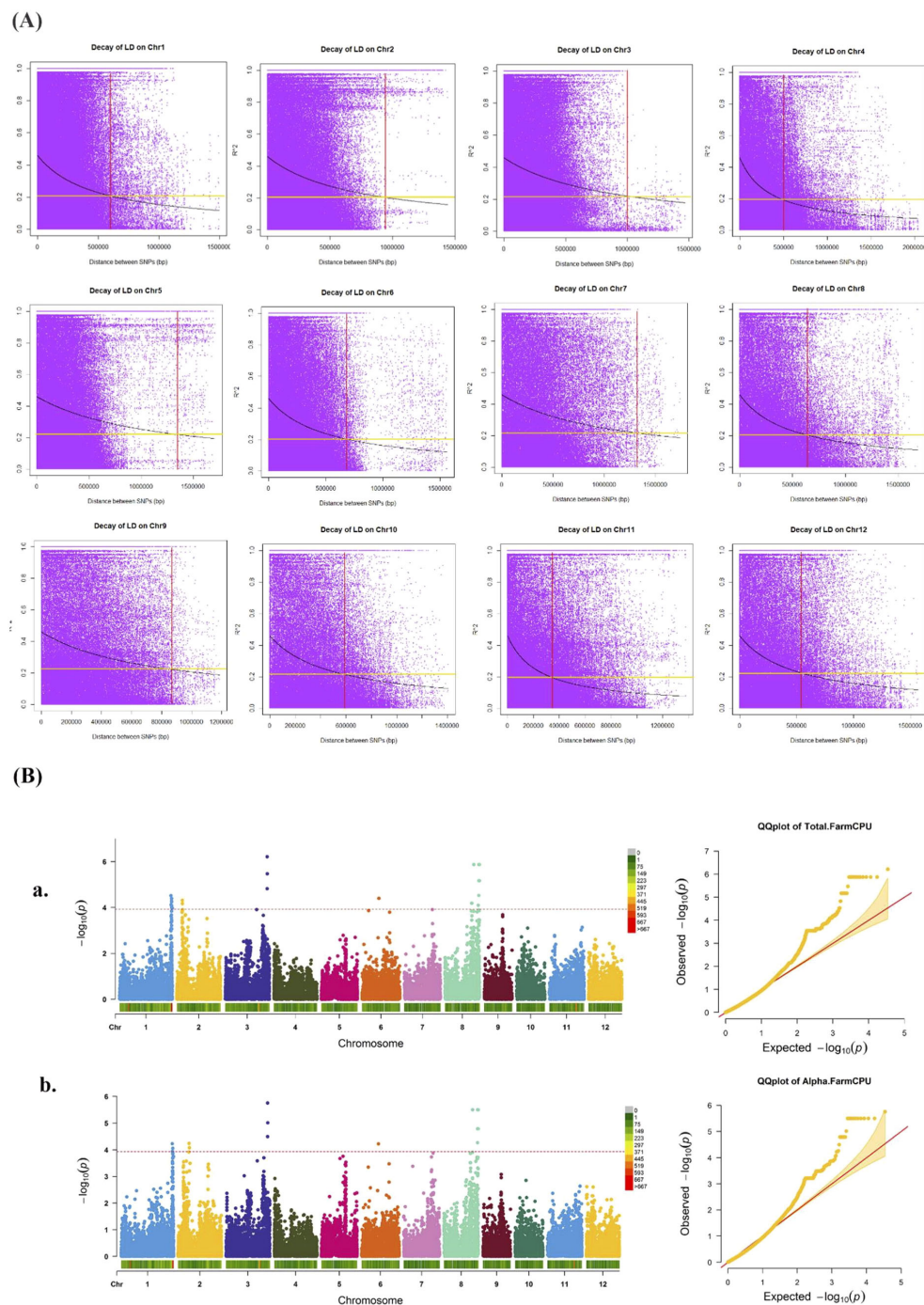


FIGURE 3

(A) The decay of LD along physical distances on 12 rice chromosomes was computed using SNP data of 179 rice accessions. A scatter  $r^2$  against physical distance showed a clean pattern of LD decay in the 179 rice accessions. A critical value of the determination coefficients  $r^2 > 0.2$  was determined to be the threshold for LD decay. (B) Manhattan and Q-Q plots. (a) Total tocopherol, (b)  $\alpha$ -tocopherol. The FarmCPU model resulted in more associations than the MLM model, which revealed 13 significant QTLs with  $-\log_{10}(p) \geq 4$  for  $\alpha$ T and total tocopherol contents.

## Promoter and co-expression analysis

Promoter analysis of the tocopherol biosynthetic pathway genes (*vte1*, *vte2*, *vte3*, *vte4*, *vte5*, *vte6*, *hppd*, *tat*, and *ggpps*) was carried out using 1500 nucleotide (Wang et al., 2015) window upstream of the start codon of each gene to determine if corresponding *cis*-

elements exists. The genes *ZFP15*, *ZFP182*, *NAC92*, *OsMADS59*, *HOS58*, and *WRKY39* have the binding site in the promoter region of different genes of the tocopherol biosynthesis pathway (Supplementary Tables S5, S6).

A co-expression analysis was performed for the different developmental stages of rice with a correlation threshold  $\geq 0.6$ .



TABLE 2 Candidate genes identified via GWAS involved in vitamin E biosynthetic pathway.

Gene	SNP	Gene name	QTL name	Distance (bp)	LOP	Functional annotation	Reference
Os08g0554900	id8007643	Nonaspanin (TM9SF)	qAlph8.2, qTot8.2	88,339	5.9	transmembrane 9 superfamily	
Os08g0555200	id8007643	Nonaspanin (TM9SF)		116,667	5.9		
Os08g0555300	id8007643	Nonaspanin (TM9SF)		119,816	5.9		
Os08g0558600	id8007643	VAMP727		251,888	5.9	knockdown of the Os08g0558600 gene could lead to a reduced tiller number phenotype with no heading	Wang et al., 2013
Os08g0564100	id8007982	ABCF4		122,842	5.2	ABC transporter f family member 4	
Os08g0564300	id8007982	ABCB22		106,884	5.2	The ABCB proteins can serve as auxin influx/efflux transporter and transporters for secondary metabolite as well as iron	Xu et al., 2014
Os08g0563300	id8007982	BET11		165,132	5.2	a critical role in the membrane fusion step of the vesicular transport system	Bao et al., 2012
Os01g0966100	dd1002535	ABCD2	qAlph1, qTot1.1	8,776	4.3	ABC transporter superfamily ABCD subgroup member 2	
Os08g0544400	id8007450	ABCG45	qTot8.2	67,773	4.1	high expression in reproductive stages as well as in the stress- and hormone- (IAA and BAP treatments) treated samples	Gupta et al., 2019
Os03g0821900	id3016979	RLCK118	qAlph3.1, qTot3.1	95,128	6.2	positively regulate chitin- and PGN-induced responses in rice/ positive regulators of rice immunity and plant development	Li et al., 2017; Fan et al., 2018; Xiao et al., 2022
Os03g0825300	id3016979	RLCK119		65,355	6.2	Serine/threonine protein kinase domain containing protein.	
Os03g0825800	id3016979	RLCK120		87,808	6.2	Protein kinase	
Os03g0828800	id3016979	SDRLK60		210,372	6.2	S-Domain receptor like kinase-60	
Os08g0540400	id8007643	CDPK21	qTot8.2	150,432	5.9	involved in multiple stress signaling pathways	Mittal et al., 2017
Os01g0957000	id1027324	NADK1	qAlph1.1, qTot1.1	31,361	4.5	NAD kinase, Mediation of intracellular redox balance, Drought tolerance	
Os01g0957100	id1027324	Protein kinase		35,288	4.5	-	
Os01g0958000	id1027324	CDKC1		80,262	4.5	key regulatory roles in diverse cellular functions, including cell-cycle progression, transcription and translation	Kitsios et al., 2008
Os01g0960400	id1027324	Protein kinase		196,309	4.5	-	
Os01g0965400	dd1002410	YGL8		43,018	4.4	chlorophyll biosynthesis gene	Peng et al., 2019
Os01g0964400	dd1002410	RLCK56-receptor-like kinases (RLKs)		76,058	4.4	Receptor-like Cytoplasmic Kinase 56	
Os02g0182600	id2002458	Similar to Protein kinase	qTol2.1	58,432	4.3	-	

(Continued)

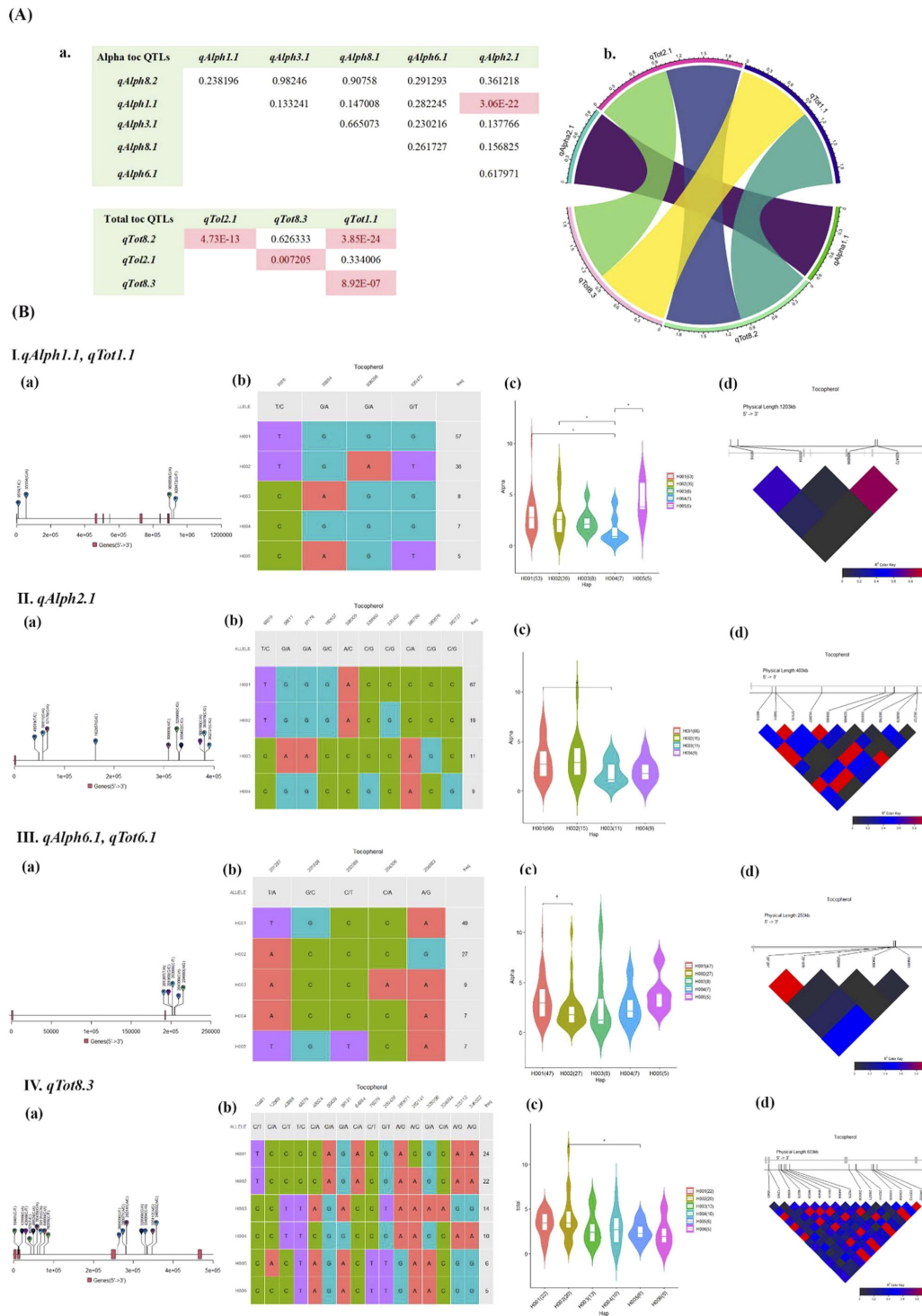
TABLE 2 Continued

Gene	SNP	Gene name	QTL name	Distance (bp)	LOP	Functional annotation	Reference
<i>Os02g0186500</i>	id2002458	<i>RLCK64</i>		154,668	4.3	Receptor-like Cytoplasmic Kinase 64	
<i>Os08g0453800</i>	id8006337	<i>CDKC</i>	<i>qTot8.3</i>	141,978	4.2	involved in developmental events as well as in the plant salt stress response mechanism through an ABA-signaling pathway	Huang et al., 2008
<i>Os08g0446200</i>	id8006210	<i>PEPR1-receptor kinase</i>		42,933	4.2	defense signaling	Shinya et al., 2018
<i>Os08g0446301</i>	id8006210	<i>INRPK1c/ST-Pks</i>		37,438	4.2	-	
<i>Os08g0450800</i>	id8006210	<i>PIP5K-lipid kinase</i>		206,815	4.2	phosphatidylinositol signaling pathway and essential functions in growth, development, and biotic and abiotic stresses responses	Zhang et al., 2020
<i>Os03g0822100</i>	id3016979	<i>Similar to Transposase</i>	<i>qAlph3.1, qTot3.1</i>	76,895	6.2	Similar to Transposase	
<i>Os01g0974701</i>	ud1001752	<i>RGP-2</i>	<i>qTot1.1</i>	21,837	4	Similar to RNA-binding glycine rich protein	
<i>Os03g0820400</i>	id3016979	<i>ZFP15</i>	<i>qAlph3.1, qTot3.1</i>	174,970	6.2	Improvements in salinity and drought tolerance	Zhang et al., 2012; Wang et al., 2022
<i>Os03g0820300</i>	id3016979	<i>ZFP182</i>		178,494	6.2	salt tolerance related genes	Ning et al., 2018
<i>Os03g0818700</i>	id3016979	<i>MYB-like</i>		252,070	6.2	Similar to MYB transcription factor	
<i>Os03g0818800</i>	id3016979	<i>EREBP33</i>		244,481	6.2	tolerance to cold stress, drought response	Yang et al., 2023; Maghraby and Alzalaty, 2024
<i>Os08g0543900</i>	id8007643	<i>bZIP68</i>	<i>qTot8.2</i>	31,372	5.9	salt responsive; positively regulates ABA-independent osmotic stress signaling	Hossain et al., 2016; Zhou et al., 2022
<i>Os08g0477900</i>	id8006747	<i>bHLH042</i>	<i>qAlph8.1, qTot8.1</i>	50,709	5.9	basic helix-loop-helix protein 042	
<i>Os08g0549600</i>	id8007643	<i>OsbZIP69</i>		99,349	5.9	promotes flowering in rice	Cerise et al., 2021
<i>Os06g0344900</i>	id6008406	<i>NAC92</i>	<i>qAlph6.1, qTot6.1</i>	201,092	4.4	role in drought stress by regulating the related genes involved in drought tolerance	Wang et al., 2022
<i>Os06g0347700</i>	id6008406	<i>OsMADS59</i>		8,696	4.4	Phenylpropanoid biosynthesis pathway	Sun et al., 2023
<i>Os06g0348800</i>	id6008406	<i>GLK1</i>		73,152	4.4	regulation of plastid development, carotenoid production	Zheng et al., 2022; Li and Yao, 2022
<i>Os02g0182800</i>	id2002458	<i>HOS58</i>	<i>qTot2.1</i>	32,964	4.3	KNOX family class 2 homeodomain protein	
<i>Os02g0265200</i>	id2004522	<i>WRKY39</i>	<i>qAlph2.1</i>	48,328	4.2	heat responsive	Li et al., 2010

The TFs identified in the GWAS analyses were checked via promoter analysis for common *cis*-elements. If they showed co-expression with tocopherol biosynthetic genes they were introduced as candidate TFs with probable roles in controlling tocopherol content. We identified 12 such TFs, including ERF, WRKY, TALE, bZIP, bHLH, NAC, MADS BOX, and MYB (Table 2).

## Epistatic interactions among QTLs

Complex traits in rice are regulated by epistatic interactions in addition to the additive effects of each gene, which makes the relationship between genotype and phenotype complex (Kato and Horibata, 2022). Here, six QTLs had epistatic effects (Figure 4A-a)



**FIGURE 4**  
**(A)** a) Evidence of epistatic effects between alpha and total tocopherol QTLs. Interactions  $\leq 0.05$  are significant. (b) Circos plot of epistasis interaction between QTLs for alpha and total tocopherol detected from the GWAS result. QTLs that had significant epistatic associations are located around the circle. The Circlize package was used to visualize the significant epistatic interactions. **(B)** Haplotype analyses for different QTLs related to alpha and total tocopherols. (I) *qAlph1*, *qTot1.1*, (II) *qAlph2.1*, (III) *qAlph6.1*, *qTot6.1* and (IV) *qTot8.3*. (a) Visualization of SNP position above gene model, the black line represents the genome and rectangles represent exons; (b) Haplotype classification, each row represents a haplotype, colored columns represent loci, and the last column shows the frequency of each haplotype; (c) Phenotype comparisons among accessions possessing different haplotypes; (d) LD-block visualization of SNP sites in the locus. \*Significant at  $p \leq 0.05$ ; † significant at  $p \leq 0.1$ . The red color indicates perfect LD, and the black color indicates no LD.

between SNPs in QTLs *qAlph1.1* and *qAlph2.1* on chromosomes 1 and 2 affecting alpha-tocopherol ( $p \leq 0.05$ ), pairs of *qTot8.2* and *qTot2.1*; *qTot8.2* and *qTot1.1*; *qTot2.1* and *qTot8.3*; *qTot8.3* and *qTot1.1* on chromosomes 8, 1 and 2 all affecting total tocopherol ( $p \leq 0.05$ ) (Figure 4A-b). The molecular basis behind these interactions may correspond with the regulation of gene expression and/or activity by *trans*-acting elements.

## Identification of favorable haplotypes of QTLs

To identify the cumulative effects of SNPs, haplotype analysis was conducted using all the QTLs although only statistically significant QTLs are being reported here. Haplotype analysis of QTLs *qAlph1.1* and *qTot1.1*, associated with alpha and total tocopherol on chromosome 1, formed a haplotype block with four SNP markers, which consisted of five haplogroups in our association panel (Figures 4B-I-a, b). All four markers within these QTLs showed substantial LD (Figures 4B-I-d), and variation in the haplotype alleles led to significant differences ( $p \leq 0.05$ ) between H001/H004, H002/H004, and H004/H005 haplotypes for both alpha and total tocopherol. The median of alpha and total tocopherol values in the violin plot of the haplogroups were 2.72 (H001), 2.65 (H002), 2.51 (H003), 1.34 (H004), 4.2 (H005)  $\mu\text{g/g}$ , respectively, and the highest amount of tocopherol was related to H005 (Figures 4B-I-c). H001 had the highest frequency (50.44%) and was mostly represented by *japonica* (TEJ) varieties based on the haplotype network (Supplementary Figure S3). Haplotype analysis of *qAlph2.1* resulted in four haplogroups with 10 SNPs in LD on chromosome 2 among the 179 rice accessions. The highest amount of tocopherol was associated with H002. The median of H001 and H003 were 2.65 and 1.32  $\mu\text{g/g}$ , respectively, showing significant differences ( $p \leq 0.05$ ). The highest haplotype frequency was observed for H001 (63.0%), represented by TEJ, TRJ, and Admix varieties (Figure 4B-II-a-d, Supplementary Figure S3). Also, haplotype analysis of QTLs *qAlph6.1*, and *qTot6.1* associated with alpha and total tocopherol on chromosome 6, formed a haplotype block with five SNP markers, which consisted of five haplogroups in the association panel. All five markers were in LD, and variation in these haplotype alleles led to significant differences between the H001 and H002 haplotypes in terms of alpha and total tocopherol. The highest amount of tocopherol was found for the H005 haplotype, while the median values for H001 and H002 were 2.83 and 1.72  $\mu\text{g/g}$ , respectively, showing a significant difference ( $p \leq 0.1$ ). H001 (32%), H002 (30%), and H003 (29%) have the highest frequency, which is represented by Indica, TRJ, and TEJ varieties, respectively (Figure 4B-III-a-d, Supplementary Figure S3). Haplotype analysis of *qTot8.3* resulted in the formation of six haplogroups based on 16 SNPs on chromosome 8. All 16 markers showed substantial LD. For total tocopherol, the median values among the H002 and H005 haplogroups were 3.51 and 2.40  $\mu\text{g/g}$ , respectively, showing significant differences ( $p \leq 0.05$ ). The highest amount of tocopherol was related to H001. H001 (30%) and H002 (27%) have the highest frequency, which is represented by TEJ varieties (Figure 4B-IV-a-d, Supplementary Figure S3).

## Expression analysis of candidate genes

To further verify the potential impact of the candidate genes on the regulation of tocopherol biosynthesis, the expression patterns of the candidate genes in various tissues [seeds (5 and 10 DPP), leaves (20-day-old) and shoots] based on data from the Rice Genome Annotation Project database were analyzed (Figure 5A). The results showed that among the transporter genes, *ABCF4* is expressed in all tissues, with the highest expression in the seed 5 DPP stage. *VAMP727* also showed a small extent of expression in all tissues, but the highest expression was in the shoot. *ABCG45* was slightly expressed in the seed 5 DPP and 10 DPP stages. The highest expression of *ABCD2* was in shoots, and *nonaspanin* was not expressed in any tissue we checked in this study. Among the signaling genes, *PIP5Ks* were expressed in all tissues; the highest expression was in seed 10 DPP. *RLCK118* was slightly expressed in all tissues. The highest expression in signaling genes belonged to *YGL8*, which was observed in leaves. Other genes that were involved in signaling were expressed in a small amount in different tissues. Transposon (*Os03g0822100*) and microprotein *RGP-2* also showed the highest expression in shoot tissue. Among TFs, the highest expression was found for *HOS58* in all tissues. *bZIP68* was weakly expressed in all tissues. The highest expression of *GLK1* was in leaves (20-day-old). The highest expression of *ZFP182* and *EREBP33* were observed in shoots. *ZFP15* was slightly expressed in shoots. Based on these results, two genes, *ABCF4*, and *PIP5Ks* from the transporter and signaling categories, respectively, were selected as candidates for validation using real-time qRT-PCR.

## Interaction network analysis

The STRING tool was employed to construct a protein association network between candidate genes and main genes (*VTE1*, *VTE4*, *HPPD*) involved in tocopherol biosynthesis. The results showed that *WRKY39*, *bHLH42*, *ZFP15*, and *ZFP182* are co-expressed with *VTE4*, and *ABCF4* and *RGP-2* are co-expressed with *VTE1*. Also, *RLCK119* is co-expressed with *HPPD* (Supplementary Figure S4). From these results, *WRKY39* and *RGP-2* genes in the transcription factor and microprotein categories were selected as the candidates for real-time qRT-PCR analysis.

## Predicting the effects of detected SNPs on gene function

Prediction of SNP effects was performed using SnpEff software. The predicted effects were categorized by impact as a modifier (with impact on noncoding regions), low (synonymous substitution), moderate (non-synonymous substitution), or high (disruptive impact on the protein). None of the associated SNPs fell within the sequence of the candidate genes, with two exceptions, namely dd1002410 and Id6008406, which were located in the intron region of *ABCD2* and *OsMADS59* genes, respectively. As a result, we selected *OsMADS59* for inclusion in the real-time qRT-PCR analyses.



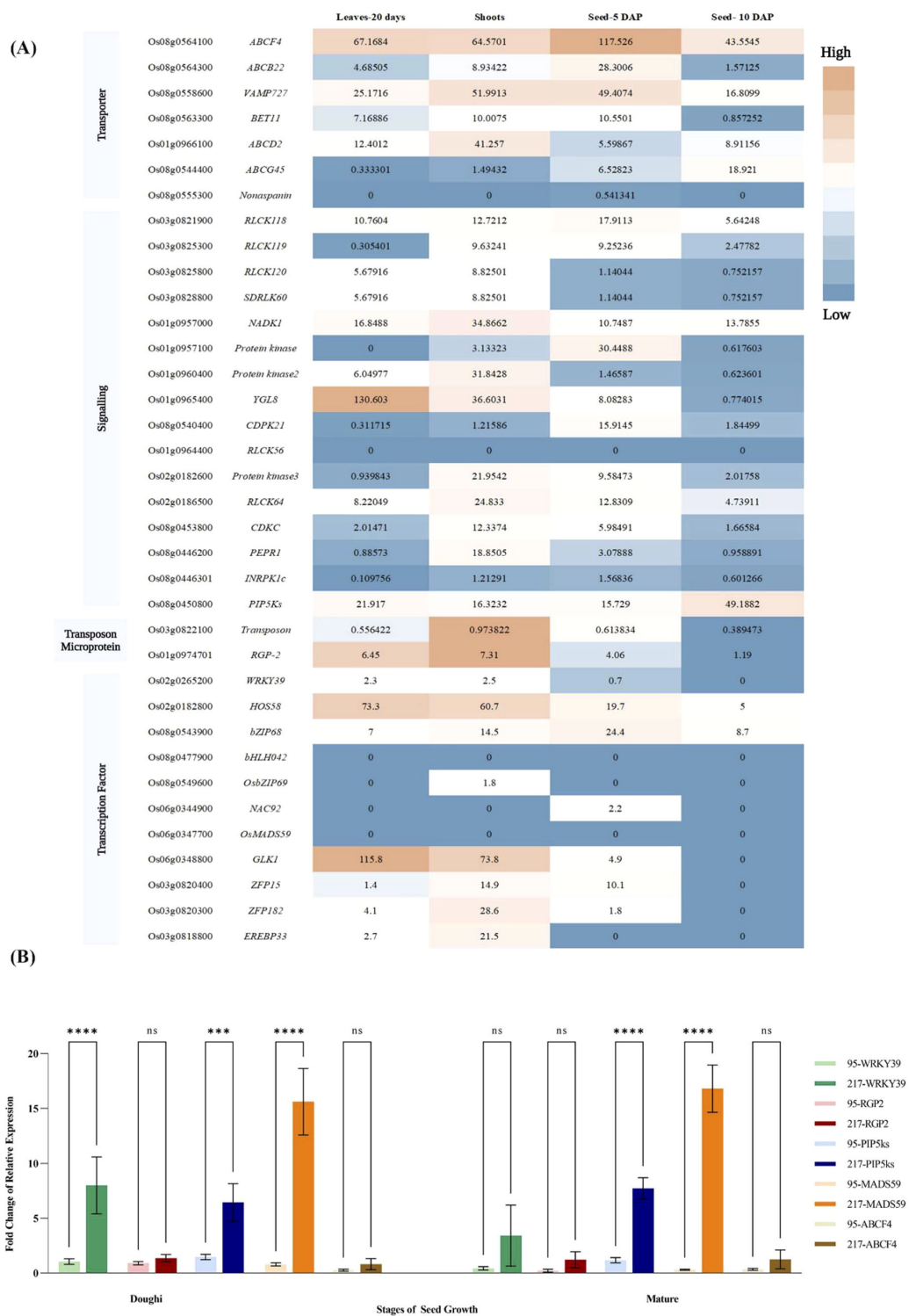


FIGURE 5

(A) Heatmap of candidate gene expression analysis by RNA-seq data obtained from RGAP database for seed 5 and 10 DPP, leaves (20-day-old), and shoots. Dark orange boxes indicate high transcript levels, and blue boxes indicate low transcript levels. (B) Expression analysis of *WRKY39*, *RGP*, *PIP5ks*, *MADS59*, and *ABCF4* by real-time qRT-PCR. Amplification of cDNA from three developmental stages of seeds, i.e., milky, doughy, and maturity in two rice low tocopherol (g95) and high tocopherol (g217) content were carried out. A housekeeping gene, *OsActin*, was used as the control, and the expression data was compared to those of the milky stage. Data were analyzed using a two-way ANOVA, followed by Tukey's multiple comparison test with a 95% confidence interval. Significantly different comparisons between #g95 vs #g217 at stages of seed growth are displayed on graphs using asterisks (\*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ ). No asterisk indicates that the difference is not significant. Error bars represent mean  $\pm$  SD ( $n = 3$  biological replicates).

## qRT-PCR validation of candidate genes

We used qRT-PCR to assess the relative expression levels of five candidate genes, including *WRKY39* (*Os02g0265200*), *RGP-2* (*Os01g0974701*), *PIP5K* (*Os08g0450800*), *OsMADS59* (*Os06g0347700*) and *ABCF4* (*Os08g0564100*) across different seed development stages (milky, doughy, maturity) in two accessions with high (g217) and low (g95) tocopherol contents, respectively (Figure 5B). The qRT-PCR results showed that the expression levels of all five selected genes were significantly higher in genotype g217 (high tocopherol) than g95 (low tocopherol) in both the doughy and mature stages. The highest expression was seen for *MADS59* in both stages. Two-way ANOVA analysis showed significant differences between the two genotypes g95 and g217 in the doughy stage for *WRKY39* ( $P < 0.0001$ ), *PIP5Ks* ( $P < 0.001$ ), and *OsMADS59* ( $P < 0.0001$ ). Also, in the mature stage, two genes *PIP5Ks* and *OsMADS59* had statistically significant differences ( $P < 0.0001$ ). *WRKY39* expression was not significantly different between the two genotypes in the mature stage. There was no significant difference between expression levels for the other genes. These results suggest that *WRKY39*, *PIP5Ks*, and *OsMADS59* may play a crucial role in tocopherol biosynthesis in rice seeds.

## Gene ontology enrichment and KEGG pathway determination

To identify the pathways that had the most significant involvement of the identified genes that displayed differential expression (DEGs), candidate genes were submitted to ShinyGO v0.75 for GO and KEGG pathway analysis. GO analysis revealed that in the biological process category, the DEGs were enriched for involvement in protein phosphorylation, phosphorylation, regulation of transcription, phosphorus metabolic process, nucleobase-containing compound biosynthetic process, organic cyclic compound biosynthetic process, regulation of RNA biosynthetic process, nucleic acid-templated transcription and aromatic compound biosynthetic process (Figure 6A-i). The GO analysis showed that the candidate genes exhibited prominent enrichment in various CC terms for cellular components. These included endosome membrane, endosome, vesicle membrane, cytoplasmic vesicle, intracellular vesicle, bounding membrane of organelle, Golgi membrane, and plasma membrane (Figure 6A-ii). The molecular functions of candidate genes included kinase activity, phosphotransferase activity, ATP binding, adenylyl nucleotide binding, purine ribonucleotide binding, protein serine/threonine kinase activity, transferase activity, DNA-binding transcription factor activity (Figure 6A-iii). KEGG pathway analysis demonstrated that candidate genes were significantly enriched in ABC transporters, biosynthesis of cofactors, nicotinate, and nicotinamide metabolism (Figure 6A-iv).

An overview of the candidate gene expression is provided in Figure 6B, which shows that they can be sorted into 9 BINs or subBINs, reflecting major cellular or functional processes. A high

proportion of genes were involved in vesicle transport and protein targeting (*VAMP727*, *BET11*), regulation of transcription (*EREBP33*, *bHLH042*, *GLK1*, *HOS58*, *OsMADS59*, *WRKY39*, *bZIP68*), regulation (*PEPR1*, *RLCK56*, *RLCK120*, *PIP5Ks*, *CDPK21*), transport (*ABCF4*, *ABCB22*, *ABCD2*, *ABCG45*) and other functions, such as DNA synthesis and development. Only one gene (*ZFP182*) was included in BIN 35.1 (not assigned) and was not placed in any group. Marked changes in the expression pattern were found in all of these categories.

## GC-MS analysis

To study whether there is a relationship between the amount of vitamin E and the other fatty acids, one-step derivatization and GC-MS for the isolated grain oil of five selected rice genotypes (g57, g93, g170, g253, g275), were carried out. Rice seed fatty acid was mostly composed of palmitic acid ( $C_{16:0}$ ), oleic acid ( $C_{18:1}$ ), linoleic acid ( $C_{18:2}$ ), and stearic acid ( $C_{18:0}$ ) (Table 3). The latter was the least abundant (3.44-9.3%) in the five selected genotypes, while oleic acid was the most abundant (30.1-36.28%). According to our limited study, a positive correlation was observed between the amount of alpha and total tocopherol and  $C16:0$ ,  $C18:1$ , and  $C18:2$  fatty acids, but the correlation between the amount of tocopherol and  $C18:0$  was negative. Overall, a positive correlation between total tocopherol and total fatty acid content was observed in rice oil in this study (Supplementary Figure S5; Supplementary Table S7).

## Discussion

Tocopherols are important for grain nutrition quality and play an important functional role in human health due to their ability to prevent the oxidation of unsaturated fatty acids in cell membranes by scavenging free radicals (Shaw and Rajcan, 2017). For humans, daily tocopherol supplementation can decrease the risk of cancer and cardiovascular disease (Shaw et al., 2016). In plants, tocopherols protect chloroplasts from photooxidative damage during seed germination and improve oil stability in oilseeds (Morelli et al., 2023). Feedstock enriched with tocopherol improves and maintains growth and health in animals (Morand-Laffargue et al., 2023).

The tocopherol biosynthetic pathway, encoded by *VTE* genes (*VTE1-6*), is highly conserved in plants (Albert et al., 2022). The activity of *VTE7* in *Arabidopsis* and maize, an  $\alpha/\beta$  hydrolase that interfaces with chlorophyll synthesis to release phytol, is unequivocal (Albert et al., 2022). Based on our alignment analysis (for both nucleotide and protein, data not shown) using in-house scripts in R, *VTE7* could not be detected in the rice genome. This study found positive relationships between total fatty acids and total tocopherol content. According to a recent study (Chu et al., 2023) a positive correlation between total tocopherol and total fatty acids has also been observed in soja. Despite the availability of considerable information about vitamin E, our knowledge about the genetic architecture of vitamin E content in

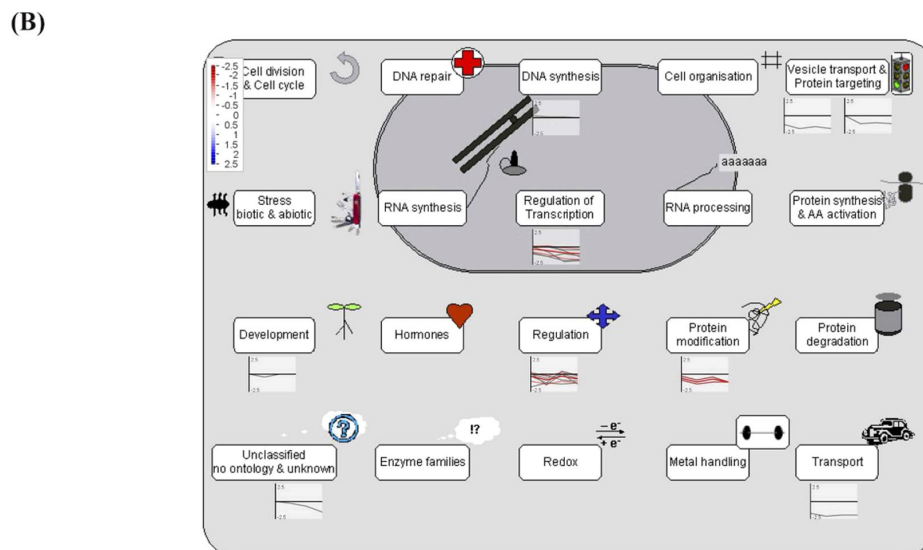
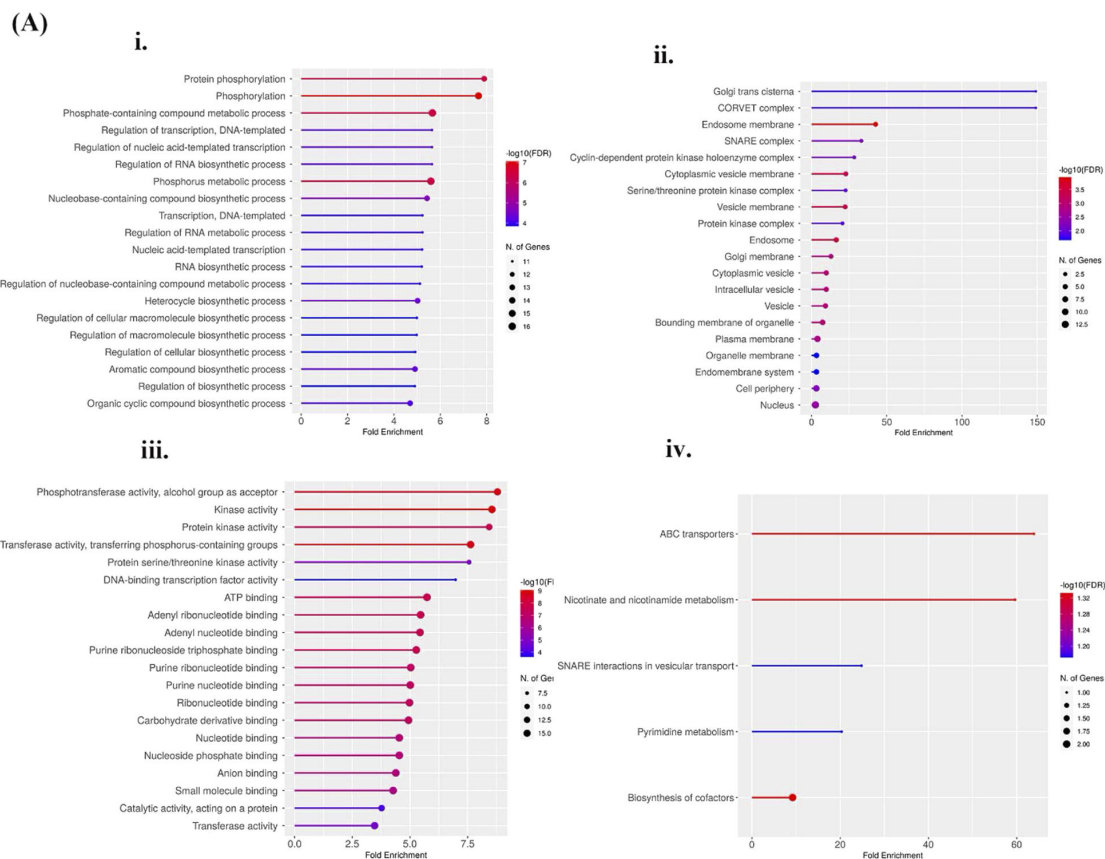


FIGURE 6

(A) Functional enrichment analysis of candidate genes given gene ontology (GO) and KEGG. GO analysis revealed that candidate genes were significantly enriched in (i) biological processes (ii) cellular components, (iii) molecular functions, and (iv) significantly enriched KEGG terms obtained from KEGG analysis. KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology. (B) MapMan analysis. Cell function overview associated with candidate gene. The studied genes are mainly effective in the regulation, protein modification, transport, and regulation of transcription, which are the precursors of tocopherol biosynthesis. For each functional group, the genes that increase and decrease more are shown as a series of blue and red lines. MapMan software (ver. 3.6.0RC1) was used to map transcriptomic data. Osa\_RAPDB\_mapping files were downloaded from the Map-Man store server (<http://mapman.gabipd.org/>).

TABLE 3 Fatty acid and tocopherol composition in rice oil (%) in selected accessions via GC-MS.

Tocopherol content ( $\mu\text{g/g}$ )			Main fatty acid in rice oil (%)			
Genotype	$\alpha$ -Toc	Total toc	Palmitic acid	Linoleic acid	Oleic acid	Stearic acid
g57	1.64	2.94	18.3	22.9	30.1	3.9
g93	0.076	0.23	17.17	28.99	36.28	3.8
g170	0.38	2.36	20.34	24.17	30.39	9.3
g253	30.83	31.76	18.14	27.78	33.75	3.44
g275	0.33	0.35	16.29	22.83	34.85	5.87

rice grains is rather rudimentary. This study aimed to employ GWAS to identify QTLs and candidate genes involved in the regulation, signaling, and transport of alpha and total tocopherols using natural rice accessions.

Association studies benefit from using genetically diverse germplasm, allowing for the examination of total allelic diversity derived from historical recombination events. In the present study, we implemented two different models, a univariate MLM model and a multivariate mixed model, FarmCPU. According to Segura et al. (2012), using the multivariate GWAS method addresses the issue of confounding between covariates and the test marker and lowers the false discovery rate (FDR) compared to univariate GWAS for the same significance threshold. This facilitates the detection of a greater number of QTLs. By examining Q-Q plots, the most suitable model for our data was FarmCPU. A panel of 179 rice genotypes was used to explore the diversity of tocopherol content in rice grain oil, using HPLC equipped with a fluorescent detector. Cluster analysis of the kinship matrix and PCA calculation indicated sufficient heterogeneity of genotypes that made it suitable for GWAS and allowed us to control for false associations due to the underlying genetic structure of the association population (Li et al., 2015).

Alpha and total tocopherol were detected similarly as in an earlier report by Lipka et al. (2015). The total tocopherol detected in this study ranged between 0.230 - 31.76 ( $\mu\text{g/g}$ ). The dominant tocopherol isoform in rice seed oil (including the hull) was alpha-tocopherol ( $\alpha$ T). Generally, tocopherol is higher in bran, but we couldn't isolate the bran in this study. Therefore, our results may be slightly different from earlier studies (Wang et al., 2015). Quantification data clearly showed natural variations in the contents of tocopherols in our rice accessions (Supplementary Figure S1). Tocopherol content is a complex trait governed by several genetic loci. Phenotypic variation in the content and composition of tocopherols is primarily due to genetic effects, as shown by the high narrow-sense heritabilities ( $H^{2n}$ ) for alpha (0.79) and total (0.76) tocopherols. The values were not significantly different from broad-sense heritabilities ( $H^{2b}$ ), which suggests a minimal effect of other genetic factors, such as epistasis and dominance. High and stable heritability has always been useful for identifying strong associations between markers and traits of interest in association studies (Courtois et al., 2013). Significant SNPs (74) were identified using FarmCPU for alpha and total tocopherol content. In addition, unique QTLs (13) were detected, and the flanking sequences were searched for the genes associated with the

putative function of defining alpha and total tocopherol content in rice grains. Another notable finding was that five pairs of epistatic QTLs were associated with tocopherol content. Epistasis is an important genetic component and a plausible feature of the genetic architecture of quantitative traits (Phillips, 2008). The change of a gene effect can often be mirrored in the placement and function of its signal-transduction pathway (Mackay et al., 2009). Here, the type of connection remains elusive. However, each QTL had several genes, and genes may have different epistatic effects. In the present study, 40 candidate genes were located within the confidence interval of the identified QTLs. Based on GWAS, 12 TFs were identified. Promoter analysis showed that *ZFP15*, *ZFP182*, *NAC92*, *OsMADS59*, *HOS58*, and *WRKY39* have binding sites in the promoter of genes involved in the biosynthesis of tocopherols. The members of the WRKY family of TFs, including *WRKY39* (*qAlph2.1*, ~48 kb away from SNP id2004522, associated with alpha-tocopherol), have been reported to participate in growth, development, metabolism, and responses to environmental cues (Phukan and Shukla, 2016; Chen et al., 2017), which again suggests the possible involvement of such TFs in regulating vitamin E content. *HOS58*, a TALE member, was identified as another candidate TF for regulating vitamin E content in rice grain. *HOS58* is located at *qTot2.1*, ~32 kb away from SNP id2002458, and was significantly associated with total tocopherol content. *ZFP15* and *ZFP182* of the C2H2 TF family were the other candidate TFs; both were common to two QTLs, including *qAlph3.1* and *qTot3.1* (~178 kb away from SNP id3016979). Members of this family in rice have been reported to be involved in response to abiotic stresses (Gourcilleau et al., 2011). *NAC92* (*qAlph6.1* and *qTot6.1*, ~201 kb away from SNP id6008406) and *OsMADS59* (*qAlph6.1* and *qTot6.1*, ~8 kb away from SNP id6008406) of MADS-Box family were the other TFs found in the close vicinity of the detected QTLs.

Haplotype analysis pinpointed the allelic variations with contributing factors in the tocopherol contents, which can be used for both pyramiding and marker-assisted breeding of rice (Varshney and Kumar, 2019). Four SNP combinations in the H005 haplogroup of *qAlph1.1* and *qTot1.1* QTLs had a greater effect than other haplotype groups. The combination of favorable alleles identified within the H002 haplotype was significantly higher for total tocopherol in *qTot8.3* QTL. Four SNP combinations within the H002 haplogroup of *qAlph2.1* QTL were more effective than other haplotype groups, and five SNP combinations within the H005 haplogroup of *qAlph6.1* and *qTot6.1* QTLs were more effective than



other haplotype groups. These favorable alleles can be used in pyramiding into a target rice line by marker-assisted selection.

Studies in humans and animals have suggested that tocopherol-binding proteins (TBPs) are important for distributing and transporting alpha-tocopherol among different tissues (Niu et al., 2022). Bermúdez et al. (2018) identified the *SITBP* (*Solanum lycopersicum* tocopherol-binding protein) as a homolog of the human alpha-tocopherol transfer protein (*HsTTP*). *In vitro* biochemical assay suggested that *SITBP* possesses alpha-tocopherol binding ability. *SITBP* is chloroplast-targeted, and knocking down *SITBP* expression in tomato conferred disorders in tocopherol, carotenoid, and lipid compositions (Bermúdez et al., 2018; Niu et al., 2022). Here, nine transporter genes from three families of ABC transporters, nonaspanins, and SNARE proteins on chromosomes 1 and 8 were identified as possible associations with tocopherol transport. The ATP-binding cassette (ABC) transporters are among the largest superfamilies involved in various biological processes (Akhtar and Turner, 2022). *ABCF4*, *ABCB22*, *ABCD2*, and *ABCG45* were found in our study (Table 2). Nonaspanins constitute a family of proteins known as TM9SF, which may function as a channel or small molecule transporter. Proteins in this group are endosomal integral membrane proteins (Sugasawa et al., 2001). *Os08g0554900*, *Os08g0555200*, and *Os08g0555300* belong to this family. Nonaspanin (*TM9SF*) and *VAMP727* co-located in QTLs *qAlph8.2* and *qTot8.2* with ~88–119 kb and 252 kb away from id8007643, respectively. SNARE protein (soluble N-ethylmaleimide-sensitive factor attachment protein receptor), is a highly conserved superfamily of proteins that mediate vesicle transport between endosomes and the trafficking to the plasma membrane of all eukaryotic cells (Jahn and Scheller, 2006). *VAMP727* and *BET11* were in this family. *BET11* is common in QTLs *qAlph8.2* and *qTot8.2*, ~165 kb away from id8007982.

Cellular signaling in plants is a complex phenomenon governed by several interconnected pathways. Protein kinases play pleiotropic roles in cells and act as predominant mediators in coordinating cellular responses to their environment. The kinase pathway is important for tocopherol biosynthesis, as the Arabidopsis *vte5* mutant contains reduced levels of tocopherol (Romer et al., 2024). In this study, *RLCK118-120* and *SDRLK60* were common in QTLs *qAlph3.1* and *qTot3.1*, ~65–95 and 210 kb away from id301697, respectively. *RLCK56* is located in QTLs *qAlph1.1* and *qTot1.1*, ~76 kb away from id1002410. *RLCK64* located in QTL *qTot1.1*, ~154 kb away from id2002458, was significantly associated with total tocopherol.

Receptor kinases (RKs) are of paramount importance in transmembrane signaling that governs plant reproduction, growth, development, and adaptation to diverse environmental conditions (Liang and Zhou, 2018). *NADK1* (*Os01g0957100*) and *CDKC1* (*Os01g0960400*) were common in two QTLs *qAlph1.1* and *qTot1.1*, ~31, 35, 80, 196 kb away from id1027324, respectively, that were significantly associated with alpha and total tocopherol. One of the key enzymes that regulate NAD(H)/NADP(H) balance is NAD kinase (NADK; EC 2.7.1.23), which catalyzes NAD phosphorylation in the presence of ATP (Takahara et al., 2010). Uridylate kinase is common in two QTLs *qAlph1.1* and *qTot1.1*, ~43 kb away from id1002410. The *UMP kinase* gene plays an important role in

regulating chloroplast development and stress response in rice (Dong et al., 2019). *CDPK21* is located in QTL *qTot8.2*, ~150 kb away from id8007643, and was significantly associated with total tocopherol. Calcium-dependent protein kinases (CDPKs) play significant roles in regulating plant growth and development in response to various stresses, including drought (Mittal et al., 2017). *CDKC* located in QTL *qTot8.3*, ~141 kb away from id8006337 was significantly associated with total-tocopherol. Cyclin-dependent protein kinases (CDKs) form a conserved superfamily of eukaryotic serine/threonine protein kinases, which require binding to a regulatory cyclin for activity. CDKs are organized in several gene families and are involved in different aspects of cell biology, such as gene transcription, cell proliferation, and differentiation (Huang et al., 2008). *PEPR1*, *INRPK1c*, and *PIP5K* are located in QTL *qTot8.3*, ~42, 37, 206 kb away from id8006210, which were significantly associated with total tocopherol. Leucine-rich repeat receptor-like protein kinase (*PEPR1*) is a plant cell membrane-localized Leucine-rich repeat (LRR) receptor kinase that plays a critical role in plant innate immunity (Song et al., 1995). *PIP5K* or *PI5K* family regulates diverse cellular processes such as G protein-coupled receptor (GPCR) signaling, vesicle trafficking, chemotaxis, and cellular movement (Doughman et al., 2003; Table 2). MapMan is a useful tool to provide global views of diverse aspects of data and can be used to functionally classify rice genes (Usadel et al., 2009). Cell function analysis using the MapMan comprehensive overview highlighted the key pathways influenced by tocopherol biosynthesis.

qRT-PCR has been proposed for preliminary verification of candidate genes identified by GWAS (Li et al., 2024). In this study, we used qRT-PCR to preliminarily verify the association of five selected genes *WRKY39* (*Os02g0265200*), *RGP-2* (*Os01g0974701*), *PIP5Ks* (*Os08g0450800*), *MADS59* (*Os06g0347700*) and *ABCF4* (*Os08g0564100*) between the two accessions: g217 with high tocopherol and g95 with low tocopherol, on three different seed developmental stages (milky, doughy, maturity) of rice grain in the context of transcript analysis. Interestingly, we found that *WRKY39*, *PIP5Ks*, and *MADS59* displayed significant differences in the low and high tocopherol rice accessions.

The present study unveiled a rich source of genetic elements, including SNPs, QTLs, and putative candidate genes associated with tocopherol biosynthesis in rice. We validated three related tocopherol biosynthetic genes, including one signaling gene (*PIP5Ks*) and two TF genes (*WRKY39* and *MADS59*) that may be responsible for tocopherol variation in the natural rice population. In the future, the identified genes and QTLs can be used to develop molecular markers for marker-assisted selection (MAS), facilitating the selection of desirable traits in breeding programs. Additionally, multiple genes or QTLs controlling tocopherol content can be pyramided into a single genetic background to achieve synergistic effects and develop rice varieties with enhanced tocopherol levels. Advanced technologies such as CRISPR/Cas9 can also be employed to precisely edit the identified genes, creating targeted mutations to improve tocopherol content in rice. Last but not least, some of the identified genes can be implemented in the metabolic engineering of microorganisms for high-throughput production of tocopherols in an industrial manner.

## Conclusions

This study conducted GWAS analysis on a mapping population consisting of 179 rice accessions. We investigated regulatory genes (TFs, miPs, and transposons) and other genes functional in the transport and signaling of tocopherol biosynthesis in this natural rice population. Finally, three important candidate genes *WRKY39* (*Os02g0265200*), *PIP5Ks* (*Os08g0450800*), and *MADS59* (*Os06g0347700*), were identified and validated by qRT-PCR analysis. This study provides fundamental information for breeding new rice varieties with high tocopherol content and high-quality oils. Further functional studies, such as overexpression or knockout of these genes, will help us better understand the molecular mechanisms of tocopherol biosynthesis in rice.

## Data availability statement

All relevant data can be found within the manuscript and its supporting materials and further inquiries can be directed to the corresponding author/s.

## Author contributions

SK: Writing – original draft, Formal analysis, Investigation. NF: Supervision, Project administration, Formal analysis, Conceptualization, Writing – review & editing. AA: Supervision, Investigation, Formal analysis, Writing – review & editing. KT: Formal analysis, Writing – review & editing. AG: Writing – review & editing, Resources. HA: Writing – review & editing, Supervision, Resources. PI: Supervision, Writing – review & editing, Resources.

## References

- Akhtar, A. A., and Turner, D. P. (2022). The role of bacterial ATP-binding cassette (ABC) transporters in pathogenesis and virulence: Therapeutic and vaccine potential. *Microbial. Pathogene.* 171, 105734. doi: 10.1016/j.micpath.2022.105734
- Albert, E., Kim, S., Magallanes-Lundback, M., Bao, Y., Deason, N., Danilo, B., et al. (2022). Genome-wide association identifies a missing hydrolase for tocopherol synthesis in plants. *Proc. Natl. Acad. Sci.* 119, e2113488119. doi: 10.1073/pnas.2113488119
- Ali, E., Hussain, S., Hussain, N., Kakar, K. U., Shah, J. M., Zaidi, S. H. R., et al. (2022). Tocopherol as plant protector: An overview of Tocopherol biosynthesis enzymes and their role as antioxidant and signaling molecules. *Acta Physiol. Plant.* 44, 20. doi: 10.1007/s11738-021-03350-x
- Alonso-Blanco, C., Andrade, J., Becker, C., Bemm, F., Bergelson, J., Borgwardt, K. M., et al. (2016). 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell.* 166, 481–491. doi: 10.1016/j.cell.2016.05.063
- Bao, Y. M., Sun, S. J., Li, M., Li, L., Cao, W. L., Luo, J., et al. (2012). Overexpression of the Qc-SNARE gene *OsSYP71* enhances tolerance to oxidative stress and resistance to rice blast in rice (*Oryza sativa* L.). *Gene.* 504, 238–244. doi: 10.1016/j.gene.2012.05.011
- Baseggio, M., Murray, M., Magallanes-Lundback, M., Kaczmar, N., Chamness, J., Buckler, E. S., et al. (2019). Genome-wide association and genomic prediction models of tocopherols in fresh sweet corn kernels. *Plant Genome.* 12, 180038. doi: 10.3835/plantgenome2018.06.0038
- Bermúdez, L., Del Pozo, T., Silvestre Lira, B., de Godoy, F., Boos, I., Romano, C., et al. (2018). A tomato tocopherol-binding protein sheds light on intracellular  $\alpha$ -tocopherol metabolism in plants. *Plant Cell Physiol.* 59, 2188–2203. doi: 10.1093/pcp/pcy191
- 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
- Cerise, M., Giaume, F., Galli, M., Khahani, B., Lucas, J., Podico, F., et al. (2021). *OsFD4* promotes the rice floral transition via florigen activation complex formation in the shoot apical meristem. *New Phytol.* 229, 429–443. doi: 10.1111/nph.v229.1
- Chen, C., Yang, J., Tong, H., Li, T., Wang, L., and Chen, H. (2019). Genome-wide analysis of fatty acid desaturase genes in rice (*Oryza sativa* L.). *Sci. Rep.* 9, 1–11. doi: 10.1038/s41598-019-55648-z
- Chen, J., Nolan, T. M., Ye, H., Zhang, M., Tong, H., Xin, P., et al. (2017). Arabidopsis *WRKY46*, *WRKY54*, and *WRKY70* transcription factors are involved in brassinosteroid-regulated plant growth and drought responses. *Plant Cell.* 29, 1425–1439. doi: 10.1105/tpc.17.00364
- Chen, W., Gao, Y., Xie, W., Gong, L., Lu, K., Wang, W., et al. (2014). Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat. Genet.* 46, 714–721. doi: 10.1038/ng.3007
- Chu, D., Zhang, Z., Hu, Y., Fang, C., Xu, X., Yuan, J., et al. (2023). Genome-wide scan for oil quality reveals a coregulation mechanism of tocopherols and fatty acids in soybean seeds. *Plant Commun.* 4, 100598. doi: 10.1016/j.xplc.2023.100598
- Courtois, B., Audebert, A., Dardou, A., Roques, S., Ghneim-Herrera, T., Droc, G., et al. (2013). Genome-wide association mapping of root traits in a japonica rice panel. *PLoS One.* 8, e78037. doi: 10.1371/journal.pone.0078037

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

## Acknowledgments

The Chemistry & Chemical Engineering Research Center of Iran (CCERCI) is acknowledged for providing facilities and laboratory space.

## Conflict of interest

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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2024.1426321/full#supplementary-material>

- Dal Cin, V., Tieman, D. M., Tohge, T., McQuinn, R., de Vos, R. C., Osorio, S., et al. (2011). Identification of genes in the phenylalanine metabolic pathway by ectopic expression of a MYB transcription factor in tomato fruit. *Plant Cell*. 23, 2738–2753. doi: 10.1105/tpc.111.086975
- DellaPenna, D., and Mène-Saffrané, L. (2011). Vitamin E<sup>3</sup> in biosynthesis of vitamins in plants part B: vitamins B6, B8, B9, C, E, K. *Adv. Bot. Res.* 59, 179–227. doi: 10.1016/B978-0-12-385853-5.00002-7
- Diepenbrock, C. H., Kandianis, C. B., Lipka, A. E., Magallanes-Lundback, M., Vaillancourt, B., Góngora-Castillo, E., et al. (2017). Novel loci underlie natural variation in vitamin E levels in maize grain. *Plant Cell*. 29, 2374–2392. doi: 10.1105/tpc.17.00475
- Dong, Q., Zhang, Y. X., Zhou, Q., Liu, Q. E., Chen, D. B., Wang, H., et al. (2019). UMP kinase regulates chloroplast development and cold response in rice. *Int. J. Mol. Sci.* 20, 2107. doi: 10.3390/ijms20092107
- Doughman, R., Firestone, A., and Anderson, R. (2003). Phosphatidylinositol phosphate kinases put P14, 5P 2 in its place. *J. Membrane. Biol.* 194, 77–89. doi: 10.1007/s00232-003-2027-7
- Enfissi, E. M., Barneche, F., Ahmed, I., Lichtlé, C., Gerrish, C., McQuinn, R. P., et al. (2010). Integrative transcript and metabolite analysis of nutritionally enhanced *DE-ETIOLATED1* downregulated tomato fruit. *Plant Cell*. 22, 1190–1215. doi: 10.1105/tpc.110.073866
- Falk, J., and Munné-Bosch, S. (2010). Tocochromanol functions in plants: antioxidant and beyond. *J. Exp. Bot.* 61, 1549–1566. doi: 10.1093/jxb/erq030
- Fan, J., Bai, P., Ning, Y., Wang, J., Shi, X., Xiong, Y., et al. (2018). The monooxygenase receptor-like kinase SDS2 controls cell death and immunity in rice. *Cell Host Microbe*. 23, 498–510. e495. doi: 10.1016/j.chom.2018.03.003
- Federer, W. T., and Raghavarao, D. (1975). On augmented designs. *Biometrics*. 31, 29–35. doi: 10.2307/2529707
- Galwey, N. W. (2023). A Q-Q plot aids interpretation of the false discovery rate. *Biometrical J.* 65, e2100309. doi: 10.1002/bimj.202100309
- Ge, S. X., Jung, D., and Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*. 36, 2628–2629. doi: 10.1093/bioinformatics/btz931
- Gourcilleau, D., Lenne, C., Armenise, C., Moulia, B., Julien, J. L., Bronner, G., et al. (2011). Phylogenetic study of plant Q-type C2H2 zinc finger proteins and expression analysis of poplar genes in response to osmotic, cold and mechanical stresses. *DNA Res.* 18, 77–92. doi: 10.1093/dnares/dsr001
- Graebner, R. C., Wise, M., Cuesta-Marcos, A., Geniza, M., Blake, T., Blake, V. C., et al. (2015). Quantitative trait loci associated with the tocochromanol (vitamin E) pathway in barley. *PLoS One*. 10, e0133767. doi: 10.1371/journal.pone.0133767
- Grusak, M. A., and DellaPenna, D. (1999). Improving the nutrient composition of plants to enhance human nutrition and health. *Annu. Rev. Plant Biol.* 50, 133–161. doi: 10.1146/annurev.arplant.50.1.133
- Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize<sup>3</sup> implements and enhances circular visualization in R. *Bioinformatics*. 30, 2811–2812. doi: 10.1093/bioinformatics/btu393
- Guo, Z., Yang, W., Chang, Y., Ma, X., Tu, H., Xiong, F., et al. (2018). Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. *Mol. Plant*. 11, 789–805. doi: 10.1016/j.molp.2018.03.018
- Gupta, B. B., Selter, L. L., Baranwal, V. K., Arora, D., Mishra, S. K., Sirohi, P., et al. (2019). Updated inventory, evolutionary and expression analyses of G (PDR) type ABC transporter genes of rice. *Plant Physiol. Biochem.* 142, 429–439. doi: 10.1016/j.plaphy.2019.08.004
- Haddadi, P., Ebrahimi, A., Langlade, N., Yazdi-Samadi, B., Berger, M., Calmon, A., et al. (2012). Genetic dissection of tocopherol and phytosterol in recombinant inbred lines of sunflower through quantitative trait locus analysis and the candidate gene approach. *Mol. Breed.* 29, 717–729. doi: 10.1007/s11032-011-9585-7
- Hagr, T., and Adam, I. (2020). Phytochemical analysis, antibacterial and antioxidant activities of essential Oil from *hibiscus sabdariffa* (L) Seeds, (Sudanese Karkadi). *Prog. Chem. Biochem. Res.* 3, 194–201.
- Havaux, M., Eymery, F., Porfiriova, S., Rey, P., and Dörmann, P. (2005). Vitamin E protects against photoinhibition and photooxidative stress in *Arabidopsis thaliana*. *Plant Cell*. 17, 3451–3469. doi: 10.1105/tpc.105.037036
- Hirochika, H. (2001). Contribution of the *Tos17* retrotransposon to rice functional genomics. *Curr. Opin. Plant Biol.* 4, 118–122. doi: 10.1016/S1369-5266(00)00146-1
- Hirochika, H., Sugimoto, K., Otsuki, Y., Tsugawa, H., and Kanda, M. (1996). Retrotransposons of rice involved in mutations induced by tissue culture. *Proc. Natl. Acad. Sci.* 93, 7783–7788. doi: 10.1073/pnas.93.15.7783
- Hossain, M. R., Bassel, G. W., and Pritchard, J. (2016). Trait specific expression profiling of salt stress responsive genes in diverse rice genotypes as determined by modified significance analysis of microarrays. *Front. Plant Sci.* 7, 183874. doi: 10.3389/fpls.2016.00567
- Hu, W., Wells, J. H., Shin, T.-S., and Godber, J. S. (1996). Comparison of isopropanol and hexane for extraction of vitamin E and oryzanol from stabilized rice bran. *J. Am. Oil Chemists' Soc.* 73, 1653–1656. doi: 10.1007/BF02517967
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42, 961–967. doi: 10.1038/ng.695
- Huang, Y. W., Tsay, W. S., Chen, C. C., Lin, C. W., and Huang, H. J. (2008). Increased expression of the rice C-type cyclin-dependent protein kinase gene, *Orysa; CDK1*, in response to salt stress. *Plant Physiol. Biochem.* 46, 71–81. doi: 10.1016/j.plaphy.2007.10.013
- Jahn, R., and Scheller, R. H. (2006). SNAREs—engines for membrane fusion. *Nat. Rev. Mol. Cell Biol.* 7, 631–643. doi: 10.1038/nrm2002
- Kato, T., and Horibata, A. (2022). QTL-by-QTL, QTL-by-environment, and QTL-by-QTL-by-environment interactions of loci controlling grain length in rice. *Euphytica*. 218, 32. doi: 10.1007/s10681-021-02962-1
- Kitsios, G., Alexiou, K. G., Bush, M., Shaw, P., and Doonan, J. H. (2008). A cyclin-dependent protein kinase, CDKC2, colocalizes with and modulates the distribution of spliceosomal components in Arabidopsis. *Plant J.* 54, 220–235. doi: 10.1111/j.1365-3113.2008.03414.x
- Lee, J. S., Kwak, J., and Hay, F. R. (2020). Genetic markers associated with seed longevity and vitamin E in diverse aus rice varieties. *Seed. Sci. Res.* 30, 133–141. doi: 10.1017/S0960258520000173
- Li, C., Song, Y., Zhu, Y., Cao, M., Han, X., Fan, J., et al. (2024). GWAS analysis reveals candidate genes associated with dense tolerance (ear leaf structure) in maize (*Zea mays* L.). *J. Integr. Agric.* In Press. doi: 10.1016/j.jia.2024.01.023
- Li, P., Guo, M., Wang, C., Liu, X., and Zou, Q. (2015). An overview of SNP interactions in genome-wide association studies. *Briefings Funct. Genomics.* 14, 143–155. doi: 10.1093/bfgp/elu036
- Li, S., Zhou, X., Chen, L., Huang, W., and Yu, D. (2010). Functional characterization of *Arabidopsis thaliana* WRKY39 in heat stress. *Mol. Cells*. 29, 475–483. doi: 10.1007/s10059-010-0059-2
- Li, Z., Ao, Y., Feng, D., Liu, J., Wang, J., Wang, H. B., et al. (2017). *OsRLCK57*, *OsRLCK107* and *OsRLCK118* positively regulate chitin- and PGN-induced immunity in rice. *Rice* 10, 1–5. doi: 10.1186/s12284-017-0145-6
- Li, Z., and Yao, Q. (2022). Rice carotenoid biofortification and yield improvement conferred by endosperm-specific overexpression of *OsGLK1*. *Front. In. Plant Sci.* 13, 951605. doi: 10.3389/fpls.2022.951605
- Liang, X., and Zhou, J.-M. (2018). Receptor-like cytoplasmic kinases: central players in plant receptor kinase-mediated signaling. *Annu. Rev. Plant Biol.* 69, 267–299. doi: 10.1146/annurev-arplant-042817-040540
- Lipka, A. E., Kandianis, C. B., Hudson, M. E., Yu, J., Drnevich, J., Bradbury, P. J., et al. (2015). From association to prediction: statistical methods for the dissection and selection of complex traits in plants. *Curr. Opin. Plant Biol.* 24, 110–118. doi: 10.1016/j.pbi.2015.02.010
- Lipka, A. E., Gore, M. A., Magallanes-Lundback, M., Mesberg, A., Lin, H., Tiede, T., et al. (2013). Genome-wide association study and pathway-level analysis of tocochromanol levels in maize grain. *G3*. 3, 1287–99. doi: 10.1534/g3.113.006148
- Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 12, e1005767. doi: 10.1371/journal.pgen.1005767
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-ΔΔCT</sup> method. *Methods*. 25, 402–408. doi: 10.1006/meth.2001.1262
- Mackay, T. F., Stone, E. A., and Ayroles, J. F. (2009). The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* 10, 565–577. doi: 10.1038/nrg2612
- Maghraby, A., and Alzalaty, M. (2024). Genome-wide identification, characterization and evolutionary analysis of the APETALA 2/ethylene responsive element-binding protein (*AP2/EREBP*), cytochrome c oxidase (*COX*) and lipid transfer protein (*LTP*) genes in *Zea mays* under drought stress. *Sci. Rep.* 14, 7610. doi: 10.1038/s41598-024-57376-5
- Mather, K. A., Caicedo, A. L., Polato, N. R., Olsen, K. M., McCouch, S., and Purugganan, M. D. (2007). The extent of linkage disequilibrium in rice (*Oryza sativa* L.). *Genetics* 177, 2223–2232. doi: 10.1534/genetics.107.079616
- Mène-Saffrané, L. (2017). Vitamin E biosynthesis and its regulation in plants. *Antioxidants*. 7, 2. doi: 10.3390/antiox7010002
- Mène-Saffrané, L., Jones, A. D., and DellaPenna, D. (2010). Plastochromanol-8 and tocopherols are essential lipid-soluble antioxidants during seed desiccation and quiescence in Arabidopsis. *Proc. Natl. Acad. Sci.* 107, 17815–17820. doi: 10.1073/pnas.1006971107
- Mittal, S., Mallikarjuna, M. G., Rao, A. R., Jain, P. A., Dash, P. K., and Thirunavukkarasu, N. (2017). Comparative analysis of CDPK family in maize, Arabidopsis, rice, and sorghum revealed potential targets for drought tolerance improvement. *Front. Chem.* 5, 115. doi: 10.3389/fchem.2017.00115
- Morand-Laffargue, L., Vairo, D., Halimi, C., Chiarello, E., Creton, B., Sabatier, D., et al. (2023). Ability of black soldier fly larvae to bioaccumulate tocopherols from different substrates and measurement of larval tocopherol bioavailability *in vitro*. *J. Insects. as Food Feed.* 1, 1–13. doi: 10.1163/23524588-20230183B
- Morelli, L., García Romañach, L., Glauser, G., Shanmugabala, V., Kessler, F., and Rodriguez-Conceptcion, M. (2023). Nutritional enrichment of plant leaves by combining genes promoting tocopherol biosynthesis and storage. *Metabolites*. 13, 193. doi: 10.3390/metabo13020193
- Muñoz, P., and Munne-Bosch, S. (2019). Vitamin E in plants: biosynthesis, transport, and function. *Trends In. Plant Sci.* 24, 1040–1051. doi: 10.1016/j.tplants.2019.08.006



- Ning, L., Wang, S., Ju, P., Bai, X., Ge, L., Qi, X., et al. (2018). Rice overexpression of millet *SiANT1* gene increases salt tolerance. *Sci. Agricult. Sin.* 51, 1830–1841. doi: 10.3864/j.issn.0578-1752.2018.10.002
- Niu, Y., Wang, J., and Bao, Y. (2022). Vitamin E synthesis and response in plants. *Front. In. Plant Sci.* 13, 994058. doi: 10.3389/fpls.2022.994058
- Pachamuthu, K., and Borges, F. (2023). Epigenetic control of transposons during plant reproduction: From meiosis to hybrid seeds. *Curr. Opin. Plant Biol.* 75, 102419. doi: 10.1016/j.pbi.2023.102419
- Peng, Y. Y., Liao, L. L., Liu, S., Nie, M. M., Li, J., Zhang, L. D., et al. (2019). Magnesium deficiency triggers SGR-mediated chlorophyll degradation for magnesium remobilization. *Plant Physiol.* 181, 262–275. doi: 10.1104/pp.19.00610
- Phillips, P. C. (2008). Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* 9, 855–867. doi: 10.1038/nrg2452
- Phukan, U. J., and Shukla, R. K. (2016). WRKY transcription factors: molecular regulation and stress responses in plants. *Front. Plant Sci.* 7, 192140. doi: 10.3389/fpls.2016.00760
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904–909. doi: 10.1038/ng1847
- Rathore, A., Martinez, T. F., Chu, Q., and Saghatelian, A. (2018). Small, but mighty? Searching for human microproteins and their potential for understanding health and disease. *Expert Rev. Proteomics.* 15, 963–965. doi: 10.1080/14789450.2018.1547194
- Romer, J., Gutbrod, K., Schuppener, A., Melzer, M., Müller-Schüssele, S. J., Meyer, A. J., et al. (2024). Tocopherol and phyloquinone biosynthesis in chloroplasts require the *phytol kinase* VITAMIN E PATHWAY GENE5 (VTE5) and the *farnesol kinase* (FOLK). *Plant Cell.* 36, 1140–1158. doi: 10.1093/plcell/koad316
- Sadre, R., Frentzen, M., Saeed, M., and Hawkes, T. (2010). Catalytic reactions of the homogentisate prenyl transferase involved in plastoquinone-9 biosynthesis. *J. Biol. Chem.* 285, 18191–18198. doi: 10.1074/jbc.M110.117929
- Sattler, S. E., Gilliland, L. U., Magallanes-Lundback, M., Pollard, M., and DellaPenna, D. (2004). Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. *Plant Cell.* 16, 1419–1432. doi: 10.1105/tpc.021360
- Schuy, C., Groth, J., Ammon, A., Eydam, J., Baier, S., Schweizer, G., et al. (2019). Deciphering the genetic basis for vitamin E accumulation in leaves and grains of different barley accessions. *Sci. Rep.* 9, 9470. doi: 10.1038/s41598-019-45572-7
- Segura, V., Vilhjálmsson, B. J., Platt, A., Korte, A., Seren, Ü., Long, Q., et al. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat. Genet.* 44, 825–830. doi: 10.1038/ng.2314
- Sen, C. K., Khanna, S., and Roy, S. (2007). Tocotrienols in health and disease: the other half of the natural vitamin E family. *Mol. Aspects. Med.* 28, 692–728. doi: 10.1016/j.mam.2007.03.001
- Shammugasamy, B., Ramakrishnan, Y., Manan, F., and Muhammad, K. (2015). Rapid reversed-phase chromatographic method for determination of eight vitamin E isomers and  $\gamma$ -oryzanol in rice bran and rice bran oil. *Food Analytical. Methods* 8, 649–655. doi: 10.1007/s12161-014-9929-0
- Shaw, E. J., Kakuda, Y., and Rajcan, I. (2016). Effect of genotype, environment, and genotype  $\times$  environment interaction on tocopherol accumulation in soybean seed. *Crop Sci.* 56, 40–50. doi: 10.2135/croscis2015.02.0069
- Shaw, E. J., and Rajcan, I. (2017). Molecular mapping of soybean seed tocopherols in the cross 'OAC Bayfield'  $\times$  'OAC Shire'. *Plant Breed.* 136, 83–93. doi: 10.1111/pbr.2017.136.issue-1
- Shinya, T., Yasuda, S., Hyodo, K., Tani, R., Hojo, Y., Fujiwara, Y., et al. (2018). Integration of danger peptide signals with herbivore-associated molecular pattern signaling amplifies anti-herbivore defense responses in rice. *Plant J.* 94, 626–637. doi: 10.1111/tpj.2018.94.issue-4
- Song, W. Y., Wang, G. L., Chen, L. L., Kim, H. S., Pi, L. Y., Holsten, T., et al. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science.* 270, 1804–1806. doi: 10.1126/science.270.5243.1804
- Sozen, E., Demirel, T., and Ozer, N. K. (2019). Vitamin E: Regulatory role in the cardiovascular system. *IUBMB Life* 71, 507–515. doi: 10.1002/iub.v71.4
- Stacey, M. G., Cahoon, R. E., Nguyen, H. T., Cui, Y., Sato, S., Nguyen, C. T., et al. (2016). Identification of homogentisate dioxygenase as a target for vitamin E biofortification in oil seeds. *Plant Physiol.* 172, 1506–1518. doi: 10.1104/pp.16.00941
- Straub, D., and Wenkel, S. (2017). Cross-species genome-wide identification of evolutionary conserved MicroProteins. *Genome Biol. Evol.* 9, 777–789. doi: 10.1093/gbe/evx041
- Sugasawa, T., Lenzen, G., Simon, S., Hidaka, J., Cahen, A., Guillaume, J. L., et al. (2001). The iodocyanopindolol and SM-11044 binding protein belongs to the TM9SF multispreading membrane protein superfamily. *Gene* 273, 227–237. doi: 10.1016/S0378-1119(01)00587-X
- Sun, M., Zhu, Z. P., Yu, J. X., Wu, K. X., Guo, Y. X., Shen, M., et al. (2023). Transcriptomic and physiological analysis reveal phytohormone and phenylpropanoid biosynthesis in root of *Cynanchum auriculatum*. *Plant Growth Regul.* 101, 67–85. doi: 10.1007/s10725-022-00953-3
- Takahara, K., Kasajima, I., Takahashi, H., Hashida, S.-n., Itami, T., Onodera, H., et al. (2010). Metabolome and photochemical analysis of rice plants overexpressing Arabidopsis NAD kinase gene. *Plant Physiol.* 152, 1863–1873. doi: 10.1104/pp.110.153098
- Tian, L., DellaPenna, D., and Dixon, R. A. (2007). The *pds2* mutation is a lesion in the Arabidopsis homogentisate solanesyltransferase gene involved in plastoquinone biosynthesis. *Planta* 226, 1067–1073. doi: 10.1007/s00425-007-0564-5
- Traber, M. G., and Atkinson, J. (2007). Vitamin E, antioxidant and nothing more. *Free Radical Biol. Med.* 43, 4–15. doi: 10.1016/j.freeradbiomed.2007.03.024
- Usadel, B., Poree, F., Nagel, A., Lohse, M., Czedik-Eysenberg, A., and Stitt, M. (2009). A guide to using MapMan to visualize and compare Omics data in plants: a case study in the crop species, Maize. *Plant. Cell Environ.* 32, 1211–1229. doi: 10.1111/j.1365-3040.2009.01978.x
- Valentin, H. E., Lincoln, K., Moshiri, F., Jensen, P. K., Qi, Q., Venkatesh, T. V., et al. (2006). The Arabidopsis vitamin E pathway gene5-1 mutant reveals a critical role for phytol kinase in seed tocopherol biosynthesis. *Plant Cell* 18, 212–224. doi: 10.1105/tpc.105.037077
- Van Eenennaam, A. L., Li, G., Venkatramesh, M., Levering, C., Gong, X., Jamieson, A. C., et al. (2004). Elevation of seed  $\alpha$ -tocopherol levels using plant-based transcription factors targeted to an endogenous locus. *Metab. Eng.* 6, 101–108. doi: 10.1016/j.ymben.2003.11.001
- Varshney, R. K., and Kumar, A. (2019). Epistatic interactions of major effect drought QTLs with genetic background loci determine grain yield of rice under drought stress. *Sci. Rep.* 9, 2616. doi: 10.1038/s41598-019-39084-7
- Vom Dorp, K., Hölzl, G., Plohm, C., Eisenhut, M., Abraham, M., Weber, A. P., et al. (2015). Remobilization of phytol from chlorophyll degradation is essential for tocopherol synthesis and growth of Arabidopsis. *Plant Cell* 27, 2846–2859. doi: 10.1105/tpc.15.00395
- Wang, B., Wang, Y., Yu, W., Wang, L., Lan, Q., Wang, Y., et al. (2022). Knocking out the transcription factor OsNAC092 promoted rice drought tolerance. *Biology* 11, 1830. doi: 10.3390/biology11121830
- Wang, Q., Xie, W., Xing, H., Yan, J., Meng, X., Li, X., et al. (2015). Genetic architecture of natural variation in rice chlorophyll content revealed by a genome-wide association study. *Mol. Plant* 8, 946–957. doi: 10.1016/j.molp.2015.02.014
- Wang, H., Xu, S., Fan, Y., Liu, N., Zhan, W., Liu, H., et al. (2018). Beyond pathways: genetic dissection of tocopherol content in maize kernels by combining linkage and association analyses. *Plant Biotechnol. J.* 16, 1464–1475. doi: 10.1111/pbi.2018.16.issue-8
- Wang, L., Zheng, J., Luo, Y., Xu, T., Zhang, Q., Zhang, L., et al. (2013). Construction of a genome-wide RNA i mutant library in rice. *Plant Biotechnol. J.* 11, 997–1005. doi: 10.1111/pbi.2013.11.issue-8
- Wenkel, S., Emery, J., Hou, B.-H., Evans, M. M., and Barton, M. (2007). A feedback regulatory module formed by LITTLE ZIPPER and HD-ZIPIII genes. *Plant Cell* 19, 3379–3390. doi: 10.1105/tpc.107.055772
- Xiao, X., Wang, R., Guo, W., Khaskhali, S., Fan, R., Zhao, R., et al. (2022). The receptor-like cytoplasmic kinase OsRLCK118 regulates plant development and basal immunity in rice (*Oryza sativa* L.). *Trop. Plants* 1, 1–10. doi: 10.48130/TP-2022-0004
- Xu, Y., Zhang, S., Guo, H., Wang, S., Xu, L., Li, C., et al. (2014). Os ABCB 14 functions in auxin transport and iron homeostasis in rice (*Oryza sativa* L.). *Plant J.* 79, 106–117. doi: 10.1111/tpj.2014.79.issue-1
- Yang, S., Zhou, J., Li, Y., Wu, J., Ma, C., Chen, Y., et al. (2023). AP2/EREBP pathway plays an important role in Chaling wild rice tolerance to cold stress. *Int. J. Mol. Sci.* 24, 14441. doi: 10.3390/ijms241914441
- Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., et al. (2021). rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics. Proteomics Bioinf.* 19, 619–628. doi: 10.1016/j.gpb.2020.10.007
- Yu, K., Miao, H., Zhao, X., and Han, Y. (2022). Genome-wide association studies reveal novel QTLs, QTL-by-environment interactions and their candidate genes for tocopherol content in soybean seed. *Front. Plant Sci.* 13, 1026581. doi: 10.3389/fpls.2022.1026581
- Zhan, W., Liu, J., Pan, Q., Wang, H., Yan, S., Li, K., et al. (2019). An allele of *Zm PORB 2* encoding a protochlorophyllide oxidoreductase promotes tocopherol accumulation in both leaves and kernels of maize. *Plant J.* 100, 114–127. doi: 10.1111/tpj.2019.100.issue-1
- Zhang, F., Boerwinkle, E., and Xiong, M. (2014). Epistasis analysis for quantitative traits by functional regression model. *Genome Res.* 24, 989–998. doi: 10.1101/gr.161760.113
- Zhang, R., Jia, G., and Diao, X. (2023). geneHapR: an R package for gene haplotypic statistics and visualization. *BMC Bioinf.* 24, 199. doi: 10.1186/s12859-023-05318-9
- Zhang, Z., Li, Y., Huang, K., Xu, W., Zhang, C., and Yuan, H. (2020). Genome-wide systematic characterization and expression analysis of the phosphatidylinositol 4-phosphate 5-kinases in plants. *Gene.* 756, 144915. doi: 10.1016/j.gene.2020.144915
- Zhang, T., Zhao, X., Wang, W., Pan, Y., Huang, L., Liu, X., et al. (2012). Comparative transcriptome profiling of chilling stress responsiveness in two contrasting rice genotypes. *PLoS One.* 7, e43274. doi: 10.1371/journal.pone.0043274



Zhao, K., Tung, C. W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., et al. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2, 467. doi: 10.1038/ncomms1467

Zheng, S., Dong, J., Lu, J., Li, J., Jiang, D., Yu, H., et al. (2022). A cytosolic pentatricopeptide repeat protein is essential for tapetal plastid development by regulating *OsGLK1* transcript levels in rice. *New Phytol.* 234, 1678–1695. doi: 10.1111/nph.v234.5

Zhou, H., Zhang, F., Zhai, F., Su, Y., Zhou, Y., Ge, Z., et al. (2022). Rice GLUTATHIONE PEROXIDASE1-mediated oxidation of bZIP68 positively regulates ABA-independent osmotic stress signaling. *Mol. Plant.* 15, 651–670. doi: 10.1016/j.molp.2021.11.006

Zingg, J. M. (2007). Vitamin E: an overview of major research directions. *Mol. Aspects. Med.* 28, 400–422. doi: 10.1016/j.mam.2007.05.004