Food Chemistry: Molecular Sciences 11 (2025) 100313

Contents lists available at ScienceDirect

Food Chemistry: Molecular Sciences

journal homepage: www.sciencedirect.com/journal/food-chemistry-molecular-sciences

Genome-wide association studies for identification of QTLs and key candidate genes to improve grain quality in rice (*Oryza sativa* L.)

Supriya Sachdeva [a], Rakesh Singh [a,*], Harshita Singh [a], Rakesh Bharadwaj [b], Antil Jain [b], Vikas K. Singh [c], Uma Maheshwar Singh [d], Arvind Kumar [e], Gyanendra Pratap Singh [a]

[a] Division of Genomic Resources, ICAR-NBPGR, Pusa, New Delhi, India
[b] Division of Germplasm Evaluation, ICAR-NBPGR, Pusa, New Delhi, India
[c] International Rice Research Institute (IRRI), South Asia Hub, ICRISAT, Hyderabad, India
[d] International Rice Research Institute (IRRI), South Asia Regional Centre (ISARC), Varanasi, India
[e] International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Telangana, India

ARTICLE INFO

Keywords:
Rice
GWAS
Grain quality
QTN
GO enrichment
Candidate gene

ABSTRACT

Grain quality is the key concern for rice breeders and is paramount to consumer acceptability. We characterized a diverse subset of 198 rice accessions of 3 K Rice Genome Project (RGP) for grain quality attributes, specifically glycemic index %, total dietary fibre, oil %, protein, amylose, moisture %, phytate, phenol, and starch content. A set of 5,53,229 single nucleotide polymorphism (SNP) markers obtained from the 3 K RG 1 M filtered SNP dataset used for genome wide association studies (GWAS). Consequently, we discovered 200 Quantitative trait nucleotides (QTNs) associated with the traits mentioned above distributed across the genome. These QTNs were grouped into 26 Quantitative Trait Loci (QTL) clusters, of which 20 clusters validated with at least three GWAS methods were considered reliable. Furthermore, 869 putative candidate genes were identified, many of which overlapped between quality traits. Integrating the GWAS, RNA-seq and qRT-PCR results, we finally identified two important genes (*LOC_Os11g303700* and *LOC_Os11g30500*) associated with rice quality, and they may affect the grain quality by regulating the textural properties, appearance and eating quality. The findings of our study highlighted the role of molecular machinery in future rice breeding.

1. Introduction

Rice is the most commonly cultivated staple food, feeding nearly 4 billion people across the globe (Ngo et al., 2023; Yoshida et al., 2023; Jukanti et al., 2025). With the improving living standards, rice trade has become ever more globalised, and the demand for healthier rice is rising (Yang et al., 2023). It is the key source of nutrition and energy, especially for people with limited access to diverse diets (Birla et al., 2017; Song et al., 2019). Whole grain rice stores a moderately lower amount of essential micronutrients compared to other staple crops such as wheat, maize, pulses, and tuber crops (Mahender et al., 2016). Rice grains, though, contain 75–78 % starch, 10–15 % protein, and less than 1 % fat, with lipids, minerals, vitamins, antioxidants, and dietary fibre in traces (Jukanti et al., 2025; Bharali et al., 2025; Butardo Jr & Sreenivasulu, 2016).

Several efforts have been undertaken to identify the genes and QTLs associated with rice grain quality (Lou et al., 2023), namely GBSSI and SS2a, controlling genetic variations in starch attributes, and OsGlu associated with grain protein content, impacting the nutritional value and eating quality of rice (Biselli et al., 2015; **Praphasanobol et al., 2023;** Yang et al., 2019; Yoshida et al., 2023). Earlier studies imply that certain genes dominantly govern rice quality. Rice quality, particularly the appearance of the grain, its milling and cooking quality, determines the preferences of the consumers and marketability (Kabange et al., 2023). Breeding programmes aim for different types of rice varieties to match the optimal demands of the health-conscious consumers belonging to diverse groups. For instance, soft and sticky rice is well suited to Japanese and South Korean populations, however, non-sticky and fluffy rice is preferred in India and South America. Nevertheless, superior quality rice with exceptional appearance and nutritional characteristics is expected to be more favoured by consumers (Anacleto et al., 2019; Ito, 2023; Li et al., 2022; Verma et al., 2022).

Major carbohydrate in rice occurs in the form of starch (amylose and amylopectin) and based on the content of amylose rice has been categorized as waxy, low, very low, intermediate and high amylose types. Amylose is a quantitative inherited trait, genes/QTL clusters for which

* Corresponding author.

https://doi.org/10.1016/j.fochms.2025.100313
Received 7 July 2025; Received in revised form 3 October 2025; Accepted 7 October 2025
Available online 8 October 2025

have been mapped across the rice genome except chromosome 5 and 12. Although QTLs have been reported, only limited genes for amylose have been cloned (Wang et al., 2023). Waxy rice has a sticky nature, however, the high amylose containing rice are fluffy and with low glycemic index beneficial to rice consumers. Particularly, mutations in *SSIIa* and *SSIIb* (starch synthase enzymes) and SBEs (starch branching enzymes) were targeted in Waxy allele, *GBSSI* to generate high amylose rice (Anacleto et al., 2019). In contrast, non-starchy and soluble forms such as dietary fibre located mainly in the bran layers is affected by numerous genes and environmental factors, making it a difficult task to map with single QTLs. Genetic engineering could be utilized to improve dietary fibre in rice, though this is still an elusive approach.

Another nutritionally important rice component is grain protein which determines the textural attributes by avoiding water absorption and swelling of starch once cooked. It has been reported that rice with lower protein content have a comparatively better flavor than high protein containing rice. About, 80 QTLs have been mapped across the rice genome, including qPro-1, qPC1.2, qPC1, qAAC7.1, qPC6.2, qPro-2 (Islam et al., 2020). Low heritability, negative relationship with yield and environmental influence has been a concern for rice breeders. Therefore, mapping of QTLs governing high protein and pyramiding of the QTLs identified in several rice breeding efforts are of much importance to the Asian population. Also, rice grains contain anti-nutritional components viz., phytic acid and polyphenolic compounds in the bran act as chelating agent for pro-oxidant metals and free radical scavengers. Phytates QTLs and genes for phytate in rice have been identified on different chromosomes, with one QTL for phytate and inorganic phosphorus on chromosome 5. A gene encoding phosphatidylinositol 4-kinase homolog (*LOC_Os08g0274775*) has been proven to be potentially involved in phytate metabolism (Gyani et al., 2020). The levels of polyphenols vary in rice cultivars and are associated with the color of rice, specifically in red and black rice varieties, which are appreciated for their higher phenolic content. Quantitative Trait Loci (QTLs) for phenolic compounds in rice have been on chromosomes 2, 7, 8, and 11 (Jin et al., 2009). 11 significant MTAs have been detected for three QTLs for antioxidant traits in rice (Sanghamitra et al., 2022) that have been presumed to assist in rice antioxidant improvement programmes. Hence, increasing the nutritional value for functional rice is an important and sustainable approach to improve consumer's nutrition with multiple benefits on human health (Huang et al., 2023; Wang et al., 2023). However, such instances are unusual because rice quality is a complex trait encompassing several attributes, and many genes contribute to each specific trait. To date, identifying new genes governing grain quality in rice remains a challenge, and the detailed molecular mechanisms modulating the biosynthetic pathways of component quality traits are largely unknown; hence, it is imperative to discover genes that breeders might deploy in modern rice breeding. Genome-wide association study is an effective tool that has been prevalently used for mapping of QTNs underlying the rice quality traits and studying their complex genetics (Badoni et al., 2024; Singh et al., 2024; Mbanjo et al., 2023; Abhijith et al., 2022; Tibbs Cortes et al., 2021; Cruz et al., 2021; Mishra et al., 2019; Zhong et al., 2021; Qiu et al., 2021; Selvaraj et al., 2021; Yang et al., 2019; Sumargo et al., 2016). Multiple loci models detect QTN effects with much higher precision when compared with single locus methods (Wen et al., 2018; Tamba et al., 2017; Zhang et al., 2017; Wang et al., 2016). Conventional methods employed for the evaluation of complex traits are undesirable because they are labour-intensive and costly (Wu et al., 2019). However, a holistic effort assessing all the grain quality component traits in a robust, cost-effective, and non-destructive way will enable us to understand the interdependence between the quality traits comprehensively and develop premium rice varieties. More recently, prediction modelling based on near-infrared spectroscopy (NIRS) technique has been widely used for estimation of a wide variety of phytochemical properties, including total phenols, phytate content, total carotenoids, anthocyanins and biochemical traits such as moisture content, starch content, dietary fibre, glycemic index, proteins, lipids and vitamins in grains, plants, and animal products.in agriculture (Bagchi et al., 2016) owing to its capacity to analyze multiple samples simultaneously and environmental friendliness (Teye et al., 2019; Yang et al., 2024). NIRS models analyze the absorption of near-infrared light and predict these properties by correlating spectral data with chemical analysis, offering an effective alternative for quality assessments and research in areas of agriculture, food processing, and ecological investigations (Fodor et al., 2024; Pandiselvam et al., 2022; Grabska et al., 2022). This approach potentially holds promise as a nutritional analysis technique addressing the food security issues worldwide.

Enormous diversity in the germplasm plays a pivotal role in breeding elite rice varieties with consumer choices (Selvaraj et al., 2021; Singh et al., 2024). The 3 K RG association panel is acknowledged as a reservoir of important genes that influence grain yield and quality-related traits, providing opportunities for their genetic improvement (Abbai et al., 2019). In this context, we assessed a range of rice quality traits on a subset of the 3 K rice genome panel of 198 accessions received from IRRI, Varanasi, India. The accessions were subjected to single-locus and multi-locus GWAS methods to identify the significant QTNs associated with grain quality traits. In addition, we reported putative and novel loci responsible for rice quality using the In-silico approach through gene ontology, expression profiles and RT-PCR that can be utilized in quality improvement programmes.

2. Materials and methods

2.1. Plant material and field trial

A set of 198 accessions of rice was received from IRRI-SARC (South Asia Regional Centre), Varanasi, Uttar Pradesh, India used in the present study. These accessions represent a subset of re-sequenced 3000 rice genomes (https://doi.org/10.1186/2047-217X-3-7) collected across the globe from 89 different countries, inclusive of two check varieties, Pusa Basmati-1121 and Pusa Basmati-1 (Table S1). The plant materials were planted at ICAR-IARI (ICAR-Indian Agricultural Research Institute) farm located at 28.04°N, 77.12°E, New Delhi during kharif 2020 and kharif 2021 following an Augmented design with 7 blocks and 4 checks. About 21-day-old rice seedlings were sown in a nursery bed and manually transplanted in the main experimental field in 2 m^2 plots with a spacing of 20 cm between the rows and 15 cm between the plants. A suggested package of agronomic practices was used to ensure uniform growth of the transplanted rice crop under irrigated conditions.

2.2. Phenotyping of grain quality traits

The paddy samples at physiological maturity were harvested, cleaned, dried in the oven at 60 °C, and dehulled using a rice mill JGMJ 8098 model. About 25 g of dried seeds were grinded and homogenized further in a Foss Cyclotech™ laboratory mill, Denmark through a 1 mm screen to obtain the flour as per John et al., 2023 and Padhi et al., 2022 for further analysis of selected grain quality related traits i.e., percentage of glycemic index (PGI), total dietary fibre (TDF), oil %, protein (Prt), amylose (Amy), moisture % (Mst), phytate (Phy), phenols (Phn), and starch content (Str).

2.2.1. Estimation of PGI

For the determination of the glycemic index, the Megazyme D-Glucose (GOPOD, glucose oxidase/peroxidase) Assay Kit (KGLUC) was used. In brief, samples were milled and ground in a tissuelyser, weighed ~50 mg, and incubated with 500 µL of ethanol 80 % (*v*/v) at 80 °C for 5 min. The tubes were then centrifuged at 10,000 rpm for 5 min, and the supernatant was carefully poured. The centrifugation was again repeated, and the supernatant was then kept in a water bath at 60 °C for 15 min. 1 ml of distilled water was added to each of the tubes and stirred on a vortex mixer. 0.1 mL of the sample was accurately transferred to tubes, and 3 ml of GOPOD reagent was prepared and added to each tube.

Mixing of the sample and solution was done for each tube by pipetting. The solution mixtures were then kept in a water bath at 40 °C for 20 min, and absorbance was measured using a Benchtop Lab Systems Spectrometer at 510 nm. The estimated glucose was converted to % starch using the tool Mega-Calc™, available online (www.megazyme.com). PGI was estimated from an index of hydrolysis (HI) obtained by dividing the AUC, the area under the starch hydrolysis curve of the rice samples, by that of glucose used as a standard (Selvaraj et al., 2021).

2.2.2. Estimation of amylose

The amylose content of rice accessions was evaluated by the iodine binding method given by John et al., 2023 where the iodine complexes with the amylose. Some modifications were made for streamlining the handling procedure. Around 0.05 g of homogenized rice powder was transferred to a test tube and suspended in 500 µl ethanol solution. 4.5 ml of 1 N NaOH solution was then added to the test tube, and the test tubes were thoroughly vortexed. The test tubes were then kept in a water bath maintained at 100 °C for 20 min, after which the test tubes were taken out and brought back to room temperature. The final volume was transferred to a volumetric flask of 50 ml, and the volume was made up with distilled water. A 500 µl aliquot was taken from the flask and moved to amber coloured Falcon tubes. To the aliquots, 200 µl glacial acetic acid and 200 µl iodine solution were added, and the total volume was made up with distilled water to 10 ml. The tubes were left undisturbed in the dark to facilitate blue coloured complex formation. The intensity of the color was estimated spectrophotometrically at 620 nm by UV- VIS double beam PC scanning spectro model UVD-3200. Standards were prepared using already known potato amylose concentrations (A0512 Amylose from Sigma Aldrich).

2.2.3. NIR based estimation of other grain quality traits

All the 198 accessions were scanned for various eating quality traits by a non-destructive and rapid method using near-infrared reflectance spectroscopy (NIRS). Before scanning, the NIR instrument was calibrated with a reference cell, white mica, followed by sample estimations for the highest accuracy. Grounded sample is then transferred to a FOSS sample circular ring with a diameter of 3.8 cm and pressed without applying much pressure for uniform packing. 5 g of each rice flour sample was scanned 32 times within the NIR spectral absorbance range of 400-2498 nm with 2 nm intervals as the logarithm of respective reflectance(1/R) on the FOSS-NIRSDS3 spectrophotometer to obtain the reflectance spectra using the analytical software viz., WinISI v1.5 (John et al., 2022). The NIR results were then compared with reference materials (BCR-465, 466, and 467) to estimate selected eating quality traits (TDF, oil, Prot, Mst, Phy, Phn and Str) in rice.

2.3. SNP genotyping, linkage disequilibrium and population structure

The 3 K rice genome 1 M SNP genotype data employed in this study were downloaded from the Rice SNP-seek database (http://snp-seek.irri.org/_download.zul). Imputation of missing SNP data was performed using the software Beaglev5.4. A set of 5,53,229 high-quality SNPs for minor allele frequency ≥ 0.05 were retained to identify the peak associations using TASSELv5.2.25, and the SNPs were considered as significantly associated with targeted rice quality traits, adopting the suggestive -$\log_{10}$ p value >5. The population genetic structure and kinship relationships of 198 accessions were inferred using STRUCTURE v2.3.4 (Pritchard et al., 2000) and TASSEL v5.2.82 (Bradbury et al., 2007), respectively. LD between pairwise 5,53,229 filtered SNPs was assessed by estimating the coefficient of determination (R^2) in LDkit software. The rate of LD decay was the corresponding distance in base pairs (bp) at which the average R^2 value decayed to one half of its average maximum value.

2.4. Association mapping

Average data, kinship matrix, and structure were utilized for association analysis using GAPITv3.1.0 (https://cran.R-project.org/). Associations were studied using six GWAS methods, viz., MLM (Zhang et al., 2010), CMLM (Li et al., 2014), MLMM (Segura et al., 2012), mrMLM (Zhang et al., 2020), FASTmrMLM (Tamba & Zhang, 2018), and FASTmrEMMA (Wen et al., 2020), considering a significant LOD score value of 3. To improve the accuracy of GWAS results, QTNs detected by more than two GWAS methodologies were considered truly significant and used in the search for candidate loci. The GWAS results were presented in -log scale in the QQ and Manhattan plots using the qqman package in R software (Turner, 2018).

2.5. Prioritization of putative candidate loci and gene set analysis

To predict genes underlying eating quality traits in rice, SNPs with $p < 0.00001$ were considered significantly associated with targeted quality traits. The search radius of putative candidates was established based on R^2 half decay results. To confirm the GWAS results, a 100 kb region flanking the significant SNPs was scanned for the presence of previously reported eating quality-related QTLs and genes based on the information in online databases., QTARO (http://qtaro.abr.affrc.go.jp/) and Gramene (https://www.gramene.org/). Consequently, candidate transcripts were subjected to functional annotations using the rice database, i.e., RAP-DB (https://rapdb.dna.affrc.go.jp/). Furthermore, the gene set enrichment analysis (GSEA) of the identified loci was conducted utilizing the web-based graphical tool ShinyGO 0.77 (https://bioinformatics.sdstate.edu/go77/).

2.6. Tissue-specific expression of candidate loci and quantitative real-time PCR

The gene expression profiles of each of the putative candidates were obtained from the Gramene database across nine different developmental tissues, including leaf, early inflorescence, emerging inflorescence, anther, pistil, seeds 5 days after pollination (5 DAP), seeds 10 days after pollination (10 DAP), plant embryo and endosperm and visualized using a heatmap in R Studio. For verifying the reliability of RNA-Seq analysis and the expression levels of candidate genes, qRT-PCR analysis was performed using SYBR Green master mixes on Biorad CFX96 Real-Time PCR System (California, USA) with UBQ as an internal reference gene. Total RNA was isolated from the seeds of two genotypes with high (10298) and low (8405) TDF, respectively using Sigma Aldrich plant Spectrum™ total RNA kit and single-stranded cDNA was then synthesized using the Verso cDNA Synthesis Kit (AB-1453-B, Thermo Scientific). Gene-specific primer pairs were designed using Primer3Plus online software (https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi), and primer information has been given in supplementary table (Table S16). All qRT-PCRs were performed in three independent replicates, and $2^{-\Delta\Delta Ct}$ was used to calculate the relative expression to obtain the fold change values. Error bars in expression graphs represent the standard deviation (SD) of three biological replicates.

2.7. Statistical data analysis and phenotypic differences

The phenotypic evaluations were done in triplicate, and mean data were used in subsequent analysis. Descriptive statistics for all nine eating quality traits were conducted in R using the describe function. Phenotypic correlations between the quality traits and their level of statistical significance were computed in R using the correlation function. A correlation network was built to understand the interconnectivity between the selected quality traits using the R package graphics. To verify the associated locus between putative genes and rice quality traits, SNP Seek software (https://snp-seek.irri.org/) was used to

perform haplotype analysis considering the non-synonymous coding SNPs, trait mean values were compared for their significant differences at a threshold of $P = 0.05$ by student's *t*-test.

3. Results

3.1. Phenotypic variations and correlation network analysis of quality traits

Rice quality is a complex trait influenced by multiple factors and, therefore, cannot be characterized precisely by a single trait. We evaluated a total of 9 quality traits, including percentage of glycemic index (PGI), total dietary fibre (TDF), oil%, protein content (Prt), amylose (Amy), moisture % (Mst), phytate (Phy), phenols (Phn), and starch content (Str) in a subset of 3 K rice genome panel for two years. The frequency histograms displayed that the quality traits under study followed a normal and approximately normal to skewed distributions, especially for PGI and Amy (Fig. S1). The selected subset depicted wide variation for the all the quality traits analysed, with the coefficient of variance ranging from 0.88 % to 29.75 % (Table 1). The PGI varied from 48.03 % to 99.83 %, whereas TDF values ranged from 4.04 to 4.66. Considerable variation was also observed for Mst, Amy, Prt, Oil, and Str, which play a crucial role in selecting high-quality rice varieties considering consumer inclinations. In our present investigation, Oil showed a range of 3.36 to 7.2 %, Prt ranged from 10.91 to 16.55 g/100 g, Amy exhibited a broad range of 5.24 to 28.36 g/100 g, Mst in the range of 11.94 to 26.35 %, Phy varied from 0.87 to 1.68 g/100 g, Phn from 0.13 to 0.69 g/100 g, and Str in the range of 71.74 to 74.95 g/100 g. We thus anticipated that the subset is suitable for GWAS studies mapping SNPs and QTLs for quality traits in rice.

Correlation analysis revealed negative correlations for PGI with phytate, amylose, total dietary fibre, starch and protein (Tuano et al., 2021; Geng et al., 2021; Selvaraj et al., 2021; Haldipur & Srividya, 2020) while it showed weak positive correlations with oil and Phn (Table S2; Fig. S2). Amy demonstrated a significant negative correlation with Str and oil, Phn, Prt, and PGI. Oil displayed negative correlations with Phn, Str, Amy, while positive correlations with Prt, Phy, Mst, and PGI. Moisture specified significant positive correlations with Prt and Amy, while it exhibited significant negative correlations with Str, Mst, PGI, and TDF, though weakly. Phn showed negative correlations with oil, Phy, Mst, Prt, Amy, however, weak positive correlations were observed with TDF, Str, and PGI. Along with Prt significantly, TDF correlated positively with Prt, Phn, Phy, Mst, and Amy, while a negative correlation was observed with Str, oil, and Mst. Prt had a significant positive correlation with TDF and Phy, Mst, whereas it was inversely correlated with Str, Phn, and PGI. These findings suggest an intricate relationship between the selected quality traits and influence quality in rice in a coordinated way.

3.2. Population structure analysis and linkage disequilibrium decay

With a view to assess the count of subpopulations in selected rice subset, this genome-wide SNP marker data using the Bayesian model revealed a sharp peak Δk at k = 2 (Fig. S3A), implying presence of two subpopulations denoted as SP1 and SP2 (Fig. S3C). SP1 was the larger one, comprising of 173 rice accessions in total belonging to *indica* subpopulations (*ind1a, ind1b, ind2, ind3* and *indx)*. The smaller SP2 cluster accommodated 25 accessions of *japonica, aus/boro* and *intermediate* type of subpopulations. The LD across the rice genome was assessed using 5,53,229 filtered SNP markers with MAF > 5 %. The average LD across the rice genome was 100 kb (Fig. S3B). These SNPs were moderately uniformly distributed on twelve chromosomes suggesting their suitability for dissecting complex quality traits in rice.

3.3. Dissecting the genetic basis of grain quality-related traits using multi-model GWAS

In overall, 200 QTNs were discovered to be significantly associated with selected nine quality traits (LOD score > 3) by six different GWAS models (Table S3; Fig. 1). Manhattan plots and Quantile-quantile plots derived from GWAS illustrating the model's goodness of fit and the chromosomal location of the SNPs identified, respectively are presented in Fig. 2. Highest number of QTNs were identified to be associated with Phn, followed by 32, 30, 23, 20, 18, 15,12 and 10 QTNs with PGI, Phy, Oil, Amy, Prt, Mst, Str and TDF, respectively. These 200 QTNs were distributed across all the twelve chromosomes with the maximum QTNs on chromosome 2, followed by chromosome 12 (Fig. S4). Then, the QTNs located within LD decay distance of 100 kb were grouped into the same cluster, corresponding to the same QTL. Hence, these QTNs were categorized into 26 QTL clusters (Table S4). These clusters contained 39 QTNs were located across the genome on chromosomes 1,2,3,4,7,8,9,10,11, and 12. Among the key clusters, 6,1,5,7,3,6,5,2 and 3 clusters were associated with PGI, TDF, Oil, Mst, Amy, Phn, Phy, Prt, and Str, respectively (Table 2). Cluster q.9–1 was co-identified for Phn and Phy traits, explaining 1.77~ 13.78 % of the phenotypic variation. Clusters q.3–1 was found to be associated with Oil and Mst traits, explaining 34.83 % of the phenotypic variation. Cluster q.1–1 was co-identified to be associated with PGI and Phn traits and had 6.40 % of phenotypic variance explained. Clusters q.12–1 was co-identified for Mst and Phn, explaining 1.45~ 7.08 % of the phenotypic variance. Clusters q.4–2 and q.3–4 co-identified for Mst, Prt, and Phy had 5.22 % and 1.74~ 3.87 % of phenotypic variance correspondingly. q.2–2 was co-identified to be associated with Mst and Phn, had 6.96 % of phenotypic variance explained. Cluster q.11–1 was co-identified to be associated with Oil and Phn traits, explaining 0.97 % of the phenotypic variation. Cluster q.11–2 was co-identified to be associated with TDF and Prt traits and had 0.79~ 4.54 % of phenotypic variance explained. For the traits PGI and Str traits, Cluster q.10–3 was found to be associated, explaining 2.80 % of the phenotypic variance. Cluster q.7–3 was co-identified for Amy and Phy traits, explaining 1.20 % of the phenotypic variation. Among these 26 clusters, 20 were considered as highly reliable key clusters, as they were detected concurrently by at least 3 GWAS methods and colocalized/located in vicinity of already known rice quality genes/QTLs such as *Pho1*(Phosphate transporter 1) gene (**Hwang et al., 2016**; Sun et al., 2012), Histidine acid phosphatases (*OsPHY2*), Ser/Thr protein phosphatase family proteins, Calmodulin binding motif domain containing protein (Chen et al., 2021), Zinc Finger domain containing proteins, SBP-box gene family members, NIN like family proteins,

Table 1
Descriptive analysis of the traits related to grain quality of the selected subset of 3 K-RG panel.

	PGI	*TDF*	*Oil*	*Protein*	*Amylose*	*Moisture*	*Phytate*	*Phenols*	*Starch*
N	198	198	198	198	198	198	198	198	198
Mean	74.63	4.29	4.85	12.84	19.82	14.8	1.17	0.52	73.36
SD	13.57	0.1	0.71	0.84	5.9	2.59	0.13	0.09	0.65
CV	18.17	2.34	14.66	6.56	29.75	15.62	11.06	18.17	0.88
Range	48.03–99.83	4.04–4.66	3.36–7.20	10.91–16.55	5.24–28.36	11.94–16.36	0.87–1.68	0.13–0.69	71.74–74.95

Percentage of glycemic index (PGI), total dietary fibre (TDF), oil%, protein content (Prt), amylose (Amy), moisture % (Mst), phytate (Phy), phenols (Phn), and starch (Str).

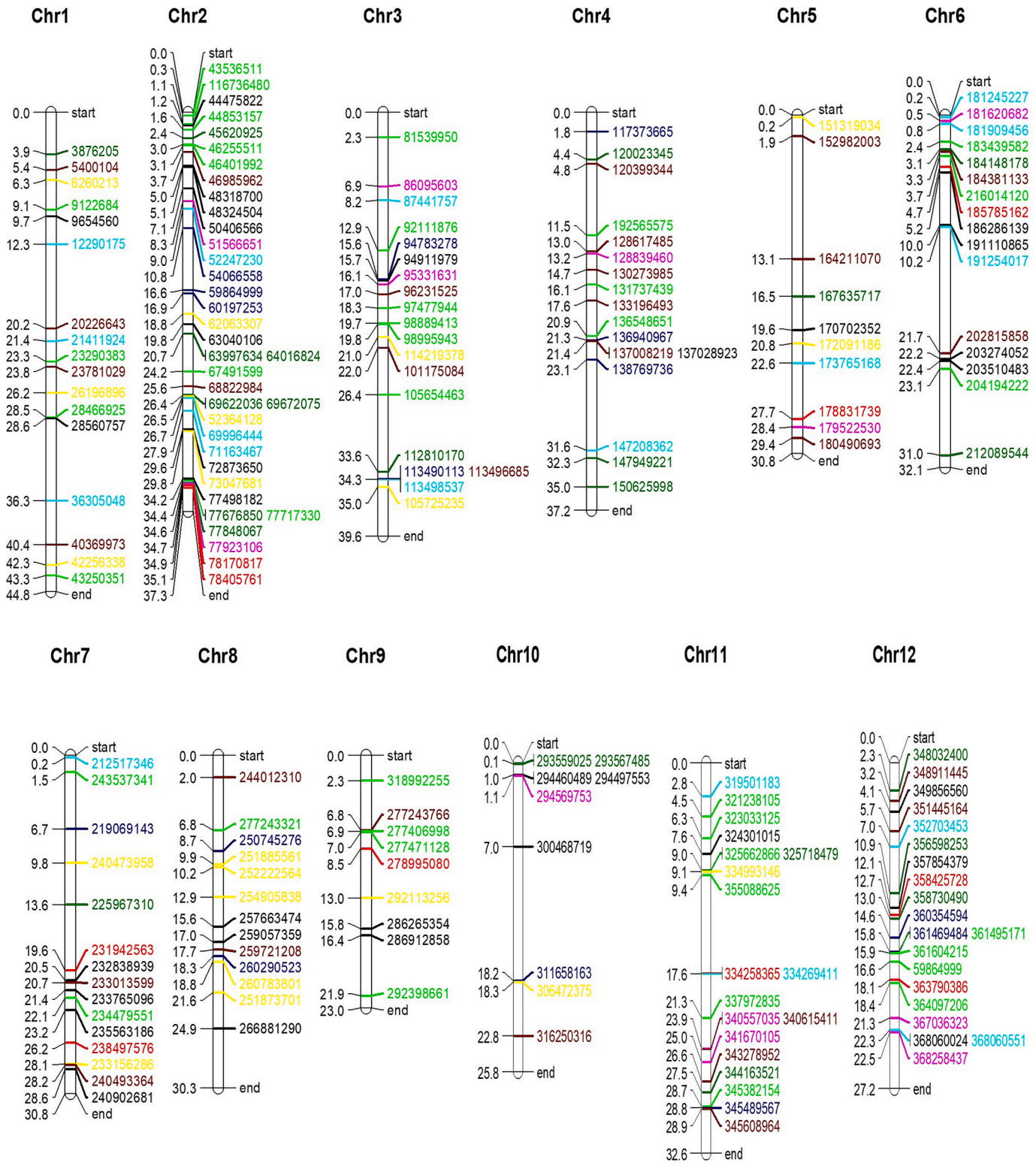

Fig. 1. Physical map of detected QTNs associated with nine quality-related traits in rice. The vertical bars represent different rice chromosomes. Horizontal lines represent the positions of SNP markers in each linkage group. Different colors represent QTNs detected for different quality traits assessed, i.e., dark green: Oil (%), brown: phytate (Phy), yellow: amylose (Amy), black: percentage of glycemic index (PGI); light blue: protein content (Prt), light green: phenols (Phn), dark blue: moisture (%) (Mst), red: total dietary fibre (TDF) and pink: starch (Str) content. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

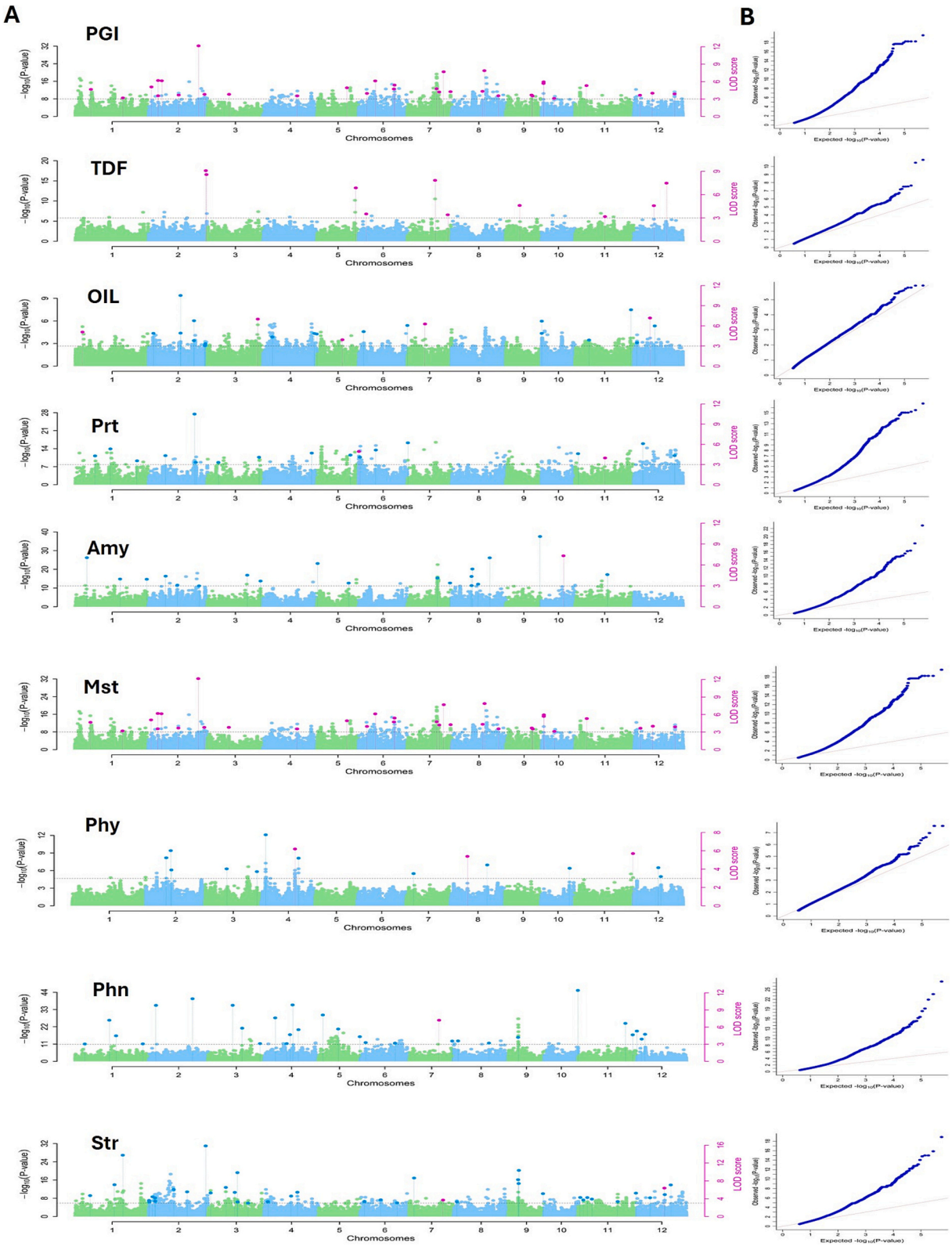

Fig. 2. Manhattan plots (left side) and Q-Q plots (right side) for eating quality traits using multi-model GWAS methods. (A) Manhattan plots of percentage of glycemic index (PGI), total dietary fibre (TDF), oil%, protein content (Prt), amylose (Amy), moisture % (Mst), phytate (Phy), phenols (Phn), and starch (Str) (starting from top to bottom). The dotted horizontal line represents the threshold LOD score > 3. Light color dots denote the QTNs identified by single GWAS model and the pink color dots depict the QTNs mapped by more than 2 GWAS models; (B) Q-Q plots of percentage of glycemic index (PGI), total dietary fibre (TDF), oil %, protein content (Prt), amylose (Amy), moisture % (Mst), phytate (Phy), phenols (Phn), and starch (Str) (starting from top to bottom). A Q-Q plot representing the observed versus expected *P* values of each identified QTN signifies the fitness of the GWAS model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Key QTL clusters identified by GWAS for selected quality traits.

Trait	QTL cluster	Peak SNP Marker	Chr	Pos	QTN effect	−log10(P)	PVE (%)	Method
PGI	q.1-1	28,466,925	1	28,466,925	−0.04	14.76	6.40	1,2,3,5
	q.2-1	48,318,700	2	5,047,777	−5.43	4.28	3.50	1,2,3,6
	q.6-1	191,110,865	6	10,027,745	−5.94	5.10	30.98	1,2,3,5
	q.7-1	232,875,954	7	20,544,047	−8.40	5.41	7.60	1,2,3,5
	q.10-2	294,460,489	10	975,219	−6.05	6.50	7.17	1,2,3,4
	q.10-3	294,497,553	10	1,012,283	−3.35	6.78	2.80	1,2,3,5
TDF	q.11-2	334,258,365	11	17,565,808	0.01	3.88	0.79	1,2,3,5
Oil	q.2-3	64,016,824	2	20,745,901	−0.17	5.71	5.12	1,2,3,4
	q.2-4	77,848,067	2	34,577,144	0.15	4.06	1.11	1,2,3,5
	q.3-1	107,010,610	3	27,802,437	0.72	5.41	34.83	1,2,3
	q.10-1	293,559,025	10	73,755	−0.10	5.67	1.85	1,2,3,5
	q.11-1	325,718,479	11	9,025,922	−0.08	4.63	0.97	1,2,3,5
Mst	q.2-2	59,864,999	2	16,594,076	1.00	6.85	6.96	1,2,3,4
	q.3-1	107,010,610	3	27,802,437	0.72	5.41	34.83	1,2,3
	q.3-3	112,810,170	3	33,601,997	−0.29	5.36	4.66	1,2,3
	q.3-4	113,490,113	3	34,281,940	−0.58	4.46	1.74	1,2,3,5
	q.4-1	117,373,665	4	1,751,673	−0.79	8.63	8.82	1,2,3,4
	q.4-2	136,940,967	4	21,318,975	−1.19	9.54	5.22	1,2,3,4
	q.12-1	361,469,484	12	15,755,821	−0.31	3.89	1.45	1,2,3,5
Amy	q.7-2	233,148,289	7	20,816,382	−2.82	6.03	0.71	1,2,3,5
	q.7-3	240,473,958	7	28,142,051	1.12	4.13	1.20	1,2,3,5
	q.8-1	251,873,701	8	9,844,173	0.98	5.12	2.75	1,2,3,5
Phn	q.1-1	28,466,925	1	28,466,925	−0.04	14.76	6.40	1,2,3,5
	q.2-2	59,864,999	2	16,594,076	1.00	6.85	6.96	1,2,3,4
	q.9-1	277,243,321	9	6,770,771	0.11	9.16	13.78	1,2,3,5
	q.9-2	277,406,998	9	6,934,448	0.05	8.27	5.61	1,2,3,4
	q.11-1	325,718,479	11	9,025,922	−0.08	4.63	0.97	1,2,3,5
	q.12-1	361,495,171	12	15,781,508	−0.04	4.25–10.16	7.08	1,2,3,4,6
Phy	q.3-4	113,496,685	3	34,288,512	0.03	3.81	3.87	1,2,3,5
	q.4-2	136,940,967	4	21,318,975	−1.19	9.54	5.22	1,2,3,4
	q.7-	240,473,958	7	28,142,051	1.12	4.13	1.20	1,2,3,5
	q.9-1	277,299,005	9	6,826,455	0.12	5.61	1.77	1,2,3,5
	q.10-4	314,820,179	10	21,334,909	−0.17	5.64	32.36	1,2,3
Prt	q.3-4	113,490,113	3	34,281,940	−0.58	4.46	1.74	1,2,3,5
	q.11-2	334,269,411	11	17,576,854	0.00–0.183	4.01–5.41	0.00–4.54	1,2,3,5,6
Str	q.2-4	77,848,067	2	34,577,144	0.151	4.0554	1.1053	1,2,3,5
	q.2-5	77,923,106	2	34,652,183	0.132	6.50	4.16	1,2,3,5
	q.10-3	294,497,553	10	1,012,283	−3.35	6.78	2.80	1,2,3,5

Percentage of glycemic index (PGI), total dietary fibre (TDF), oil%, protein content (Prt), amylose (Amy), moisture % (Mst), phytate (Phy), phenols (Phn), and starch (Str).

OsMADS6 (Lou et al., 2023; Ren et al., 2023).and consequently, were considered in our GWAS study for candidate gene search.

3.4. Identification of putative candidates based on key QTL clusters

869 putative candidate genes were annotated from the 26 key QTL clusters mentioned above. Out of which 206 candidate genes were

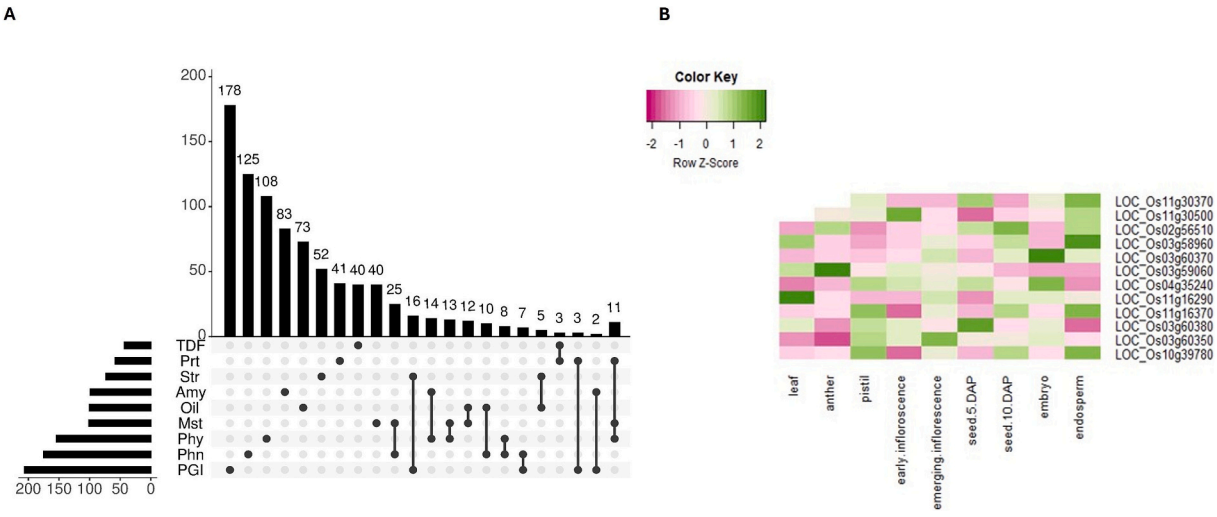

Fig. 3. Candidate genes identified for different quality traits. (A) Venn diagram showing an overlap of significant genes identified from GWAS. (B) Heatmap of the expression levels of putative candidates identified for different quality traits in nine different tissues (leaf, anther, pistil, early inflorescence, emerging inflorescence, seeds 5 days after pollination, seeds 10 days after pollination, embryo and the endosperm) of rice. Where Percentage of glycemic index (PGI), total dietary fibre (TDF), oil%, protein content (Prt), amylose (Amy), moisture % (Mst), phytate (Phy), phenols (Phn), and starch (Str).

identified in the QTL clusters for PGI, followed by 175, 100, 154, 99, 101, 73, 43, and 58 putative candidates for Phn, Oil, Phy, Amy, Mst, Str, TDF, and Prt, respectively (Table S5). 25 candidate genes were shared by Mst and Phn; 16 candidate genes by Str and PGI; 14 candidate genes by Phy and Amy; 2 candidate genes by Amy and PGI; 13 candidate genes by Phy and Mst; 12 candidate genes by Oil and Mst; 10 candidate genes by Oil and Phn; 11 candidate genes by Prt, Phy and Mst; 8 candidate genes by Phn and Phy; 7 candidate genes by PGI and Phn; 5 candidate genes by Str and Oil; 3 candidate genes by TDF and Prt; and 3 candidate genes by PGI and Prt (Fig. 3A). These results showed that the putative candidate genes for quality traits were overlapped, emphasizing the reliability of these candidates.

3.5. Gene set analysis and expression analysis of candidate loci

GO enrichment analysis was then performed to detect the functional categories across the grain quality putative candidates identified by GWAS. The results demonstrated that these genes associated with PGI were enriched in GO terms related to ubiquitin molecules, galactokinase domain, phospholipid acyltransferase and phospholipase d family, GPAT domain superfamily, nascent polypeptide-associated complex alpha-like UBA domain, PO loop repeats, and other biological functions (Fig. S5A, Table S6). The putative candidates associated with TDF were enriched in GO terms primarily in NADP activity, beta-1,4-mannosyltransferase activity, Glycosyltransferase family 21, Duf639 protein and SBP domain, AUX family and auxin binding, and other biological functions (Fig. S5B, Table S7).

The candidate genes related to oil were enriched in the U-box domain, NADH cytochrome P450 reductase and DHHC palmitoyl-transferase, cupin domain, TAP42-like family and TIP-41-like family, TB2/DP1, HVA22 and PAPD-like superfamily and molecular functions (Fig. S5C, Table S8). The candidate genes related to Prt were enriched with TIM10-like protein and the TIM22 translocase, GDSL/SGNH-like acyl transferase activity, PMR5 N-terminal domain and cellular functions (Fig. S5D, Table S9). The candidates related to Amy were enriched largely in nucleoside phosphate binding, carbohydrate derivative binding, phospholipid/glycerol acyltransferase domain-containing proteins, ribonucleotide binding, ion binding, glycerol acyltransferase and GPAT N-terminal domain superfamily, and cellular functions (Fig. S5E, Table S10).

The candidate genes putatively associated with Mst were enriched in Glutaredoxin-like and protein Duf2232, methyl-sterol oxidase activity and squalene epoxidase activity, ACT domain-containing protein ACR1–12 and UV-B-induced protein At3g17800-like, NADPH-cytochrome P450 reductase and DHHC palmitoyl-transferase, cupin domain, TAP42-like family and TIP-41-like family, PMR5 N-terminal domain, and other cellular functions (Fig. S5F, Table S11). The candidate genes related to Phy were enriched in GO terms related to DNA replication origin binding, DNA replication initiation, ion binding, ribonucleotide binding, and carbohydrate derivative binding (Fig. S5G, Table S12).

The candidate genes associated with Phn were enriched principally in NmrA-like domain and ACT domain, STAM Vps27/Hrs/STAM domain, oxylipin biosynthesis and acyl transferase/acyl hydrolase, TMEM 14 family and Mce/MlaD, FAD/NAD(P) binding domain, NAP-like superfamily, E3 binding domain, and molecular functional category (Fig. S5H, Table S13). The candidate genes related to Str were enriched in PHO1 and SPX domains, GAGA binding protein-like family, EXS family, along with their molecular functions (Fig. S5I, Table S14).

Among the putative candidate genes associated with Oil, two candidates *LOC_Os02g56510* (*OsPHO 1*, Phosphate transporter 1) and *LOC_Os03g58960* (DHHC zinc finger domain containing protein) were detected in q.2–4 and q.3–3, respectively (Table 3). These genes showed high expression levels in the emerging inflorescence, seeds 5 days after pollination, 10 days after pollination and the endosperm, and hence may play a vital role in regulating oil biosynthesis and accumulation. Of the two candidate genes for q.11–1 associated with Phn, *LOC_Os11g16290* encodes a NIN-like family protein, and *LOC_Os11g16370* encodes a uridine/cytidine kinase-like 1 gene. These genes displayed higher expression values in the inflorescence, seeds at 5 and 10 days after pollination, embryo together with the leaf (*LOC_Os11g16290*) and pistil (*LOC_Os11g16370*) (Table 3; Fig. 3B).

Table 3
Putative candidate genes associated with rice quality traits identified by multiple GWAS models.

QTL cluster	*SNP*	*Trait*	*GeneID*	*Annotation*
q.2–4	77,848,067	Oil, Str	*LOC_Os02g56510*	phosphate transporter 1 (*OsPHO1*)
q.3–3	112,810,170	Oil, Mst	*LOC_Os03g58960*	DHHC zinc finger domain-containing protein
q.3–4	113,490,113	Prt, Mst, Phy	*LOC_Os03g60370*	Histidine acid phosphatases, putative, expressed
q.4–2	136,940,967	Mst, Phy	*LOC_Os04g35240*	Calcium/calmodulin-dependent protein kinase (*OsSAPK7*)
q.10–4	314,820,179	Phy	*LOC_Os10g39780*	Protein Phosphatase 2C (*OsPP2C72*)
q.11–1	325,718,479	Phn	*LOC_Os11g16290*	NIN like family protein
q.11–2	334,258,365	TDF	*LOC_Os11g30500*	HVA22-related protein
q.3–2	112,810,170	Mst	*LOC_Os03g59060*	Phosphatase 2 A isoform 2 belonging family 2 (*OsPP2 Ac-2*)
q.3–4	113,490,113	Prt, Phy	*LOC_Os03g60350*	Leaf senescence related protein (*OsTBL5*)
q.11–1	325,718,479	Phn	*LOC_Os11g16370*	uridine/cytidine kinase-like 1
q.11–2	334,258,365	TDF	*LOC_Os11g30370*	SBP-box gene family member (*OsSPL19*)
q.3–4	113,490,113	Prt, Phy	*LOC_Os03g60380*	cinnamoyl CoA reductase

Percentage of glycemic index (PGI), total dietary fibre (TDF), oil%, protein content (Prt), amylose (Amy), moisture % (Mst), phytate (Phy), phenols (Phn), and starch (Str).

Among the putative genes associated with Mst, four genes *LOC_Os03g60370* (Histidine acid phosphatase), *LOC_Os03g59060* (*OsPP2 Ac-2*, Phosphatase 2 A isoform 2 belonging to family 2), and *LOC_Os04g35240* (*OsSAPK7*, Camk_Camk_Like.25 - Camk Includes Calcium/Calmodulin Dependent Protein Kinases), *LOC_Os03g58960* (DHHC zinc finger domain containing protein) were found in q.3–4, q.3–2, q.4–2 and q.3–3, respectively. The expression of *LOC_Os04g35240* was high in all the tissues except the leaf, anthers and the endosperm, while the expression of *LOC_Os03g59060* was comparatively higher in the leaf, anther, pistil, early inflorescence and the emerging inflorescence (Table 3; Fig. 3B). The expression of *LOC_Os03g60370* and *LOC_Os03g58960* was significantly higher in emerging inflorescence, seeds 10 days after pollination, embryo and the endosperm, signifying their potential roles in regulating moisture availability to the developing embryo and seedling which is crucial for grain filling and grain quality.

Among the putative candidates associated with Prt, three genes *LOC_Os03g60370* (*OsPHY2*, histidine acid phosphatase), *LOC_Os03g60350* (*OsTBL5*, leaf senescence related protein), and *LOC_Os03g60380* (*OsCCR2*, cinnamoyl CoA reductase) were detected in the cluster q.3–4 (Table 3; Fig. 3B). The genes *LOC_Os03g60370*, *LOC_Os03g60350* and *LOC_Os03g60380* were highly expressed in the pistil, inflorescence, seeds at different stages and the embryo and, also leaf in case of *LOC_Os03g60380* suggesting their critical role in protein metabolism, mobilizing release/transport of amino acids specifically during seed germination and eventually affecting the grain size and protein quality.

Another candidate gene for q.2–4 associated with Str, *LOC_Os02g56510* encoding *OsPHO1*, phosphate transporter-1 gene exhibited significantly higher expression in anther, seeds 5 and 10 days after pollination, and the endosperm pointing towards its involvement

in uptake and transport of phosphates and thus, is considered as the most likely candidate gene governing starch biosynthesis in rice grains (Table 3; Fig. 3B). Our GWAS results suggested five candidate genes for Phy, including *LOC_Os03g60370* (histidine acid phosphatase), *LOC_Os04g35240* (*OsSAPK7*, Camk_Camk_Like.25 - Camk Includes Calcium/Calmodulin Dependent Protein Kinases), *LOC_Os03g60350* (*OsTBL5*, leaf senescence related protein), *LOC_Os10g39780* (*OsPP2C72*, Protein Phosphatase 2C), and *LOC_Os03g60380* (cinnamoyl CoA reductase) which were found in q.3–4, q.4–2, q.3–4, q.10–4 and q.3–4, respectively. Notably, these genes facilitate remobilisation of nitrogen and phosphorus from the vegetative tissues to the developing rice grains, influencing the bioavailability of nutrients at the grain filling stage (**HajibaraT and Saidi, 2022**; **Fuchs et al., 2013**). Their higher expression in the pistil, early and emerging inflorescence, seeds at 5 and 10 days after pollination, embryo and the endosperm also indicates their potential roles impacting the overall rice grain quality and phytate content (Table 3; Fig. 3B).

Among the candidate genes associated with TDF, two genes including *LOC_Os11g30370* (*OsSPL19*, SBP-box gene family member) and *LOC_Os11g30500* (HVA22-related protein) located in q.11–2 exhibited relatively high expression in the pistil, early inflorescence, seeds 5 days after pollination, embryo and the endosperm indicating their regulatory role in flowering, panicle architecture, and the endosperm development, and might influence dietary fibre profile of the rice grains (Table 3; Fig. 3B). The qRT-PCR results also showed the expression levels of *LOC_Os11g30370 and LOC_*Os11g30500 were extensively higher in IRIS_313–10,298 with high TDF than that in the genotype IRIS_313–8405 with low TDF (Fig. 5). The introduction of the *OsSPL19* molecular module into high-yielding rice varieties could substantially improve rice quality, providing a novel approach for the synergistic improvement of rice quality. Hence, these twelve candidate genes were determined as promising for haplotype analysis.

3.6. Haplotype analysis of putative candidates

A set of twelve candidate loci were subjected to haplotype analysis considering the non-synonymous coding SNPs, then *t*-test was conducted to test the significant variations among the haplotypes. The haplotypes revealed and phenotypic distribution of each grain quality trait were then visualized as boxplots in R Studio using ggplot2 package.

For the gene *LOC_Os03g60370* encoding Histidine acid phosphatases (*OsPHY2*) associated with Prt, Mst and Phy, three haplotypes were identified based on eight SNPs in the genic region. For Phy, the genotypes having HapB (TTTGC) was significantly superior to the genotypes with HapA (TTTGC) and HapC (GCCAC) with a mean value of 1.14 g/100 g and 1.16 g/100 g phytate respectively. For the gene *LOC_Os11g16290* encoding NIN like family protein, present within the LD region of q11–1 five haplotypes were identified based on forty one SNPs in the coding region of the gene (Fig. 4). The accessions possessing HapC showed higher mean value of 0.603 g/100 g while the genotypes having HapE displayed the least mean value of 0.456 g/100 g. For Oil% and Str, Mst and Phy, Prt and Phy, 3 genes *LOC_Os02g56510*, *LOC_Os04g35420* and *LOC_Os03g60380* encoding *OsPHO1*, *OsSAPK7* and *OsCCR2* were identified for haplotype analysis. Three, three and four haplotypes were identified for *LOC_Os02g56510*, *LOC_Os04g35420* and *LOC_Os03g60380* respectively. However, none of the haplotypes showed significant associations with aforementioned traits. The details of the haplotypes identified are given in the supplementary table (Table S15). For the gene *LOC_Os11g30500* encoding HVA-22 related protein involved in governing TDF, HapC (GC) with a mean value of 4.366 g/100 g depicted superiority over Hap A(CT) and HapB(CC). Similarly, HapC and HapE contributed to the respective highest and lowest mean values of TDF for another candidate gene *LOC_Os11g30370* encoding *OsSPL19*. Also, significant differences in TDF among the gene haplotypes were consistent with the RNA-Seq and qRT-PCR results (Fig. 5) signifying that SPL19 gene and HVA22 gene may play an imperative role in rice quality.

4. Discussion

Rice is the second most widely consumed cereal worldwide, and grain quality has been a key concern for consumers and rice breeders. Multiple genes with several gene expression profiles and regulatory pathways contribute to quality-related traits; henceforth, geneticists incessantly develop approaches to dissect their complexity. The intricacy at the genetic level and an inadequate number of germplasm

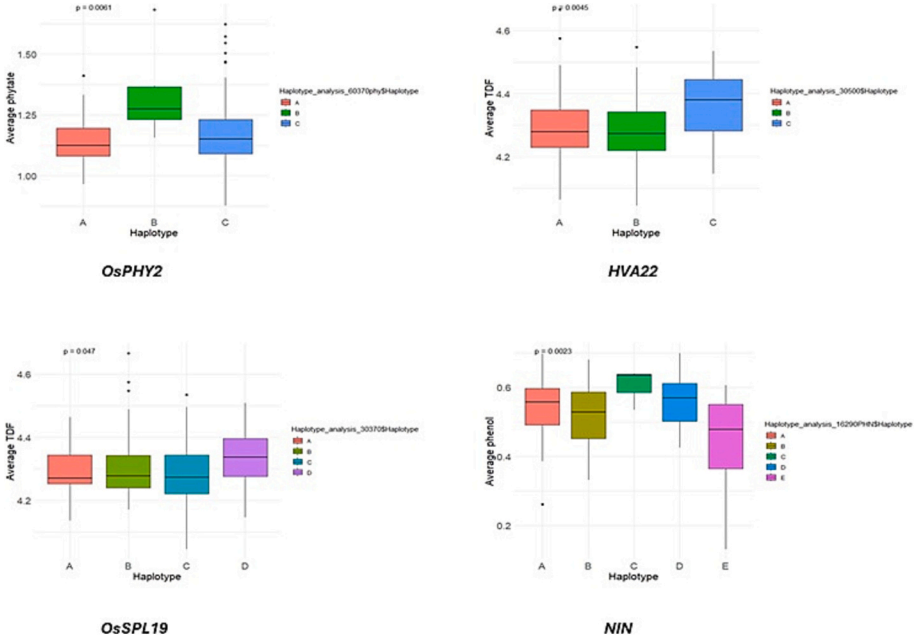

Fig. 5. Verification of the relative expression of LOC_Os11g30370 and LOC_Os11g30500 in IRIS_313–10,298 and IRIS_313–8405 by RT-PCR. Ubiquitin was used as the control. Amplification of cDNA in high (IRIS_313–10,298) and low (IRIS_313–8405) TDF genotypes was done. Bars represent standard deviation of three biological replicates. *$P < 0.05$; ***$P < 0.001$ (two-tailed Student's t-test).

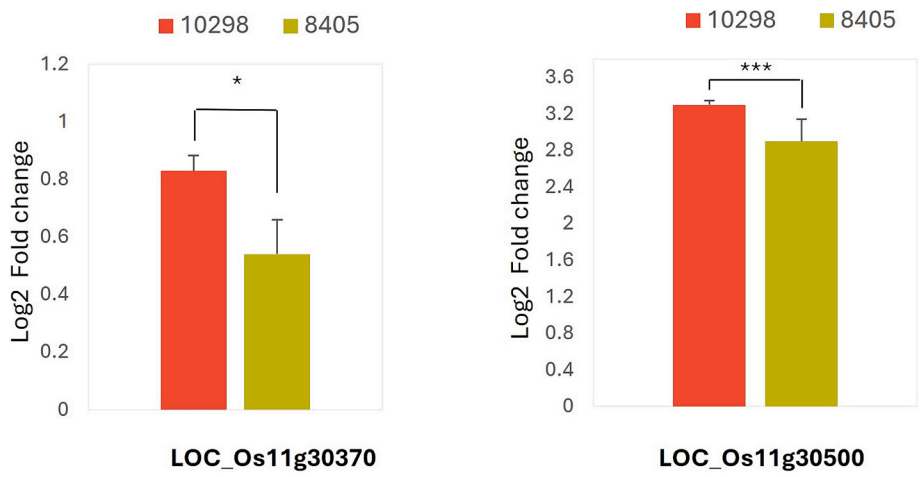

Fig. 4. Boxplots depicting the significant differences among the haplotypes identified for the genes *OsPHY2, HVA22* (upper panel) and *OsSPL19, NIN* (lower panel) for selected grain quality traits in rice.

lines with the required eating and cooking quality attributes are the major barriers to grain quality improvement in rice. Using a subpanel of 198 accessions and 5,53,229 SNP markers, significant QTNs and candidate genes associated with grain quality traits were identified. In this study, based on the frequency distribution, it may be determined that a wide range of variability existed for PGI, Amy, Str, Mst and other quality related traits in the selected subset (Fig. S1). Likewise, our correlation analysis revealed that Prt exhibited significant negative correlations with Str, Amy, Phn and PGI however, positive correlation was observed with TDF and Phy, consistent with several reports (Pinta et al., 2024; John et al., 2023; Qadir & Wani, 2023; Song et al., 2019; Balindong et al., 2018; Zhang et al., 2017; Lee et al., 2015). Phn exhibited negative correlations with Oil, Amy, Prt, Phy, and Mst, though positively correlated with TDF and PGI, similar to earlier studies (John et al., 2023; Liu et al., 2020). Str showed negative correlation with TDF, Prt, Phy, Oil, Amy, PGI, and Mst, suggesting its vital role in mitigating the detrimental effects of Type II diabetes globally (Takahashi et al., 2024). Additionally, we observed negative correlations between PGI and Amy, Str, Prt and Phy (Anacleto et al., 2019; Badoni et al., 2024; Guzman et al., 2017; Jabeen et al., 2021; John et al., 2023; Kumar et al., 2020; Ngo et al., 2023; Pautong et al., 2022; Selvaraj et al., 2021; Zhou et al., 2016) confirming the hypothesis that higher proportions of Amy, Str and Prt are believed to be desirable for public health with lower PGI (Badoni et al., 2024; Jukanti et al., 2025). Phy displayed negative correlations with TDF, Str, Amy, Phn, and PGI, and positive correlations with Prt and oil, which corroborate with the study done by Lee et al. (2015). Interestingly, it has been found that dehusked rice cultivars with high phenolic compounds and low phytic acid possess antioxidant activity and hence, are believed to be suitable for infants and low GI diets (Lee et al., 2015). Moreover, TDF, oil, and moisture displayed moderately low to moderately high non-significant correlations with the other traits, indicating these traits cannot determine the grain quality exclusively, and it is imperative to take into consideration other quality traits. Pleiotropic SNPs identified by GWAS supported well Pearson's correlations and network visualization approach, further highlighting the interdependence between the selected quality traits (Fig. S2).

Previous studies have reported genetic dissection of different complex traits (Abbai et al., 2019; Anacleto et al., 2019; Badoni et al., 2024; Mbanjo et al., 2023; Selvaraj et al., 2021; Singh et al., 2024; Song et al., 2019). However, the understanding of the genetic mechanisms regulating rice grain quality is extremely limited. Therefore, we conducted a comprehensive GWAS study and identified 200 QTNs associated with rice quality. Particularly, 32 QTNs associated with PGI were mapped to 6 QTL clusters located on chromosomes 1,2,3,5,6,7,10, and 12 (Table S3, Fig. 1) that differed from those already reported loci for PGI, which were mainly positioned on chromosomes 1,2,6,8,9, and 11 (Anacleto et al., 2019; Badoni et al., 2024; Selvaraj et al., 2021). 20QTNs detected for Amy (Table S3) were mapped to 3 QTL clusters located on chromosomes 1,2,3,5,7,8,9,10,11 and contrasted with the known Amy genes at chromosome 6 (Chen et al., 2024; Okpala et al., 2022; Zhong et al., 2021). Similarly, 30QTNs identified for Phy were mapped to 5 QTL clusters located across the genome were different from the genes documented for phytic acid on chromosomes 3,4,6, and 8 (Gyani et al., 2020; Perera et al., 2018; Perera et al., 2019; Tp et al., 2022). 40 QTNs detected for Phn mapped to 6 QTL clusters located on chromosomes 1,2,3,4,6,7,8,9,11, and 12 differed from previously known Phn genes on chromosome 1,2,4, 6,7,9,11, and 12 (Xia et al., 2021; Xu, Bao, He, & Park, 2016a; Shao et al., 2014). Remarkably, 118 of 200 QTNs identified in the current study were located in close vicinity of already reported QTLs/grain quality genes whereas the remaining 82 QTNs can be considered novel (Table S3).

Additionally, 13 QTL clusters were simultaneously detected by more than one quality trait (Table 2), such as q.9–1, q.7–3, q.11–1 and q.11–2. For PGI, 2 QTL clusters (q.1–1 and q.10–3) were co-detected by PGI, Phn, and Str, and for Mst, 4 of 8 QTL clusters (q.2–2, q.3–4, q.4–2, and q.12–1) were co-detected by Mst, Phn, Prt, and Phy. Likewise, many putative candidates were also found to be associated with QTL clusters detected with multiple quality traits (Fig. 3A), emphasizing the complexity of rice grain quality and pleiotropic effects of the identified candidates. These findings collectively contribute to our understanding of the genetic basis of the assessed quality traits and multifaceted developmental processes regulating preferred nutritional attributes in rice.

The QTL cluster governing Str and oil% (q.2–4) was found near to the position of *OsMADS6* gene encoding MADS-box transcription factor. Previous reports demonstrate the significant influence of *OsMADS6* on the expression of seed storage protein (SSP) genes and starch synthesis related gene (*SBEIII*) encoding 1,4-Alpha-Glucan-Branching Enzyme resulting in impaired grain filling, grain shape and reduced rice quality (Ren et al., 2023; Xu, Bao, Kim, & Park, 2016b). Furthermore, the above cluster harbours *PHO1* gene (*LOC_Os02g56510*) and has been reported to mediate the uptake and translocation of inorganic phosphates essential for grain filling in developing rice seeds (Che et al., 2020; Secco et al., 2010; Yan et al., 2024). Primarily, previous reports revealed over-accumulated inorganic phosphate in growing seeds of the *Ospho1* mutants inhibited the activity of *AGPase* (ADP-glucose pyrophosphorylase) gene, crucial for starch synthesis, and the grain-filling defects were reduced by the overexpression of *AGPase* in *Ospho1* mutants (Bao et al., 2012; Ma et al., 2021). Grain filling is a rate limiting step for seed development and determines the significance of a starch to many food

and non-food applications. Considering these findings, we anticipate the role of *PHO1* gene in affecting the rice quality.

In addition, we found 6 QTL clusters novel to this study expanding the established genetic landscape linked to grain quality traits in rice. The accessions harbouring superior haplotypes associated with the selected grain quality traits can be utilized in marker assisted rice breeding programmes. For the Phy, 3 QTL clusters identified on chromosomes 3,4 and 10 are novel and specific to this GWAS study as previous studies reported the QTLs in chromosomes 2,3, 5,7,8 and 12 (Gyani et al., 2020; Perera et al., 2018; Perera et al., 2019; Tp et al., 2022).

The significant SNP 325718479 of the QTL cluster q.11–1 associated with Phn i.e., was found within the candidate gene *NIN* (*LOC_Os11g16290*), a member of NIN-like proteins (NLPs). Transgenic studies strongly advocate their role as nitrate responsive transcription factors improves yield and NUE in rice by participating in developmental processes, stress signalling, and defense mechanisms (Alfatih et al., 2020; Pradhan et al., 2019). Loss of *OsNLP4* mutation caused significant reduction in rice yield and NUE in comparison to wild type rice under different N conditions (Wu et al., 2021). All these results suggest that these genes, influence protein content by regulating nitrogen uptake and potentially phenolic content of the rice grains. The current GWAS study for the first time shows the association of NIN protein in the phyate metabolism in rice while previous studies have highlighted the role of *OsPHY1* and other inositol phosphate kinases and synthases (Ali et al., 2013; Karmakar et al., 2020; Suzuki et al., 2007).

We also found an association of q.3–4 with Phy harbouring Histidine acid phosphatases (*LOC_Os03g60370, OsPHY2*) gene (Table 2) that plays a crucial role in hydrolysis of phytates found in rice grains (Scott et al., 2024) and was reported to catalyse the transfer of a phosphoryl group from phosphomonoesters to water driving the release of phosphates from phytate (Acquistapace et al., 2020). Thus, it is plausible that the moisture content of rice grains influences the activity of phytases and the breakdown of phytic acid by HAPhys.

Dietary fibre is an important trait to assess the grain quality and dimensions for selecting appropriate breeding lines for the rice eating population and market acceptance. For TDF, two genes i.e., HVA-22 related protein (*LOC_Os11g30500*) and SBP-box gene family protein (*LOC_Os11g303700*) were identified on chromosome 11 with 3 and 4 haplotypes respectively (Fig. 4). The roles HVA-22 related protein and *OsSPL19* in fibre quality including the content of dietary fibre in the rice hull, bran layer, textural characteristics in different crops and grain protein respectively were deciphered previously (Zhang et al., 2023; Ren et al., 2023; Wattanavanitchakorn et al., 2023; Ren et al., 2023; Wang et al., 2012; Yang et al., 2008). Their association with dietary fibre has not been reported in any of the previous researches. Xie et al., 2011 also reported limited QTLs for dietary fibre on chromosomes 2,4,6, 8 and 11 by utilizing the rice BIL population (Xie et al., 2011). High expression in pistil, inflorescence and seed tissues also suggested these genes might be a major determinant of rice grain quality (Fig. 3B). Quantitative RT-PCR has been anticipated for validation of candidate genes identified by association studies (Cao et al., 2024). In this study, we used qRT-PCR to assess the relative expression levels of the candidate genes associated with TDF content viz., *OsSPL19* (*LOC_Os11g30370*) and *HVA-22* (*LOC_Os11g30500*) between the two genotypes: IRIS_313–10,298 with high TDF and IRIS_313–8405 with low TDF. Notably, we found that both the genes displayed significant differences in the contrasting genotypes (Fig. 5). The QTL clusters and putative candidates, therefore, identified in our GWAS study would be valuable molecular tools for gaining insights into rice quality improvement.

5. Conclusion

Through GWAS, 20 key QTL clusters and several rice grain quality genes were identified. Finally, two important candidate genes *OsSPL19* (*LOC_Os11g30370*) and *HVA-22* (*LOC_Os11g30500*) were functionally characterized and validated by RT-PCR analysis. The identified QTLs and candidate genes signify reliable sources for developing new, improved rice varieties with enhanced nutritional value in future rice improvement and might be used in value-added trait introgression breeding for faster deployment of premium quality rice grains consistent with the needs of the consumers and market trade.

CRediT authorship contribution statement

Supriya Sachdeva: Writing – original draft, Visualization, Validation, Software, Formal analysis, Data curation. **Rakesh Singh:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Harshita Singh:** Formal analysis, Data curation. **Rakesh Bharadwaj:** Writing – review & editing, Resources, Methodology. **Antil Jain:** Formal analysis, Data curation. **Vikas K. Singh:** Writing – review & editing. **Uma Maheshwar Singh:** Writing – review & editing, Resources. **Arvind Kumar:** Writing – review & editing. **Gyanendra Pratap Singh:** Writing – review & editing.

Funding

This work was funded by the Department of Biotechnology, Government of India, grant number "BT/PR32853/AGIII/103/1059/2019". The funder has no role in the design of the GWAS study and data collection, analysis, interpretation, and writing the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to thank the International Rice Research Institute South Asia Regional Centre (IRRI-SARC), Varanasi, India, for providing the rice accessions for the GWAS study. The facility provided by the Director, ICAR-NBPGR, is also duly acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochms.2025.100313.

Data availability

Data will be made available on request.

References

Abbai, R., Singh, V. K., Nachimuthu, V. V., Sinha, P., Selvaraj, R., Vipparla, A. K., … Kumar, A. (2019). Haplotype analysis of key genes governing grain yield and quality traits across 3K RG panel reveals scope for the development of tailor-made rice with enhanced genetic gains. *Plant Biotechnology Journal, 17*(8), 1612–1622.

Abhijith, K. P., Gopala Krishnan, S., Ravikiran, K. T., Dhawan, G., Kumar, P., Vinod, K. K., … Badhran, S. K. (2022). Genome-wide association study reveals novel genomic regions governing agronomic and grain quality traits and superior allelic combinations for basmati rice improvement. *Frontiers in Plant Science, 13*, Article 994447.

Acquistapace, I. M., Zi, M. A., Li, A. W., Salmon, M., Kühn, I., Bedford, M. R., Brearley, C. A., & Hemmings, A. M. (2020). Snapshots during the catalytic cycle of a histidine acid phytase reveal an induced-fit structural mechanism. *Journal of Biological Chemistry, 295*(51), 17724–17737.

Alfatih, A., Wu, J., Zhang, Z. S., Xia, J. Q., Jan, S. U., Yu, L. H., & Xiang, C. B. (2020). Rice NIN-LIKE PROTEIN 1 rapidly responds to nitrogen deficiency and improves yield and nitrogen use efficiency. *Journal of Experimental Botany, 71*(19), 6032–6042.

Ali, N., Paul, S., Gayen, D., Sarkar, S. N., Datta, K., & Datta, S. K. (2013). Development of low phytate rice by RNAi mediated seed-specific silencing of inositol 1, 3, 4, 5, 6-pentakisphosphate 2-kinase gene (IPK1). *PLoS One, 8*(7), Article e68161.

Anacleto, R., Badoni, S., Parween, S., Butardo, V. M., Jr., Misra, G., Cuevas, R. P., … Bird, A. R. (2019). Integrating a genome-wide association study with a large-scale transcriptome analysis to predict genetic regions influencing the glycemic index and texture in rice. *Plant Biotechnology Journal, 17*(7), 1261–1275.

Anacleto, R., Badoni, S., Parween, S., Butardo, V. M., Jr., Misra, G., Cuevas, R. P., … Bird, A. R. (2019). Integrating a genome-wide association study with a large-scale transcriptome analysis to predict genetic regions influencing the glycemic index and texture in rice. *Plant biotechnology journal, 17*(7), 1261–1275. https://doi.org/10.1111/pbi.13051

Badoni, S., Pasion-Uy, E. A., Kor, S., Kim, S. R., Tiozon, R. N., Jr., Misra, G., … Slamet-Loedin, I. (2024). Multiomics of a rice population identifies genes and genomic regions that bestow low glycemic index and high protein content. *Proceedings of the National Academy of Sciences, 121*(36), Article e2410598121.

Bagchi, T. B., Sharma, S., & Chattopadhyay, K. (2016). Development of NIRS models to predict protein and amylose content of brown rice and proximate compositions of rice bran. *Food Chemistry, 191*, 21–27.

Balindong, J. L., Ward, R. M., Liu, L., Rose, T. J., Pallas, L. A., Ovenden, B. W., … Waters, D. L. (2018). Rice grain protein composition influences instrumental measures of rice cooking and eating quality. *Journal of Cereal Science, 79*, 35–42. https://doi.org/10.1016/j.jcs.2017.09.008

Bao, J., Lu, Y., Yang, F., Zhang, G., Shao, Y., Corke, H., & Sun, M. (2012). Nucleotide polymorphisms in OsAGP genes and their possible association with grain weight of rice. *Journal of Cereal Science, 55*(3), 312–317.

Bharali, V., Yadla, S., Thati, S., Bitra, B., Karapati, D., Chirravuri, N. N., … Jukanti, A. K. (2025). Mapping genomic regions for grain protein content and quality traits in milled Rice (Oryza sativa L.). *Plants, 14*(6), 905.

Birla, D. S., Malik, K., Sainger, M., Chaudhary, D., Jaiwal, R., & Jaiwal, P. K. (2017). Progress and challenges in improving the nutritional quality of rice (Oryza sativa L.). *Critical Reviews in Food Science and Nutrition, 57*(11), 2455–2481.

Biselli, C., Bagnaresi, P., Cavalluzzo, D., Urso, S., Desiderio, F., Orasen, G., … Ben Hassen, M. (2015). Deep sequencing transcriptional fingerprinting of rice kernels for dissecting grain quality traits. *BMC Genomics, 16*, 1–28.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics, 23*(19), 2633–2635.

Butardo, V. M., Jr., & Sreenivasulu, N. (2016). Tailoring grain storage reserves for a healthier rice diet and its comparative status with other cereals. *International Review of Cell and Molecular Biology, 323*, 31–70.

Cao, H., Gong, R., Xiong, L., Wang, F., Gu, H., Li, S., … Qiu, X. (2024). Comparative metabolome and transcriptome analysis reveals the possible roles of Rice phospholipase a genes in the accumulation of oil in grains. *International Journal of Molecular Sciences, 25*(21), 11498.

Che, J., Yamaji, N., Miyaji, T., Mitani-Ueno, N., Kato, Y., Shen, R. F., & Ma, J. F. (2020). Node-localized transporters of phosphorus essential for seed development in rice. *Plant and Cell Physiology, 61*(8), 1387–1398.

Chen, M., Ni, L., Chen, J., Sun, M., Qin, C., Zhang, G., Zhang, A., & Jiang, M. (2021). Rice calcium/calmodulin-dependent protein kinase directly phosphorylates a mitogen-activated protein kinase kinase to regulate abscisic acid responses. *The Plant Cell, 33*(5), 1790–1812.

Chen, X., Guo, Q., Yang, X., Yuan, M., Song, J., Fu, H., Zhang, H., Xu, P., Liao, Y., Ali, A., & Du, K. (2024). Triple gene mutations boost amylose and resistant starch content in rice: insights from sbe2b/sbe1/OE-Wxa mutants. *Frontiers in Plant Science, 15*, 1452520.

Cruz, M., Arbelaez, J. D., Loaiza, K., Cuasquer, J., Rosas, J., & Graterol, E. (2021). Genetic and phenotypic characterization of rice grain quality traits to define research strategies for improving rice milling, appearance, and cooking qualities in Latin America and the Caribbean. *The Plant Genome, 14*(3), Article e20134.

Fodor, M., Matkovits, A., Benes, E. L., & Jókai, Z. (2024). The role of near-infrared spectroscopy in food quality assurance: A review of the past two decades. *Foods, 13*(21), 3501.

Fuchs, S., Grill, E., Meskiene, I., & Schweighofer, A. (2013). Type 2C protein phosphatases in plants. *The FEBS Journal, 280*(2), 681–693.

Geng, D. H., Liu, L., Lin, Z., Zhu, L., Deng, J., Chen, J., Xiang, Z., Yao, H., Su, X., Xia, C., & Tong, L. T. (2021). Effects of red lentil protein addition on textural quality and starch digestibility of brown rice noodles. *International Journal of Food Science and Technology, 56*(12), 6656–6666.

Grabska, J., Beć, K. B., & Huck, C. W. (2022). Chapter Two - Current and future applications of IR and NIR spectroscopy in ecology, environmental studies, wildlife and plant investigations. *Comprehensive analytical chemistry, 98*, 45–76.

Guzman, M. K. D., Parween, S., Butardo, V. M., Alhambra, C. M., Anacleto, R., Seiler, C., … Sreenivasulu, N. (2017). Investigating glycemic potential of rice by unraveling compositional variations in mature grain and starch mobilization patterns during seed germination. *Scientific Reports, 7*(1), 5854.

Gyani, P. C., Bollinedi, H., Gopala Krishnan, S., Vinod, K. K., Sachdeva, A., Bhowmick, P. K., Ellur, R. K., Nagarajan, M., & Singh, A. K. (2020). Genetic analysis and molecular mapping of the quantitative trait loci governing low phytic acid content in a novel LPA rice mutant, PLM11. *Plants, 9*(12), 1728.

Hajibarat, Z., & Saidi, A. (2022). Senescence-associated proteins and nitrogen remobilization in grain filling under drought stress condition. *Journal of Genetic Engineering and Biotechnology, 20*(1), 101.

Haldipur, A. C., & Srividya, N. (2020). In vitro glycemic response of indigenous pigmented Rice cultivars from South India and influence of different carbohydrate components. *Current Research in Nutrition and Food Science Journal, 8*(3), 815–828.

Huang, M., Liao, C., Xie, J., Chen, J., & Cao, F. (2023). Lysine content and its relationship with protein content in indica rice landraces of China. *Food Chemistry: X, 17*, Article 100549.

Hwang, S. K., Koper, K., Satoh, H., & Okita, T. W. (2016). Rice endosperm starch phosphorylase (Pho1) assembles with disproportionating enzyme (Dpe1) to form a protein complex that enhances synthesis of malto-oligosaccharides. *Journal of Biological Chemistry, 291*(38), 19994–20007.

Islam, M. Z., Arifuzzaman, M., Banik, S., Hossain, M. A., Ferdous, J., Khalequzzaman, M., Pittendrigh, B. R., Tomita, M., & Ali, M. P. (2020). Mapping QTLs underpin nutrition components in aromatic rice germplasm. *PLoS One, 15*(6), Article e0234395.

Ito, V. C. (2023). *Black rice (oryza sativa L.): characterization of extract fractions and evaluation of the effects of ionizing radiation on the composition, physical-chemical, thermal, structural properties and antioxidant activity*.

Jabeen, I., Hafiz, M., Ratnam, W., Karupaiah, T., Dipti, S., Fujita, N., Umemoto, T., Li, Z., & Rahman, S. (2021). Differential expression of three key starch biosynthetic genes in developing grains of rice differing in glycemic index. *Journal of Cereal Science, 99*, Article 103187.

Jin, L., Xiao, P., Lu, Y., Shao, Y., Shen, Y., & Bao, J. (2009). Quantitative trait loci for brown rice color, phenolics, flavonoid contents, and antioxidant capacity in rice grain. *Cereal Chemistry, 86*(6), 609–615.

John, R., Bhardwaj, R., Jeyaseelan, C., Bollinedi, H., Singh, N., Harish, G. D., Singh, R., Nath, D. J., Arya, M., Sharma, D., & Singh, S. (2022). Germplasm variability-assisted near infrared reflectance spectroscopy chemometrics to develop multi-trait robust prediction models in rice. *Frontiers in Nutrition, 2022*, Article 946255.

John, R., Bollinedi, H., Jeyaseelan, C., Padhi, S. R., Sajwan, N., Nath, D., Singh, R., Ahlawat, S. P., Bhardwaj, R., & Rana, J. C. (2023). Mining nutri-dense accessions from rice landraces of Assam, India. *Heliyon, 9*(7).

Jukanti, A. K., Karapati, D., Bharali, V., Gudla, M., Thati, S., Yadla, S., … Sundaram, R. M. (2025). From gene to plate: Molecular insights into and health implications of Rice (Oryza sativa L.) grain protein. *International Journal of Molecular Sciences, 26*(7), 3163.

Kabange, N. R., Dzorkpe, G. D., Park, D. S., Kwon, Y., Lee, S. B., Lee, S. M., … Lee, J. H. (2023). Rice (Oryza sativa L.) grain size, shape, and weight-related QTLs identified using GWAS with multiple GAPIT models and high-density SNP Chip DNA markers. *Plants, 12*(23), 4044.

Karmakar, A., Bhattacharya, S., Sengupta, S., Ali, N., Sarkar, S. N., Datta, K., & Datta, S. K. (2020). RNAi-mediated silencing of ITPK gene reduces phytic acid content, alters transcripts of phytic acid biosynthetic genes, and modulates mineral distribution in rice seeds. *Rice Science, 27*(4), 315–328.

Kumar, A., Sahu, C., Panda, P. A., Biswal, M., Sah, R. P., Lal, M. K., Baig, M. J., Swain, P., Behera, L., Chattopadhyay, K., & Sharma, S. (2020). Phytic acid content may affect starch digestibility and glycemic index value of rice (*Oryza sativa* L.). *Journal of the Science of Food and Agriculture, 100*(4), 1598–1607.

Lee, H. H., Loh, S. P., Bong, C. F. J., Sarbini, S. R., & Yiu, P. H. (2015). Impact of phytic acid on nutrient bioaccessibility and antioxidant properties of dehusked rice. *Journal of Food Science and Technology, 52*, 7806–7816.

Li, M., Liu, X., Bradbury, P., Yu, J., Zhang, Y. M., Todhunter, R. J., … Zhang, Z. (2014). Enrichment of statistical power for genome-wide association studies. *BMC Biology, 12*, 1–10.

Li, P., Chen, Y. H., Lu, J., Zhang, C. Q., Liu, Q. Q., & Li, Q. F. (2022). Genes and their molecular functions determining seed structure, components, and quality of rice. *Rice, 15*(1), 18.

Liu, J., Liu, M., Liu, Y., Jia, M., Wang, S., Kang, X., … Zhou, Z. (2020). Moisture content is a key factor responsible for inducing rice yellowing. *Journal of Cereal Science, 94*, Article 102988.

Lou, G., Bhat, M. A., Tan, X., Wang, Y., & He, Y. (2023). Research progress on the relationship between rice protein content and cooking and eating quality and its influencing factors. *Seed Biology, 2*(1).

Ma, B., Zhang, L., Gao, Q., Wang, J., Li, X., Wang, H., Liu, Y., Lin, H., Liu, J., Wang, X., & Li, Q. (2021). A plasma membrane transporter coordinates phosphate reallocation and grain filling in cereals. *Nature Genetics, 53*(6), 906–915.

Mahender, A., Anandan, A., Pradhan, S. K., & Pandit, E. (2016). Rice grain nutritional traits and their enhancement using relevant genes and QTLs through advanced approaches. *Springerplus, 5*, 1–18.

Mbanjo, E. G. N., Pasion, E. A., Jones, H., Carandang, S., Misra, G., Ignacio, J. C., Kretzschmar, T., Sreenivasulu, N., & Boyd, L. A. (2023). Unravelling marker trait associations linking nutritional value with pigmentation in rice seed. *The Plant Genome, 16*(4), Article e20360.

Mishra, A., Kumar, P., Shamim, M., Tiwari, K. K., Fatima, P., Srivastava, D., … Yadav, P. (2019). Genetic diversity and population structure analysis of Asian and African aromatic rice (Oryza sativa L.) genotypes. *Journal of Genetics, 98*, 1–19.

Ngo, T. V., Kunyanee, K., & Luangsakul, N. (2023). Insights into recent updates on factors and technologies that modulate the glycemic index of rice and its products. *Foods, 12*(19), 3659.

Okpala, N. E., Aloryi, K. D., An, T., He, L., & Tang, X. (2022). The roles of starch branching enzymes and starch synthase in the biosynthesis of amylose in rice. *Journal of Cereal Science, 104*, Article 103393.

Padhi, S. R., John, R., Bartwal, A., Tripathi, K., Gupta, K., Wankhede, D. P., Mishra, G. P., Kumar, S., Rana, J. C., Riar, A., & Bhardwaj, R. (2022). Development and optimization of NIRS prediction models for simultaneous multi-trait assessment in diverse cowpea germplasm. *Frontiers in Nutrition, 9*, Article 1001551.

Pandiselvam, R., Prithviraj, V., Manikantan, M. R., Kothakota, A., Rusu, A. V., Trif, M., & Mousavi Khaneghah, A. (2022). Recent advancements in NIR spectroscopy for assessing the quality and safety of horticultural products: A comprehensive review. *Frontiers in Nutrition, 9*, Article 973457.

Pautong, P. A., Anonuevo, J. J., de Guzman, M. K., Sumayao, R., Jr., Henry, C. J., & Sreenivasulu, N. (2022). Evaluation of in vitro digestion methods and starch structure components as determinants for predicting the glycemic index of rice. *LWT, 168*, Article 113929.

Perera, I., Fukushima, A., Akabane, T., Horiguchi, G., Seneweera, S., & Hirotsu, N. (2019). Expression regulation of myo-inositol 3-phosphate synthase 1 (INO1) in determination of phytic acid accumulation in rice grain. *Scientific Reports, 9*(1), 14866.

Perera, I., Seneweera, S., & Hirotsu, N. (2018). Manipulating the phytic acid content of rice grain toward improving micronutrient bioavailability. *Rice, 11*, 1–13.

Pinta, W., Aninbon, C., Kaewtaphan, P., & Kunyanee, K. (2024). Effects of parboiling on chemical properties, phenolic content and antioxidant capacity in colored landrace Rice. *Foods, 13*(3), 393.

Pradhan, S. K., Pandit, E., Nayak, D. K., Behera, L., & Mohapatra, T. (2019). Genes, pathways and transcription factors involved in seedling stage chilling stress tolerance in indica rice through RNA-Seq analysis. *BMC Plant Biology, 19*(1), 352.

Praphasanobol, P., Purnama, P. R., Junbuathong, S., Chotechuen, S., Moung-Ngam, P., Kasettranan, W., Paliyavuth, C., Comai, L., Pongpanich, M., Buaboocha, T., & Chadchawan, S. (2023). Genome-wide association study of starch properties in local thai rice. *Plants, 12*(18), 3290.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945–959.

Qadir, N., & Wani, I. A. (2023). Functional properties, antioxidant activity and in-vitro digestibility characteristics of brown and polished rice flours of Indian temperate region. *Grain & Oil Science and Technology, 6*(1), 43–57.

Qiu, X., Yang, J., Zhang, F., Niu, Y., Zhao, X., Shen, C., … Xu, J. (2021). Genetic dissection of rice appearance quality and cooked rice elongation by genome-wide association study. *Crop Journal, 9*(6), 1470–1480.

Ren, D., Ding, C., & Qian, Q. (2023). Molecular bases of rice grain size and quality for optimized productivity. *Science Bulletin, 68*(3), 314–350.

Sanghamitra, P., Barik, S. R., Bastia, R., Mohanty, S. P., Pandit, E., Behera, A., Mishra, J., Kumar, G., & Pradhan, S. K. (2022). Detection of genomic regions controlling the antioxidant enzymes, phenolic content, and antioxidant activities in Rice grain through association mapping. *Plants, 11*(11), 1463.

Scott, B.M., Koh, K., & Rix, G.D. (2024). Structural and functional profile of phytases across the domains of life. Current Research in Structural Biology, 7,100139.

Secco, D., Baumann, A., & Poirier, Y. (2010). Characterization of the rice PHO1 gene family reveals a key role for OsPHO1; 2 in phosphate homeostasis and the evolution of a distinct clade in dicotyledons. *Plant Physiology, 152*(3), 1693–1704.

Segura, V., Vilhjálmsson, B. J., Platt, A., Korte, A., Seren, Ü., Long, Q., & Nordborg, M. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature Genetics, 44*(7), 825–830.

Selvaraj, R., Singh, A. K., Singh, V. K., Abbai, R., Habde, S. V., Singh, U. M., & Kumar, A. (2021). Superior haplotypes towards development of low glycemic index rice with preferred grain and cooking quality. *Scientific Reports, 11*(1), 10082.

Shao, Y., Tang, F., Huang, Y., Xu, F., Chen, Y., Tong, C., … Bao, J. (2014). Analysis of genotype× environment interactions for polyphenols and antioxidant capacity of rice by association mapping. *Journal of Agricultural and Food Chemistry, 62*(23), 5361–5368.

Singh, N., Panda, K. K., Bhardwaj, R., Krishnan, S. G., Bhowmick, P. K., Nagarajan, M., … Bollinedi, H. (2024). Exploring the glycemic impact and culinary qualities of rice through genome-wide association studies on starch composition and viscosity profiles. *Food Bioscience, 59*, Article 103919.

Song, J. M., Arif, M., Zhang, M., Sze, S. H., & Zhang, H. B. (2019). Phenotypic and molecular dissection of grain quality using the USDA rice mini-core collection. *Food Chemistry, 284*, 312–322.

Sumargo, F., Gulati, P., Weier, S. A., Clarke, J., & Rose, D. J. (2016). Effects of processing moisture on the physical properties and in vitro digestibility of starch and protein in extruded brown rice and pinto bean composite flours. *Food Chemistry, 211*, 726–733.

Sun, S., Gu, M., Cao, Y., Huang, X., Zhang, X., Ai, P., … Xu, G. (2012). A constitutive expressed phosphate transporter, OsPht1; 1, modulates phosphate uptake and translocation in phosphate-replete rice. *Plant Physiology, 159*(4), 1571–1581.

Suzuki, M., Tanaka, K., Kuwano, M., & Yoshida, K. T. (2007). Expression pattern of inositol phosphate-related enzymes in rice (*Oryza sativa* L.): implications for the phytic acid biosynthetic pathway. *Gene, 405*(1–2), 55–64.

Tamba, C. L., Ni, Y. L., & Zhang, Y. M. (2017). Iterative sure independence screening EM-Bayesian LASSO algorithm for multi-locus genome-wide association studies. *PLoS Computational Biology, 13*(1), Article e1005357.

Takahashi, K., Hirano, T., Kunieda, M., Hirata, Y., Imaizumi, T., & Nishizu, T. (2024). Effect of lipase addition on inhibition of starch retrogradation in rice. *Food Hydrocolloids, 155*, Article 110216.

Tamba, C.L., & Zhang, Y.M. (2018). A fast mrMLM algorithm for multi-locus genome-wide association studies. biorxiv, 341784.doi: https://doi.org/10.1101/341784.

Teye, E., Amuah, C. L., McGrath, T., & Elliott, C. (2019). Innovative and rapid analysis for rice authenticity using hand-held NIR spectrometry and chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 217*, 147–154.

Tibbs Cortes, L., Zhang, Z., & Yu, J. (2021). Status and prospects of genome-wide association studies in plants. *The Plant Genome, 14*(1), Article e20077.

Tp, M. A., Kumar, A., Anilkumar, C., Sah, R. P., Behera, S., & Marndi, B. C. (2022). Understanding natural genetic variation for grain phytic acid content and functional marker development for phytic acid-related genes in rice. *BMC Plant Biology, 22*(1), 446.

Tuano, A. P. P., Castrillo, G. A., & Viola, G. A. V. (2021). Analysis of apparent amylose content of market milled rice via digital image photometry using a smartphone camera. *Current Research in Food Science, 4*, 852. https://doi.org/10.1016/j.crfs.2021.11.011

Turner. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *Journal of Open Source Software, 3*(25), 731.

Verma, R. K., Chetia, S. K., Sharma, V., Baishya, S., Sharma, H., & Modi, M. K. (2022). GWAS to spot candidate genes associated with grain quality traits in diverse rice accessions of Northeast India. *Molecular Biology Reports, 49*(6), 5365–5377.

Wang, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., … Zhang, G. (2012). Control of grain size, shape and quality by OsSPL16 in rice. *Nature Genetics, 44*(8), 950–954.

Wang, X., Xie, L., Fang, J., Pang, Y., Xu, J., & Li, Z. (2023). Identification of candidate genes that affect the contents of 17 amino acids in the rice grain using a genome-wide haplotype association study. *Rice, 16*(1), 40.

Wang, Y., Zheng, Y., Cai, Q., Liao, C., Mao, X., Xie, H., Zhu, Y., Lian, L., Luo, X., Xie, H., & Zhang, J. (2016). Population structure and association analysis of yield and grain quality traits in hybrid rice primal parental lines. *Euphytica, 212*, 261–273.

Wattanavanitchakorn, S., Wansuksri, R., Chaichoompu, E., Kamolsukyeunyong, W., & Vanavichit, A. (2023). Dietary fibre impacts the texture of cooked whole grain rice. *Foods, 12*(4), 899.

Wen, Y., Zhang, Y., Zhang, J., Feng, J., & Zhang, Y. (2020). The improved FASTmrEMMA and GCIM algorithms for genome-wide association and linkage studies in large mapping populations. *The Crop Journal, 8*(5), 723–732.

Wen, Y. J., Zhang, H., Ni, Y. L., Huang, B., Zhang, J., Feng, J. Y., … Wu, R. (2018). Methodological implementation of mixed linear models in multi-locus genome-wide association studies. *Briefings in Bioinformatics, 19*(4), 700–712.

Wu, D., Guo, Z., Ye, J., Feng, H., Liu, J., Chen, G., Zheng, J., Yan, D., Yang, X., Xiong, X., & Liu, Q. (2019). Combining high-throughput micro-CT-RGB phenotyping and genome-wide association study to dissect the genetic architecture of tiller growth in rice. *Journal of Experimental Botany, 70*(2), 545–561.

Wu, J., Zhang, Z. S., Xia, J. Q., Alfatih, A., Song, Y., Huang, Y. J., Wan, G. Y., Sun, L. Q., Tang, H., Liu, Y., & Wang, S. M. (2021). Rice NIN-LIKE PROTEIN 4 plays a pivotal role in nitrogen use efficiency. *Plant Biotechnology Journal, 19*(3), 448–461.

Xia, D., Zhou, H., Wang, Y., Li, P., Fu, P., Wu, B., & He, Y. (2021). How rice organs are colored: The genetic basis of anthocyanin biosynthesis in rice. *The Crop Journal, 9*(3), 598–608.

Xie, J. K., Kong, X. L., Chen, J., Hu, B. L., Wen, P., Zhuang, J. Y., & Bao, J. S. (2011). Mapping of quantitative trait loci for fiber and lignin contents from an interspecific cross *Oryza sativa*× *Oryza rufipogon*. *Journal of Zhejiang University Science B, 12*(7), 518–526.

Xu, F., Bao, J., He, Q., & Park, Y. J. (2016a). Genome-wide association study of eating and cooking qualities in different subpopulations of rice (Oryza sativa L.). *BMC Genomics, 17*(1), 663.

Xu, F., Bao, J., Kim, T. S., & Park, Y. J. (2016b). Genome-wide association mapping of polyphenol contents and antioxidant capacity in whole-grain rice. *Journal of Agricultural and Food Chemistry, 64*(22), 4695–4703.

Yan, M., Xie, M., Chen, W., Si, W. J., Lin, H. H., & Yang, J. (2024). Transcriptome analysis with different leaf blades identifies the phloem-specific phosphate transporter OsPHO1; 3 required for phosphate homeostasis in rice. *The Plant Journal, 118*(3), 905–919.

Yang, H. E., Kim, N. W., Lee, H. G., Kim, M. J., Sang, W. G., Yang, C., & Mo, C. (2024). Prediction of protein content in paddy rice (Oryza sativa L.) combining near-infrared spectroscopy and deep-learning algorithm. *Frontiers in Plant Science, 15*, Article 1398762.

Yang, Y., Guo, M., Sun, S., Zou, Y., Yin, S., Liu, Y., … Yan, C. (2019). Natural variation of OsGluA2 is involved in grain protein content regulation in rice. *Nature Communications, 10*(1), 1949.

Yang, Y., Zhang, Y., Sun, Z., Shen, Z., Li, Y., Guo, Y., Feng, Y., Sun, S., Guo, M., Hu, Z., & Yan, C. (2023). Knocking out OsAAP11 to improve rice grain quality using CRISPR/Cas9 system. *International Journal of Molecular Sciences, 24*(18), 14360.

Yang, Z., Wang, X., Gu, S., Hu, Z., Xu, H., & Xu, C. (2008). Comparative study of SBP-box gene family in *Arabidopsis* and rice. *Gene, 407*(1–2), 1–11.

Yoshida, H., Okada, S., Wang, F., Shiota, S., Mori, M., Kawamura, M., Zhao, X., Wang, Y., Nishigaki, N., Kobayashi, A., & Miura, K. (2023). Integrated genome-wide differentiation and association analyses identify causal genes underlying breeding-selected grain quality traits in japonica rice. *Molecular Plant, 16*(9), 1460–1477.

Zhang, J., Feng, J. Y., Ni, Y. L., Wen, Y. J., Niu, Y., Tamba, C. L., Yue, C., Song, Q., & Zhang, Y. M. (2017). pLARmEB: integration of least angle regression with empirical Bayes for multilocus genome-wide association studies. *Heredity, 118*(6), 517–524.

Zhang, Y. W., Tamba, C. L., Wen, Y. J., Li, P., Ren, W. L., Ni, Y. L., … Zhang, Y. M. (2020). mrMLM v4. 0.2: an R platform for multi-locus genome-wide association studies. *Genomics, Proteomics & Bioinformatics, 18*(4), 481–487.

Zhang, Y., Zhang, S., Zhang, J., Wei, W., Zhu, T., Qu, H., … Xu, G. (2023). Improving rice eating and cooking quality by enhancing endogenous expression of a nitrogen-dependent floral regulator. *Plant Biotechnology Journal, 21*(12), 2654–2670.

Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., … Buckler, E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics, 42*(4), 355–360.

Zhong, H., Liu, S., Zhao, G., Zhang, C., Peng, Z., Wang, Z., Yang, J., & Li, Y. (2021). Genetic diversity relationship between grain quality and appearance in rice. *Frontiers in Plant Science, 12*, Article 708996.

Zhou, H., Wang, L., Liu, G., Meng, X., Jing, Y., Shu, X., … Wu, D. (2016). Critical roles of soluble starch synthase SSIIIa and granule-bound starch synthase waxy in synthesizing resistant starch in rice. *Proceedings of the National Academy of Sciences, 113*(45), 12844–12849.