



OPEN ACCESS

EDITED BY

Nimai Prasad Mandal,
NRRI-Central Rainfed Upland Rice Research
Station (CRURRS), India

REVIEWED BY

Jindong Liu,
Chinese Academy of Agricultural Sciences,
China
Munmi Phukon,
Assam Agricultural University, India

*CORRESPONDENCE

Faliang Zeng
✉ zeng9210@163.com
Faming Yu
✉ yfm4679@163.com

[†]These authors have contributed equally to
this work

RECEIVED 28 August 2025

REVISED 14 November 2025

ACCEPTED 17 November 2025

PUBLISHED 11 December 2025

CITATION

Yang L, Yang R, Wang Y, Asad MAU, Wang Y,
Chang Z, Miao T, Yang S, Wei Y, Wu S, Bao J,
Wu M, Ye J, Zhai R, Ye S, Zhang X, Zeng F and
Yu F (2025)

Integrative genomic and transcriptomic
approaches decipher pre-harvest
sprouting resistance in rice.

Front. Plant Sci. 16:1692942.

doi: 10.3389/fpls.2025.1692942

COPYRIGHT

© 2025 Yang, Yang, Wang, Asad, Wang, Chang,
Miao, Yang, Wei, Wu, Bao, Wu, Ye, Zhai, Ye,
Zhang, Zeng and Yu. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](#).
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.

Integrative genomic and transcriptomic approaches decipher pre-harvest sprouting resistance in rice

Lv Yang^{1†}, Renyuan Yang^{1†}, Yuxuan Wang^{2†},
Muhammad Asad Ullah Asad¹, Yueying Wang³, Zhiyuan Chang²,
Tianhui Miao⁴, Shudong Yang⁴, Yiting Wei⁵, Shanshan Wu¹,
Jiaxue Bao¹, Mingming Wu¹, Jing Ye¹, Rongrong Zhai¹,
Shenghai Ye¹, Xiaoming Zhang¹, Faliang Zeng^{6*}
and Faming Yu^{1*}

¹Institute of Crop and Nuclear Technology Utilization, Zhejiang Academy of Agricultural Science, Hangzhou, China, ²Wenzhou Vocational College of Science and Technology, Wenzhou, China, ³State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou, China, ⁴Shengzhou Seed Multiplication Farm, Shengzhou, China, ⁵Longyou Agriculture and Rural Affairs Bureau, Quzhou, China, ⁶Jiangxi Academy of Agricultural Sciences Rice Research Institute, Nanchang, China

Pre-harvest sprouting (PHS) seriously compromises rice yield and quality, increase susceptibility to insect pest and reduce seed viability. Beside agronomic control measures, the genetic makeup of rice plants serves as a fundamental determinant in conferring resistance to PHS. Therefore, integrating multi-omics strategies to construct high-resolution genetic variation maps, screen extreme-phenotype germplasm, and identify causal genes are pivotal for generating PHS-resistant breeding material. In this study, we performed whole-genome re-sequencing of 165 highly diverse indica rice accessions to construct a high-density genetic variation map, obtaining a dataset comprising 1,584,905 high-quality SNPs for subsequent association analysis. Genome-wide association studies (GWAS) further uncovered 21 candidate loci and multiple candidate genes associated with PHS, from which key candidate genes were prioritized. In particular, previously cloned PHS-related genes—*OsCDP3.10*, *OsWRKY50*, *UGT74J1*, *OsJAZ6*, and *IPA1*. Additionally, we investigated the transcriptional analyses in cultivars Z33 and Z216 under high-humidity conditions and identified 19,087 differentially expressed genes (DEGs). Notably, by integrating GWAS and transcriptomic analyses, we identified *UGT74J1* as a promising candidate gene, and haplotype analysis further revealed *UGT74J1*-Hap3 as a superior haplotype associated with PHS resistance. This multi-omics dataset and the candidate genes identified will provide valuable genetic resources for molecular breeding toward improved PHS resistance in rice.

KEYWORDS

rice, pre-harvest sprouting, GWAS, transcriptome analysis, molecular breeding

Introduction

Pre-harvest sprouting (PHS) is a phenomenon where mature grain crops may sprout prematurely in their panicles if not harvested in time, especially under high-temperature and high-humidity conditions (Chen et al., 2021). With global climate change, the frequency of extreme weather events has increased, leading to more unpredictable rainfall and high-humidity conditions during harvest (Tai et al., 2021). Consequently, the incidence of PHS in rice (*Oryza sativa* L.) has significantly increased. This phenomenon directly diminishes grain yield and quality, leading to substantial economic losses and jeopardizing food security, especially in humid rice-growing areas.

Seed dormancy is a key agronomic trait governing PHS resistance in rice. Varieties with strong dormancy can maintain a quiescent state for an extended period after seed maturation, which prevents germination under permissive conditions and thereby confers enhanced PHS resistance. A central challenge in contemporary rice breeding lies in orchestrating a precise equilibrium between robust seed dormancy and rapid, uniform germination (Holdsworth et al., 2008). Consequently, the strategic development of novel cultivars endowed with an optimal dormancy profile is imperative for the dual objectives of sustaining yield potential and preserving grain quality. This complex agronomic trait is governed by a sophisticated molecular regulatory network (Finkelstein et al., 2008; Gubler et al., 2005). The regulatory mechanisms governing seed dormancy and seed vigor critically influence crop yield and grain quality through complex interactions involving phytohormonal signaling, environmental stimuli, and seed physiological status. Several phytohormones such as Absciscic acid (ABA), jasmonic acid (JA), and brassinosteroids (BR) serve as pivotal regulators orchestrating dormancy establishment and sustainability (Khan, 2025; Thilakarathne et al., 2025). These hormones coordinate intra-seed signaling networks to precisely modulate dormancy depth and germination competency (Wang et al., 2020a). ABA suppresses precocious germination under adverse conditions by inducing stomatal closure to minimize water dissipation, thereby maintaining seed desiccation status (Chen et al., 2022; Fang et al., 2008). JA potentiates dormancy intensity, enabling seeds to retain quiescence even in permissive environments, consequently enhancing seed viability and ecological adaptability (Dave et al., 2011; Liu et al., 2019). BR dynamically regulate the dormancy-germination transition through modulation of endogenous hormonal homeostasis and acceleration of water uptake dynamics (Xiong et al., 2022).

The molecular architecture governing pre-harvest sprouting constitutes an intricate regulatory network entailing synergistic interactions among multiple genetic components (Wang et al., 2020b). *OsMFT2* (Mother of FT and TFL2) belongs to phosphatidylethanolamine-binding protein (PEBP) family and regulate flowering time and seed dormancy. *OsMFT2* exhibits seed-specific expression in rice and enhances ABA signaling through interacting with downstream proteins and regulating expression and activity of transcription factors (TFs) *OsbZIP23*, *OsbZIP66*, and *OsbZIP72*, thereby regulating seed dormancy and

germination processes (Song et al., 2020). The C2H2-type zinc finger protein encoded by *OsZFP15* accelerates seed germination via regulating the increase of ABA catabolism (Wang et al., 2022). Under ABA accumulation, *SAPK10* phosphorylates TF *bZIP72* at Ser71, which both stabilizes the *bZIP72* protein and augments its binding affinity for the promoter of AOC, —a key JA biosynthesis gene (Kobayashi et al., 2005; Kobayashi et al., 2004). This regulatory cascade consequently elevates endogenous JA levels and suppresses germination (Wang et al., 2022). Moreover, loss activity of *OsBZR1*, a core component of BR signaling, *OsBZR1* due to functional mutations, moderately attenuate coleoptile elongation but significantly enhance pre-harvest sprouting resistance through delayed germination (Xiong et al., 2022).

Identification of PHS-resistant genes, superior haplotypes, and associated markers are the crucial objectives in rice breeding to enhance crop resilience and reduce grain yield losses (Aloryi et al., 2025). High-throughput functional marker-assisted selection enables the precise identification of resistant germplasm and significantly boost breeding efficiency (Robinson et al., 2010). This approach accelerates the introgression of favorable alleles, enabling the rapid development of broadly adaptable, high-yielding rice varieties with enhanced stability against environmental adversities (Yi et al., 2025). The development of such varieties is vital for ensuring global food security in the face of increasing population (Zhang et al., 2020; Zhu et al., 2019). Meanwhile, multidimensional regulatory networks of PHS obtained by integrating genomic, transcriptomic, and population analyses, not only advances gene discovery and functional validation but also enhances our mechanistic understanding of PHS (Min et al., 2024). Consequently, this PHS regulatory network provides a robust theoretical foundation for genomic selection and intelligent design in rice breeding.

The integration of multi-omics analyses utilizing high-density SNPs in plant populations has emerged as a robust strategy for identifying candidate genes (Lv et al., 2022). The advent of methodologies such as GWAS, transcriptomics, metabolomics, and proteomics has provided powerful tools for systematically elucidating the molecular mechanisms underlying PHS resistance. Compared to traditional linkage analysis, GWAS can detect a broader spectrum of alleles and genetic variations (Min et al., 2021). For instance, the combined use of GWAS and linkage mapping identified *Sdr4* as a key regulator of PHS in rice. This gene exhibits additive effects with *Rc* and *SD1*, its genotypic distribution shows significant association with climatic conditions, and molecular markers have been developed for breeding applications (Zhao et al., 2022). Furthermore, integrated transcriptomic and metabolomic analyses revealed that the *SbPP2C33* gene influences seed germination in sorghum by modulating ABA signaling and carbohydrate mobilization, providing new insights into the molecular regulatory network of PHS (Ju et al., 2025). In this study, whole-genome resequencing of 165 diverse indica rice accessions was conducted to generate a high-density genetic variation map from high-quality SNPs. By integrating GWAS and transcriptome profiling, we identified multiple significant PHS-associated loci and candidate QTLs.

Further multi-omics integration revealed critical PHS-related regulatory pathways and differentially expressed genes. Our finding establishes a comprehensive molecular framework for understanding PHS and provide valuable genetic targets and technical foundations for breeding PHS-resistant rice varieties.

Materials and methods

Plant materials

The indica population was selected for its high genetic diversity and direct agricultural relevance. An assemblage of 165 *indica* accessions was established for this study (Supplementary Table 1). These genetic resources are maintained at the Zhejiang Academy of Agricultural Sciences (ZAAS, Hangzhou, Zhejiang Province, China) and cultivated in ZAAS experimental fields.

Statistical analysis of phenotypic variation

A panel of 165 *indica* rice varieties was assessed for panicle sprouting evaluation in 2024. Main panicles were collected from mature plants at harvest, 35 days after heading. For each variety, main panicles from three plants were randomly sampled as biological replicates. The collected panicles were soaked in water for 3 hours, then drained, gently dried, and transferred to the artificial climate chamber at ZAAS under a 12/12-hour light/dark cycle with temperatures of 30 °C and 25 °C, respectively. The relative humidity was set to > 90% for the high-humidity treatment group and 70% for the control group (Liu et al., 2022; Zhao et al., 2022). Germination rates, indicated by the emergence of white sprouts, were recorded at 24-hours interval. The final result for each accession was calculated as the average across the three replicates. The descriptive statistical analysis was performed in R 4.0.3 software. Based on the phenotypic data, the traits were validated for the analysis by GraphPad Prism.9 software. The significant differences between the control and treatment of the samples were analyzed by Student's *t* test. The results with $P < 0.05$ and $P < 0.01$ were considered statistically significant and denoted by one star and two stars, respectively.

Variant calling

For each accession, DNA sequencing was carried out on a single individual. Total DNA was extracted from the leaves of one-month-old rice plants and subjected to genome sequencing with 15× coverage on the DNBSEQ-T7 platform. Fastp (v0.23.2) with default parameters was used for sequencing data quality control (Chen, 2023). The processed reads were aligned to the Nipponbare reference genome (MSU v7.0) using bwa (v0.7.17-r1188) (mem-M-R-K 10000000). samtools (v1.9) facilitated format conversion and index construction (Li et al., 2009). Variant calling employed GATK4 (HaplotypeCaller-ERC GVCF), with CombineGVCFs for

merging gvcf files and GenotypeGVCFs for genotyping (Heldenbrand et al., 2019). Quality control of single nucleotide polymorphisms (SNPs) utilized GATK4's variant filter (QD2, QUAL30, FS60, MQ40, MQRankSum-12.5, ReadPosRankSum-8). SNP filtering was accomplished by vcftools (maf 0.05, max missing 0.8) (Danecek et al., 2011).

Population structure analysis and LD analysis

A total of 1,584,905 high-quality SNPs were filtered and used for population structure analysis. SnpEff (V4.3m) was employed for functional annotation of the population genotypes (Kawahara et al., 2013). Based on SNP site information, population principal component analysis (PCA) was performed using plink (v1.90b6.21), population structure was assessed using mixture v1.3.0 (Alexander and Lange, 2011), and a phylogenetic tree was constructed using iqtree v2.2.5 (Minh et al., 2020). Visualization was carried out using iTol (<http://itol.embl.de/>).

Linkage disequilibrium (LD) decay analysis was performed for the entire population using PopLDdecay with the following parameters: -MaxDist 500, -MAF 0.05, -Het 0.88, and -Miss 0.999 (Zhang et al., 2019).

Genome-wide association study

The PHS traits were measured. For genome-wide association analysis (GWAS) of 165 *indica* accessions, we used the Farm CPU, MLM, and MLMM (PC = 5, vc.method="BRENT", maxLoop=20, method.bin="static", priority="speed") models from the R package rMVP (Kang et al., 2010). To correct for population structure, PCA was first conducted using PLINK, and the top five PCs were incorporated as covariates. Manhattan and Quantile-Quantile (Q-Q) plots illustrating the GWAS results were generated using the R package 'qqman' (v1.4.5) in R version 3.4.2. A total of 1,584,905 high-quality SNPs with missing data ≤20% and MAF ≥0.05 were selected (Visscher et al., 2012). The significant association threshold was $-\lg(p) = 5$ (Gao et al., 2020; Lv et al., 2022; Wei et al., 2021). Based on the linkage disequilibrium (LD) decay threshold ($r^2 = 0.2$, Distance =250Kb), candidate genes were defined as those flanking 200 kb upstream and downstream of the significant SNPs (Kang et al., 2010; Klinsawang et al., 2024; Lv et al., 2021).

Transcriptome analysis

For transcriptomic analysis, seeds from cultivars Z216 and Z33 were collected from spikes at 24 hours post-germination. As described in the phenotypic investigation, samples from three experimental replicates were pooled to form three biological replicates per cultivar. Total RNA was extracted from these pooled samples using TRIzol reagent. Based on RNA integrity (RIN >6) and purity (concentration >50 ng/μL), per sample were

selected for subsequent RNA-seq library construction and sequencing (Lv et al., 2022; Wei et al., 2024). Raw data were quality-controlled with fastp (v0.23.2). The processed reads were aligned to the Nipponbare reference genome (Msu7.0) using HISAT2 (v2.2.1) (Kim et al., 2015) with parameters (-dta -t -p 1) and indexed with samtools (v1.9) (Li et al., 2009). Gene expression was normalized using FPKM (Trapnell et al., 2012). featureCounts (v2.0.1) quantified reads with parameters (-T 2 -p -t exon -g gene_id) (Liao et al., 2014).

Differentially expressed genes (DEGs) between the control and treatment groups of cultivars Z216 and Z33 were identified using the edgeR package (Robinson et al., 2010). DEGs were identified based on stringent thresholds (CPM >1, bcv = 0.2, P -value <0.01, $|\log_2FC| \geq 2$) (Robinson et al., 2010). Genes consistently differentially expressed in both replicates were considered core DEGs and annotated using TBtools through GO and KEGG analysis (Chen et al., 2020).

Haplotype analysis

Candidate genes were annotated with snpEff (v4.3m) based on the variation map of 165 samples. Variants annotated as missense, frameshift, and stop-gain mutations were selected, followed by further screening for genotype differences within the coding regions of the corresponding genes. Haplotypes were then correlated with phenotypes to identify superior haplotypes (Hothorn et al., 2008). The significant differences between different haplotypes were analyzed by Student's t test. The results with $P < 0.05$ and $P < 0.01$ were considered statistically significant and denoted by one star and two stars, respectively (Wei et al., 2024; Lv et al., 2022).

Results

Evaluation of PHS phenotypes

We evaluated the pre-harvest sprouting (PHS) phenotypes in 165 *indica* accessions to identify germplasms with extreme PHS traits and to uncover novel candidate genes with PHS resistance. Seeds were exposed to a controlled artificial climate chamber simulating a high-humidity maturation environment. The dynamic progression of pre-harvest sprouting (PHS) was monitored at 24-hour intervals. Based on the calculated panicle germination rate, cultivars were categorized by PHS severity. This classification revealed 18 susceptible cultivars (80–100% germination), such as Ai Bai Qiu and Zhe Zao 33, and 64 resistant cultivars (0–20% germination), including Guang Xuan 3 Hao (Figure 1A). The results indicated significant variation in PHS susceptibility among the accessions (Figure 1B).

Genomic variation and population structure analysis

To identify the variants and candidate genes associated with PHS, we re-sequenced 165 *indica* accessions, generating 799 Gb total sequencing length with an average depth of $\sim 13\times$. After evaluating/filtering, we obtained 1,584,905 high-quality SNPs covering the whole genome, with 4.25 SNPs per kb on average (Figure 2). SNP distribution showed that 22.32% were in intergenic regions, 63.11% within 2 kb of transcription start/end sites, and 14.56% in gene bodies. Chromosome 1 had the most SNPs (176,605 sites), while chromosome 9 had the fewest (97,806 sites). These sites provide valuable data sets that significantly advances rice biology and breeding research.

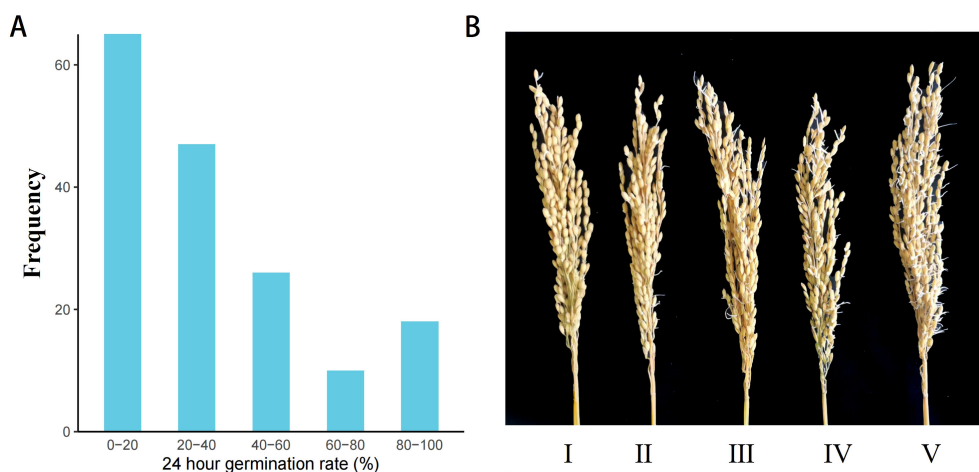
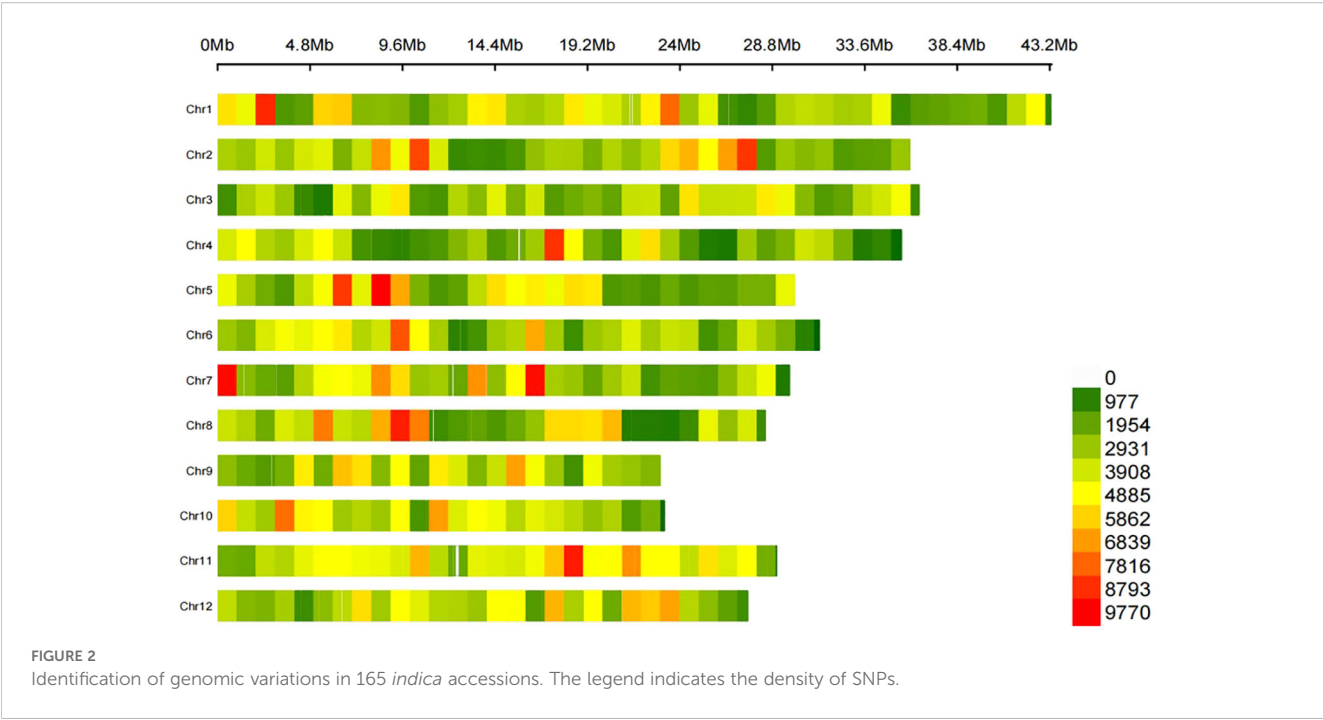


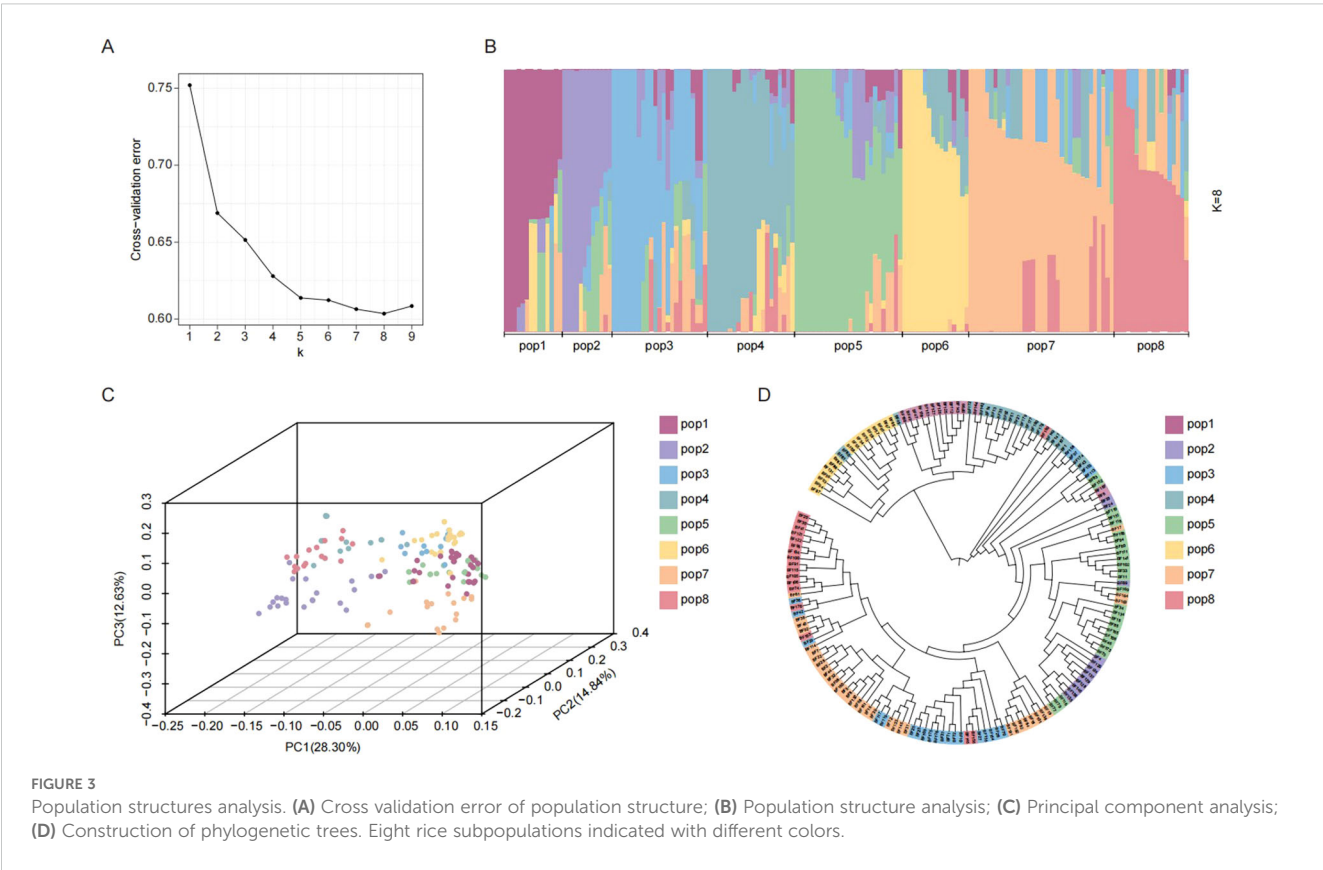
FIGURE 1

Comparative phenotypic difference of PHS in 165 *indica* accessions. (A) Frequency distribution of 24-hour PHS rate in the population; (B) Phenotypic characteristics of varieties with different PHS types. Varieties were classified into PHS types I–V based on the spike germination rate, with each grade representing an approximately 20% increment.



Using filtered SNPs, we analyzed the population structure of 165 *indica* accessions to assess genetic diversity and relationships. Optimal population structure analysis ($K = 8$) classified the 165 *indica* accessions into eight distinct subpopulations (pop1–pop8) (Figures 3A, B). Given the minor structural differences within the

population, we further assessed it using SNP-based principal component analysis (PCA). The results showed that the first 3 principal components among the top 10 only explained 55.77% of the variation, while principal components 4 to 10 had similar variation explanation rates. This indicates that there are no single



or few dominant genetic variation dimensions that can strongly differentiate the subpopulations within this group (Figure 3C). Phylogenetic reconstruction using the NJ method also resolved eight primary clusters (Figure 3D). This observed multi-dimensionality underscores the complex and diverse genetic architecture of the rice population.

GWAS

A total of 21 significant loci were identified through GWAS for PHS (Figure 4). The MLM, MLM, and FarmCPU models contributed 17, 3, and 1 locus, respectively. Importantly, three loci were co-detected by at least two models, highlighting their consistent genetic effects. This multi-model approach demonstrates the potential to uncover a comprehensive set of candidate genes associated with PHS.

Z216 and Z33 exhibit differences in PHS phenotypes

We further evaluated the PHS-resistance level of 165 rice accessions under high-humidity conditions and found significant differences in PHS levels between Z33 and Z216 cultivars (Figure 5A). At 24 hours, Z216 had an average germination rate of 24%, compared to Z33's significantly higher rate of 90% (Figure 5B). This suggests that Z216 cultivar is more resistant to panicle sprouting, while Z33 cultivar is more sensitive.

RNA-seq and DEGs analysis

To explore the expression of PHS regulated genes in rice, we conducted transcriptome sequencing on Z33 and Z216 cultivars following 24-hour high-humidity treatment. Through FPKM

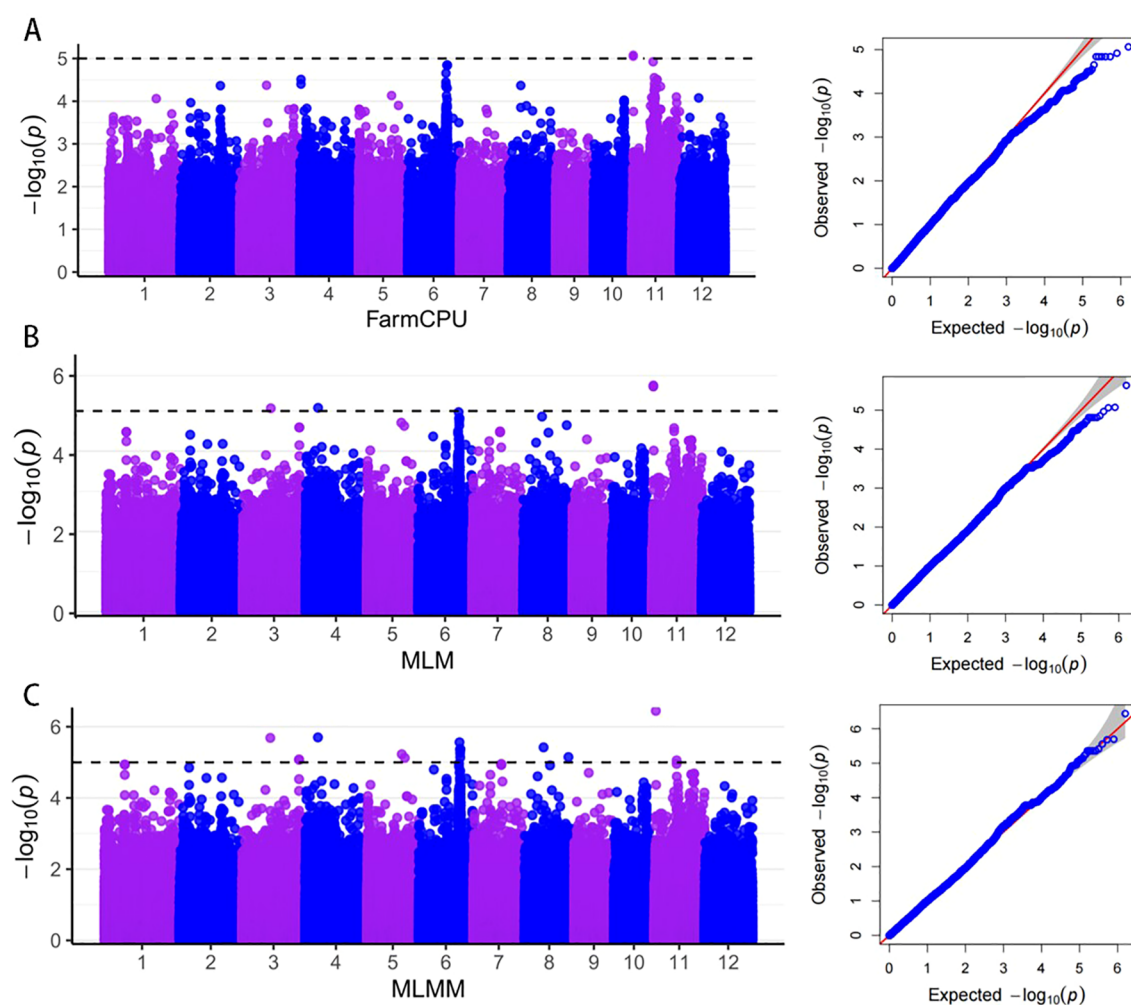


FIGURE 4
Genome-wide association analysis based on PHS rates. (A–C) are Manhattan and QQ plots of GWAS FarmCPU, MLM, and MLM models.

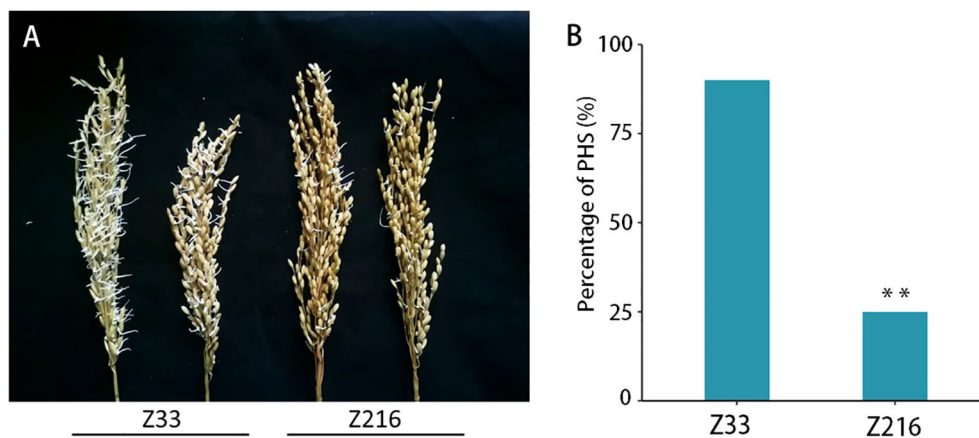


FIGURE 5

Comparison of the PHS phenotype in Z33 and Z216 cultivars under high humidity conditions. (A) PHS phenotype in Z33 and Z216 lines; (B) Comparison of PHS after 24 hours of high humidity treatment ($n=5$, T-test, $**P < 0.01$).

analysis, we detected a total of 25,776 expressed genes. Hierarchical clustering revealed distinct transcriptomic differences between the two materials (Figure 6A). A total of 19,087 DEGs were identified (Figure 6B), with Z33 having 3,975 upregulated and 7,213 down-regulated DEGs, and Z216 having 3,105 upregulated and 4,794 down-regulated DEGs. These DEGs overlapped with 21 previously reported PHS-related genes.

DEG analysis further highlighted the top three enriched GO terms as metabolic process (GO:0008152), stimulus response (GO:00508962), and membrane system (GO:0016020). The most enriched KEGG pathways were starch and sucrose metabolism (A09100), plant hormone signal transduction (B09183), and phenylpropanoid biosynthesis (B09101) (Figures 6C, D). The upregulated amylase genes (e.g., *LOC_Os08g36910*) accelerate endosperm breakdown to

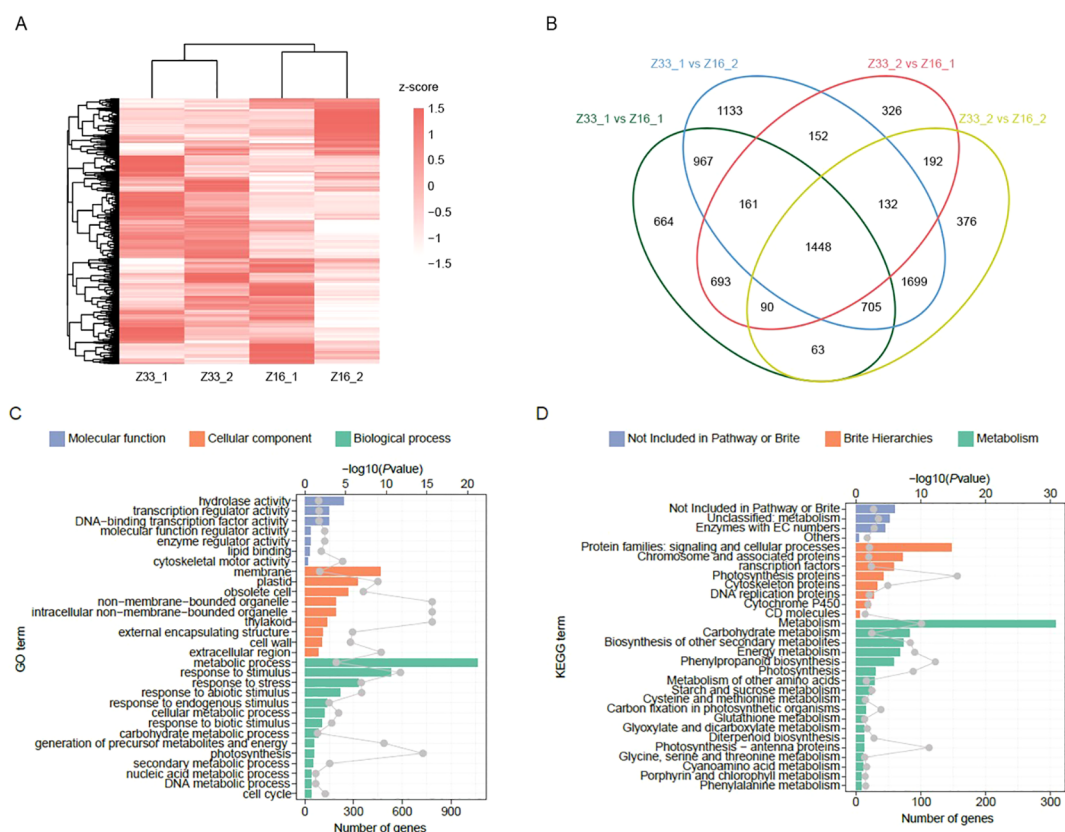


FIGURE 6

Transcriptome analysis of PHS between Z33 and Z216 cultivars. (A) Gene expression profile; (B) DEG analysis of Z33 and Z216 cultivars; (C) GO classification and (D) KEGG classification of Z33 and Z216.

supply energy and carbon skeletons for germination, while differentially expressed ABA-related genes (e.g., *LOC_Os01g02120*, *LOC_Os01g64000*) can change the hormone balance. Analyzing the regulatory network offers new insights into PHS mechanisms, provides key targets for molecular breeding, and improves breeding efficiency for panicle sprouting resistance.

Candidate gene identification and haplotype analysis

GWAS analysis has provided a new dataset of candidate genes for PHS-related, defining twenty-one candidate QTLs. These intervals contain several previously cloned PHS-related genes: *OsCDP3.10* (*LOC_Os03g57960*; cupin domain-containing protein; Huang et al., 2021), *OsWRKY50* (*LOC_Os11g02540*; transcription factor repressor; Peng et al., 2022), *UGT74J1* (*LOC_Os04g12980*; UDP-glucosyltransferase; Tezuka et al., 2021), *OsJAZ6* (*LOC_Os03g28940*; jasmonate ZIM-domain protein; Wang et al., 2024), and *IPA1* (*LOC_Os08g39890*; ideal plant architecture gene; Chen et al., 2023). Furthermore, integrating transcriptomic analysis facilitated further gene mining, revealing an overlap between differentially expressed genes (DEGs) and GWAS candidate genes. Specifically, *UGT74J1*, encoding a UDP-glucosyltransferase, was identified as a candidate gene on chromosome 4 through both MLM and MLM models. Haplotype analysis, defined by a missense mutation in the coding region, identified *UGT74J1*-Hap3 as a superior haplotype conferring PHS resistance (Figure 7), highlighting the functional relevance of this coding sequence variation and underscoring the value of *UGT74J1* as a promising candidate for PHS-related marker development.

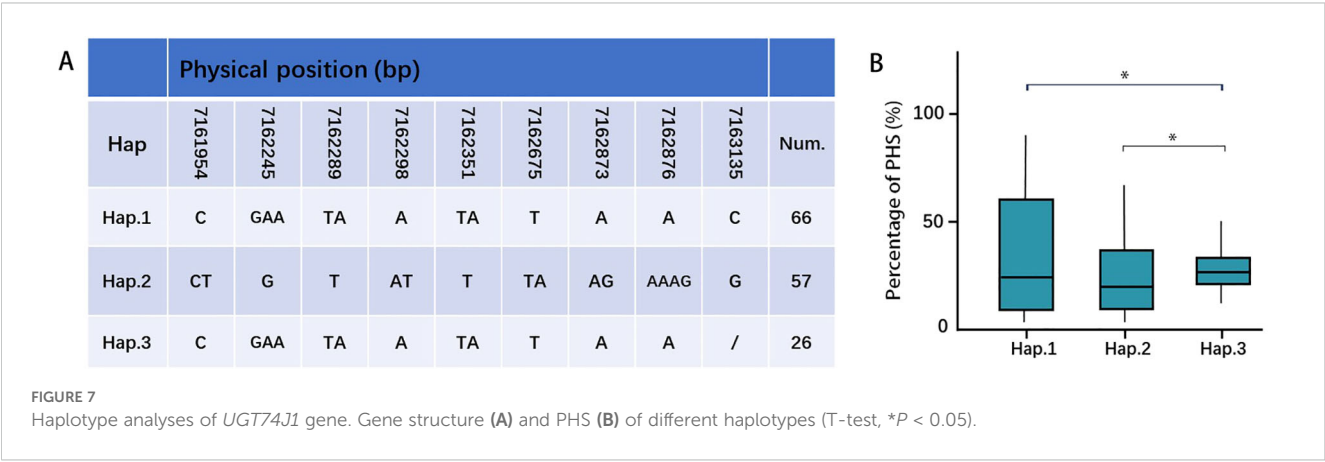
Discussion

Pre-harvest sprouting, the premature germination of rice grains in humid conditions before harvest, compromises grain quality and yield while elevating susceptibility to mold and pests, posing a threat to global food security (Lee et al., 2021; Sohn et al., 2021). Given that PHS susceptibility is influenced by both environmental factors and genetic background (Chen et al., 2021), identifying

resistance-associated genetic variations across diverse germplasms offers significant potential for breeding improvement (He et al., 2021). This study therefore aimed to identify PHS-associated variants, candidate genes, and regulatory pathways to facilitate the development of resilient rice varieties.

We selected a population comprising *indica* accessions based on their high genetic diversity and agricultural relevance. *Indica* constitutes a predominant portion of worldwide rice output and is largely cultivated in the warm and humid environments of Asia, a climatic profile that strongly predisposes the crop to PHS. Nevertheless, the genetic foundations of PHS resistance in *indica* have been less thoroughly investigated relative to *japonica*. Therefore, the genetic enhancement of PHS resistance in *indica* emerges as a pressing breeding goal (Xu et al., 2025). We evaluated PHS susceptibility in a diverse panel of 165 *indica* rice accessions under controlled high-humidity conditions. Phenotypic assessment revealed a broad spectrum of PHS susceptibility, with cultivars Z33 and Z216 identified as highly susceptible and resistant extremes, respectively, and selected for further investigation. This careful phenotyping not only establishes a reliable foundation for subsequent genomic analyses but also identifies valuable germplasm donors, such as Z216 and Guang Xuan 3 Hao, which can be directly utilized in crossing programs to introduce PHS resistance into elite genetic backgrounds.

Whole-genome resequencing generated a high-density variation map of over 1.58 million high-quality SNPs. This dataset itself constitutes a valuable community resource, enabling future genetic studies and marker development in *indica* rice. GWAS using the MLM, MLM, and FarmCPU models identified 21 significant loci associated with PHS. The co-detection of three loci by at least two independent models underscores the reliability of these associations. The integration of GWAS with transcriptomic data significantly refined our candidate gene list. This multi-omics strategy effectively shifted the focus from large genomic intervals to high-probability candidates, thereby enhancing the efficiency of gene discovery. Notably, several previously cloned PHS-related genes, including *OsCDP3.10* (Huang et al., 2021), *OsWRKY50* (Peng et al., 2022), *UGT74J1* (Tezuka et al., 2021), *OsJAZ6* (Wang et al., 2024), and *IPA1* (Chen et al., 2023), were located within our candidate intervals, which cross-validates the effectiveness of our approach. Among the novel candidates, *UGT74J1* emerged as a high-priority gene, being



highlighted by both GWAS and differential expression analysis. The identification of *UGT74J1*-Hap3 as a superior haplotype associated with PHS resistance has direct and immediate implications for molecular breeding. This finding allows for the development of a co-dominant, functional marker based on the haplotype-specific SNPs. This marker can be used in MAS to rapidly introgress the resistant *UGT74J1*-Hap3 allele into susceptible elite varieties, significantly accelerating the breeding of PHS-resistant cultivars. This strategy is particularly crucial for *indica* rice improvement, as it enables the enhancement of a complex trait without compromising on other agronomic traits.

Transcriptomic profiling under PHS-inducing conditions revealed 19,087 DEGs in Z33 and Z216, revealing a complex regulatory network and highlighting significant enrichment in critical metabolic and stimulus-response pathways. Beyond confirming known pathways, our transcriptome data revealed specific regulatory nodes, such as the coordinated expression of amylase and ABA-related genes. This detailed view of the PHS-related regulatory network provides a theoretical foundation for pyramiding multiple favorable alleles across different pathways, which is a promising strategy for breeding durable and broad-spectrum PHS resistance.

In conclusion, this study elucidates the genetic architecture of pre-harvest sprouting resistance in *indica* rice through an integrated genomic, transcriptomic, and population genetic framework. Our findings provide actionable resources for breeding, including a high-density SNP dataset for genomic selection, superior haplotypes for marker-assisted selection, and characterized resistant germplasm. The deployment of these resources will accelerate the development of high-yielding, climate-resilient rice varieties, thereby contributing to global food security.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#), further inquiries can be directed to the corresponding author/s. Genome data in this study have been deposited in the China National Genomics Data Center (<https://ngdc.cncb.ac.cn/>) under the project accession numbers PRJCA047790.

Author contributions

LY: Writing – original draft, Data curation, Conceptualization. RY: Software, Investigation, Writing – review & editing. YXW: Writing – review & editing, Data curation, Methodology. MAUA: Funding acquisition, Project administration, Formal Analysis, Writing – original draft, Visualization. YYW: Formal Analysis, Writing – review & editing, Investigation, Software. ZC: Project administration, Formal Analysis, Writing – review & editing. TM: Formal Analysis, Writing – review & editing, Data curation. SDY: Methodology, Project administration, Writing – review & editing. YTW: Methodology, Validation, Supervision, Writing – review & editing. SW: Formal Analysis, Writing – review & editing, Methodology. JB: Writing –

review & editing, Methodology, Project administration. MW: Methodology, Supervision, Writing – review & editing. JY: Formal Analysis, Data curation, Methodology, Writing – review & editing. RZ: Project administration, Supervision, Writing – review & editing. SHY: Methodology, Validation, Writing – review & editing. XZ: Methodology, Writing – review & editing, Investigation, Project administration. FZ: Funding acquisition, Formal Analysis, Methodology, Writing – review & editing, Data curation. FY: Investigation, Project administration, Funding acquisition, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by grants from the Zhejiang Natural Science Foundation project (LQN25C130002); Major Science and Technology project of Zhejiang Province (2021C02063-4); Natural Science Foundation of Jiangxi Province funding project (20242BAB20284); National Natural Science Foundation of China (W2433049) China Postdoctoral Science Foundation (2024M763601).

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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.2025.1692942/full#supplementary-material>

References

- Alexander, D. H., and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinf.* 12, 246. doi: 10.1186/1471-2105-12-246
- Aloryi, K. D., Okpala, N. E., Amenyogbe, M. K., Bimpong, D., Karikari, B., Guo, H., et al. (2025). Whole-genome meta-analysis coupled with haplotype analysis reveal new genes and functional haplotypes conferring pre-harvest sprouting in rice. *BMC Plant Biol.* 25, 527. doi: 10.1186/s12870-025-06551-5
- Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009
- Chen, F., Zhang, H., Li, H., Lian, L., Wei, Y., Lin, Y., et al. (2023). IPA1 improves drought tolerance by activating SNAC1 in rice. *BMC Plant Biol.* 23, 55. doi: 10.1186/s12870-023-04062-9
- Chen, S. (2023). Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *Imeta* 2, e107. doi: 10.1002/imt2.107
- Chen, W., Wang, W., Lyu, Y., Wu, Y., Huang, P., Hu, S., et al. (2021). OsVP1 activates Sdr4 expression to control rice seed dormancy via the ABA signaling pathway. *Crop J.* 9, 68–78. doi: 10.1016/j.cj.2020.06.005
- Chen, Y., Xiang, Z., Liu, M., Wang, S., Zhang, L., Cai, D., et al. (2022). ABA biosynthesis gene OsNCED3 contributes to preharvest sprouting resistance and grain development in rice. *Plant Cell Environ.* 46, 1384–1401. doi: 10.1111/pce.14480
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330
- Dave, A., Hernandez, M. L., He, Z., Andriotis, V. M., Vaistij, F. E., Larson, T. R., et al. (2011). 12-oxo-phytodienoic acid accumulation during seed development represses seed germination in Arabidopsis. *Plant Cell* 23, 583–599. doi: 10.1105/tpc.110.081489
- Fang, J., Chai, C., Qian, Q., Li, C., Tang, J., Sun, L., et al. (2008). Mutations of genes in synthesis of the carotenoid precursors of ABA lead to pre-harvest sprouting and photo-oxidation in rice. *Plant J.* 54, 177–189. doi: 10.1111/j.1365-313X.2008.03411.x
- Finkelstein, R., Reeves, W., Ariizumi, T., and Steber, C. (2008). Molecular aspects of seed dormancy. *Annu. Rev. Plant Biol.* 59, 387–415. doi: 10.1146/annurev.arplant.59.032607.092740
- Gao, H., Zhang, C., He, H., Liu, T., Zhang, B., Lin, H., et al. (2020). Loci and alleles for submergence responses revealed by GWAS and transcriptional analysis in rice. *Mol. Breed.* 40, 75. doi: 10.1007/s11032-020-01160-6
- Gubler, F., Millar, A. A., and Jacobsen, J. V. (2005). Dormancy release, ABA and pre-harvest sprouting. *Curr. Opin. Plant Biol.* 8, 183–187. doi: 10.1016/j.pbi.2005.01.011
- He, P., Wang, C., Zhang, N., Liu, B., Yang, Y., Zhu, Y., et al. (2021). Multi-genotype varieties reduce rice diseases through enhanced genetic diversity and show stability and adaptability in the field. *Phytopathol. Res.* 3, 28. doi: 10.1186/s42483-021-00105-x
- Heldenbrand, J. R., Baheti, S., Bockol, M. A., Drucker, T. M., Hart, S. N., Hudson, M. E., et al. (2019). Recommendations for performance optimizations when using GATK3.8 and GATK4. *BMC Bioinf.* 20, 557. doi: 10.1186/s12859-019-3169-7
- Holdsworth, M. J., Bentsink, L., and Soppe, W. J. J. (2008). Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytol.* 179, 33–54. doi: 10.1111/j.1469-8137.2008.02437.x
- Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical J.* 50, 346–363. doi: 10.1002/bimj.200810425
- Huang, S., Hu, L., Zhang, S., Zhang, M., Jiang, W., Wu, T., et al. (2021). Rice osWRKY50 mediates ABA-dependent seed germination and seedling growth, and ABA-independent salt stress tolerance. *Int. J. Mol. Sci.* 22, 8625. doi: 10.3390/ijms22168625
- Ju, L., Liu, R., Cheng, X., Wang, Y., Lv, X., Chu, J., et al. (2025). Integrated transcriptome, GWAS, and metabolome revealed the mechanism of seed germination in sorghum. *Front. Plant Sci.* 16. doi: 10.3389/fpls.2025.1601899
- Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S. Y., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42, 348–354. doi: 10.1038/ng.548
- Kawahara, Y., De La Bastide, M., Hamilton, J. P., Kanamori, H., McCombie, W. R., Ouyang, S., et al. (2013). Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6, 4. doi: 10.1186/1939-8433-6-4
- Khan, N. (2025). Decoding phytohormone signaling in plant stress physiology: Insights, challenges, and future directions. *Environ. Exp. Bot.* 231, 106099. doi: 10.1016/j.envexpbot.2025.106099
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Klinsawang, S., Aesomnuk, W., Mangkalasane, P., Ruanjaichon, V., Siangliw, J. L., Pandey, B. K., et al. (2024). Genome-wide association study identifies key F-box genes linked to ethylene responsiveness and root growth in rice (*Oryza sativa* L.). *Front. Plant Sci.* 15. doi: 10.3389/fpls.2024.1501533
- Kobayashi, Y., Murata, M., Minami, H., Yamamoto, S., Kagaya, Y., Hobo, T., et al. (2005). Abscissic acid-activated SNRK2 protein kinases function in the gene-regulation pathway of ABA signal transduction by phosphorylating ABA response element-binding factors. *Plant J.* 44, 939–949. doi: 10.1111/j.1365-313X.2005.02583.x
- Kobayashi, Y., Yamamoto, S., Minami, H., Kagaya, Y., and Hattori, T. (2004). Differential activation of the rice sucrose nonfermenting1-related protein kinase2 family by hyperosmotic stress and abscisic acid. *Plant Cell* 16, 1163–1177. doi: 10.1105/tpc.019943
- Lee, H., Choi, M., Hwang, W., Jeong, J., Yang, S., and Lee, C. (2021). Occurrence of rice preharvest sprouting varies greatly depending on past weather conditions during grain filling. *Field Crops Res.* 264, 108087. doi: 10.1016/j.fcr.2021.108087
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656
- Liu, X., Li, Z., Hou, Y., Wang, Y., Wang, H., Tong, X., et al. (2019). Protein interactomic analysis of SAPKs and ABA-inducible bZIPs revealed key roles of SAPK10 in rice flowering. *Int. J. Mol. Sci.* 20, 1427. doi: 10.3390/ijms20061427
- Liu, D., Zeng, M., Wu, Y., Du, Y., Liu, J., Luo, S., et al. (2022). Comparative transcriptomic analysis provides insights into the molecular basis underlying pre-harvest sprouting in rice. *BMC Genomics* 23, 771. doi: 10.1186/s12864-022-08998-4
- Lv, Y., Ma, J., Wang, Y. Y., Wang, Q., Lu, X. L., Hu, H. T., et al. (2021). Loci and natural alleles for low-nitrogen-induced growth response revealed by the genome-wide association study analysis in rice (*Oryza sativa* L.). *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.770736
- Lv, Y., Ma, J., Wei, H., Xiao, F., Wang, Y., Jahan, N., et al. (2022). Combining GWAS, genome-wide domestication and a transcriptomic analysis reveals the loci and natural alleles of salt tolerance in rice (*Oryza sativa* L.). *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.912637
- Min, M. H., Khaing, A. A., Chu, S.-H., Nawade, B., and Park, Y. J. (2024). Exploring the genetic basis of pre-harvest sprouting in rice through a genome-wide association study-based haplotype analysis. *J. Integr. Agric.* 23, 2525–2540. doi: 10.1016/j.jia.2023.12.004
- Min, M. H., Maung, T. Z., Cao, Y., Phitaktansakul, R., Lee, G. S., Chu, S. H., et al. (2021). Haplotype analysis of BADH1 by next-generation sequencing reveals association with salt tolerance in rice during domestication. *Int. J. Mol. Sci.* 22, 7578. doi: 10.3390/ijms22147578
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015
- Peng, L., Sun, S., Yang, B., Zhao, J., Li, W., Huang, Z., et al. (2022). Genome-wide association study reveals that the cupin domain protein OsCDP3.10 regulates seed vigour in rice. *Plant Biotechnol. J.* 20, 485–498. doi: 10.1111/pbi.13731
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Sohn, S. I., Pandian, S., Kumar, T. S., Zoclanclounon, Y. A. B., Muthuramalingam, P., Shilpha, J., et al. (2021). Seed dormancy and pre-harvest sprouting in rice—an updated overview. *Int. J. Mol. Sci.* 22, 11804. doi: 10.3390/ijms222111804
- Song, S., Wang, G., Wu, H., Fan, X., Liang, L., Zhao, H., et al. (2020). OsMFT2 is involved in the regulation of ABA signaling-mediated seed germination through interacting with OsbZIP23/66/72 in rice. *Plant J.* 103, 532–546. doi: 10.1111/tj.14748
- Tai, L., Wang, H. J., Xu, X. J., Sun, W. H., Ju, L., Liu, W. T., et al. (2021). Pre-harvest sprouting in cereals: genetic and biochemical mechanisms. *J. Exp. Bot.* 72, 2857–2876. doi: 10.1093/jxb/erab024
- Tezuka, D., Matsuura, H., Saburi, W., Mori, H., and Imai, R. (2021). A ubiquitously expressed UDP-glucosyltransferase, UGT74J1, controls basal salicylic acid levels in rice. *Plants* 10, 1875. doi: 10.3390/plants10091875
- Thilakarathne, A. S., Liu, F., and Zou, Z. (2025). Plant signaling hormones and transcription factors: key regulators of plant responses to growth, development, and stress. *Plants (Basel)* 14, 1070. doi: 10.3390/plants14071070
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016
- Visscher, P. M., Brown, M. A., McCarthy, M. L., and Yang, J. (2012). Five years of GWAS discovery. *Am. J. Hum. Genet.* 90, 7–24. doi: 10.1016/j.ajhg.2011.11.029
- Wang, J., Deng, Q., Li, Y., Yu, Y., Liu, X., Han, Y., et al. (2020b). Transcription factors rc and osVP1 coordinately regulate preharvest sprouting tolerance in red pericarp rice. *J. Agric. Food Chem.* 68, 14748–14757. doi: 10.1021/acs.jafc.0c04748

- Wang, Y., Hou, Y., Qiu, J., Wang, H., Wang, S., Tang, L., et al. (2020a). Absciscic acid promotes jasmonic acid biosynthesis via a 'SAPK10-bZIP72-AOC' pathway to synergistically inhibit seed germination in rice (*Oryza sativa*). *New Phytol.* 228, 1336–1353. doi: 10.1111/nph.16774
- Wang, Y., Liao, Y., Quan, C., Li, Y., Yang, S., Ma, C., et al. (2022). C2H2-type zinc finger OsZFP15 accelerates seed germination and confers salinity and drought tolerance of rice seedling through ABA catabolism. *Environ. Exp. Bot.* 199, 104873. doi: 10.1016/j.envexpbot.2022.104873
- Wang, W., Xie, Z., Wu, Y., Sun, Y., Zhan, C., Jin, L., et al. (2024). The JA-OsJAZ6-DELLA module controls the tillering and drought stress response in rice. *Environ. Exp. Bot.* 222, 105776. doi: 10.1016/j.envexpbot.2024.105776
- Wei, H., Wang, X., Zhang, Z., Yang, L., Zhang, Q., Li, Y., et al. (2024). Uncovering key salt-tolerant regulators through a combined eQTL and GWAS analysis using the super pan-genome in rice. *Natl. Sci. Rev.* 11, nwae043. doi: 10.1093/nsr/nwae043
- Wei, Z., Yuan, Q., Lin, H., Li, X., Zhang, C., Gao, H., et al. (2021). Linkage analysis, GWAS, transcriptome analysis to identify candidate genes for rice seedlings in response to high temperature stress. *BMC Plant Biol.* 21, 85. doi: 10.1186/s12870-021-02857-2
- Xiong, M., Yu, J., Wang, J., Gao, Q., Huang, L., Chen, C., et al. (2022). Brassinosteroids regulate rice seed germination through the BZR1-RAmy3D transcriptional module. *Plant Physiol.* 189, 402–418. doi: 10.1093/plphys/kiac043
- Xu, F., Yoshida, H., Chu, C., Matsuoka, M., and Sun, J. (2025). Seed dormancy and germination in rice: Molecular regulatory mechanisms and breeding. *Mol. Plant* 18, 960–977. doi: 10.1016/j.molp.2025.05.010
- Yi, X., Hua, W., Zhang, Z., Liu, L., Liu, X., Liu, F., et al. (2025). Dissecting the genetic basis of preharvest sprouting in rice using a genome-wide association study. *J. Agric. Food Chem.* 73, 3257–3267. doi: 10.1021/acs.jafc.4c09230
- Zhang, C., Dong, S. S., Xu, J. Y., He, W. M., and Yang, T. L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinf. (Oxford England)* 35, 1786–1788. doi: 10.1093/bioinformatics/bty875
- Zhang, C., Zhou, L., Lu, Y., Yang, Y., Feng, L., Hao, W., et al. (2020). Changes in the physicochemical properties and starch structures of rice grains upon pre-harvest sprouting. *Carbohydr Polym* 234, 115893. doi: 10.1016/j.carbpol.2020.115893
- Zhao, B., Zhang, H., Chen, T., Ding, L., Zhang, L., Ding, X., et al. (2022). Sdr4 dominates pre-harvest sprouting and facilitates adaptation to local climatic condition in Asian cultivated rice. *J. Integr. Plant Biol.* 64, 1246–1263. doi: 10.1111/jipb.13266
- Zhu, D., Qian, Z., Wei, H., Guo, B., Xu, K., Dai, Q., et al. (2019). The effects of field pre-harvest sprouting on the morphological structure and physicochemical properties of rice (*Oryza sativa* L.) starch. *Food Chem.* 278, 10–16. doi: 10.1016/j.foodchem.2018.11.017