frontiers | Frontiers in Plant Science

Check for updates

Mining SNP loci and candidate genes for sheath blight resistance in *Indica* rice using genome-wide association studies

Xin Gu[1†], Junjie Ding[2†], Hui Xu[1], Wei Liu[1], Hui Wu[1], Qiang Tang[1], Shuai Yang[3], Ling Wang[3], Suhua Zhang[1], Long Chen[1], Xue Meng[1], Ye Tao[1], Jing Hao[1], Zhao Chen[1], Shuhua Jiang[1], Ruoting Yu[1], Liying Xu[1*] and Muhammad Ahmad Hassan[4*]

[1]College of Landscape Architecture and Horticulture, Wuhu Vocational University, Wuhu, China, [2]Jiamusi Branch of Heilongjiang Academy of Agricultural Sciences/ Observation and Experiment Station of Crop Pests of Jiamusi, Ministry of Agriculture and Rural Affairs, Jiamusi, Heilongjiang, China, [3]Shenyang Polytechnic College, Shenyang, China, [4]College of Resource and Environment, Anhui Agricultural University, Hefei, China

Rice is a vital component of the global food supply chain. Sheath blight (ShB) poses a severe threat to *indica* rice production, resulting in substantial yield and quality deteriorations. Although chemical control strategies are widely employed, their prolonged use raises concerns about the development of pathogenic resistance. Genome-wide association studies (GWAS) serve as a critical tool for identifying genetic loci associated with ShB resistance. However, current research efforts are constrained by limitations in sample size, environmental influences, and analytical methodologies. Notably, studies on ShB resistance have focused mainly on *japonica* rice, with *indica* rice receiving limited attention in GWAS-based investigations. Prioritizing GWAS research targeting ShB resistance in *indica* rice is essential for enriching disease resistance gene resources and facilitating molecular breeding programs. Eighty-four *indica* accessions were subjected to *in vitro* phenotyping for ShB resistance, followed by whole-genome resequencing using a magnetic bead-based DNA extraction protocol. Population structure analysis revealed high genetic diversity (K = 5) with normally distributed resistance phenotypes. From 2.93 million raw SNPs, 904,708 high-quality variants were retained for multi-model GWAS (GLM/MLM/EMMAX/GEMMA), identifying 22 consensus loci. Haplotype blocks delineated three critical regions harboring 15 candidate genes, including ABC transporters and calmodulin-like proteins with putative roles in pathogen defense, among others. Functional annotation revealed that ABCG transporters, calmodulin-like proteins (CMLs), non-specific lipid transfer proteins (nsLTPs), and receptor-like serine/ threonine kinases orchestrate defense responses through calcium-mediated signaling, antimicrobial metabolite translocation, and maintenance of membrane integrity. qPCR analysis revealed that the three GWAS-associated candidate genes were significantly upregulated after *R. solani* inoculation in the resistant line X-11, suggesting that variations in their non-coding regions

contribute to disease resistance by enhancing both basal and pathogen-induced expression. These loci constitute a genetic toolkit for deciphering *Rhizoctonia solani* Kühn (*R. solani*) resistance mechanisms and facilitate precision breeding in *indica* rice.

1 Introduction

Rice (*Oryza sativa* L.) is a crucial component in ensuring global food security, as it fulfills the dietary needs of half of the world's population (Hassan et al., 2023; Zhu et al., 2024). Rice productivity is endangered by several biotic (pathogenic diseases) and abiotic factors (temperature extremes, heavy metals, salinity, drought, etc.) (Turaidar et al., 2018; Khoshkdaman et al., 2021). Among biotic factors, diseases (i.e., blast, blight, etc.) are significant constraints for yield and quality reduction in rice (Turaidar et al., 2018). Rice sheath blight (ShB) is a prevalent and frequently occurring disease in rice, causing substantial yield reductions (20~50%), particularly in *indica* rice (*Oryza sativa* ssp. *indica*) (Richa et al., 2016; Srivastava et al., 2024). The *Rhizoctonia solani* Kühn (*R. solani*), a soil-borne fungal pathogen, is the causative agent of rice ShB (Ganesh and Roka, 2024). It was first identified in Japan and subsequently detected in other rice-producing regions, including China, Bangladesh, Vietnam, Thailand, Pakistan, and others (Srivastava et al., 2024). Prolonged high humidity (>80%) and temperatures of 28~35°C are favorable conditions that accelerate epidemic development of rice ShB (Uppala and Zhou, 2018). Long, uneven, greyish green spots with brown margins on leaves and developing panicles are the symptoms of rice ShB, which is also known as snakeskin disease due to its symptoms resemblance with snakeskin (Molla et al., 2020; Hunjan et al., 2022) Fungicide application remains a primary control strategy, its overuse has triggered the widespread evolution of resistant *R. solani* strains, as it remains dormant under the soil for up to 2 years and become active on favorable conditions, such conditions jeopardizing the sustainable rice cultivation systems (Hunjan et al., 2022; Ganesh and Roka, 2024; Zhai et al., 2025).

Genome-wide association studies (GWAS), empowered by high-density genotyping and next-generation sequencing, have revolutionized the dissection of polygenic trait architectures in plant systems (Chen et al., 2014; Wang et al., 2020). GWAS has proven instrumental in decoding the genetic architecture of ShB resistance, with recent advances enabling systematic mining of resistance-associated alleles (Zhang et al., 2019). The widespread infestation of ShB in *indica* rice-growing regions becomes more severe due to humid climatic conditions that are conducive to *R. solani* proliferation (Yu et al., 2025). This epidemiological pressure has driven rigorous efforts to identify ShB-resistant genetic loci and utilize them for developing elite cultivars, a strategic priority in contemporary disease-resistance breeding plans (Hunjan et al., 2022). Advances in high-throughput sequencing and population genetics algorithms have substantially enhanced the power of GWAS in decoding the genetic complexity of rice ShB resistance (Lu et al., 2025). Wang et al (Wang et al., 2021). have conducted an investigative study and performed the whole-genome resequencing of 259 rice accessions, integrating SNP-GWAS, haplotype-based GWAS (Hap-GWAS), and weighted gene co-expression network analysis (WGCNA) to identify 653 core candidate genes associated with ShB resistance. Multi-environment trials conducted across diverse locations and seasons strengthened the ecological validity; however, the inherent environmental heterogeneity of field conditions may introduce confounding factors, potentially obscuring the causal relationships between genotype and phenotype. Building on this methodological consistency, subsequent research analyzed 563 rice accessions from 47 countries, representing *indica*, *japonica*, and *aus-type* subpopulations (Zhang et al., 2019). Utilizing ~3 million SNPs, they employed both single-locus GWAS (EMMAX) and multi-locus approaches (FarmCPU) to dissect ShB resistance traits, i.e., culm length (CL), lesion height (LH), and relative lesion height (RLH). Cross-validation between methods enhanced the reliability of identified resistance-associated loci (Chen et al., 2019). Complementing these multi-model approaches, a subsequent investigation conducted genotyping of 228 *indica* accessions using a high-density rice SNP array (700K markers), identifying multiple significant loci associated with sheath blight resistance on chromosomes 1, 4, and 11 through GWAS (Oreiro et al., 2020). The study further integrated transcriptomic profiling with GWAS signals to elucidate resistance mechanisms across both positional genomics and gene regulatory dimensions. This dual-system validation strengthens mechanistic plausibility beyond statistical associations. Echoing this multi-system strategy, Li et al (Li et al., 2022). implemented a multi-environment trial framework that combined field evaluations in Arkansas (USA) and Nanning (China) with greenhouse micro-chamber assays to dissect ShB resistance alongside agronomic traits, including plant height, heading date, tiller number, and panicle count. GWAS identified 22 ShB-associated SNPs across chromosomes (excluding Chr10 and Chr12), with 13 intragenic SNPs encoding proteins such as

protease inhibitors, ATPases, and cyclins, several of which exhibit pathogen-responsive functions and have been validated in prior studies. Notably, seven resistance loci co-localized with previously reported ShB resistance genes or quantitative trait loci (QTLs), reinforcing their biological relevance. However, environmental heterogeneity (divergent climates and soils) and distinct pathogen isolates across varying test locations complicate the precise resolution of genotype-by-environment (G × E) interactions. Building on this genetic dissection, Chen et al (Chen et al., 2019). utilized the Rice Diversity Panel 1 (RDP1, comprising 299 accessions and 44K SNPs) to elucidate ShB resistance through controlled greenhouse inoculation and standardized phenotyping. GWAS revealed 11 significant loci, with two robust QTLs (i.e., *qSB-3* and novel *qSB-6*) on chromosomes 3 and 6, respectively. Haplotype analysis demonstrated additive effects: accessions carrying dual superior haplotypes (AGC) exhibited a 1.21-unit reduction in lesion score. Particularly, the aromatic (ARO) and aus subgroups displayed higher resistance than the tropical *japonica* (TRJ) and *indica* (IND) subgroups. However, the limited sample size of the ARO subgroup (n = 11) may have constrained statistical validation.

Whereas GWAS generally offer stronger statistical authentication in large populations, identifying disease-resistant loci with limited samples is feasible via optimized experimental designs and methodological improvements (Wang et al., 2019; Uffelmann et al., 2021). A genomic analysis conducted GWAS on 150 accessions from the Ting core collection genotyped with 5.1 million SNPs (Fu et al., 2022). While phenotyping across two seasons (2016-2017) aimed to validate the robustness of the results, the limited cohort size constrained statistical power and generalizability. Notably, no SNPs showed replicable associations across years, underscoring the environment-dependent nature of resistance mechanisms. Similarly, another research investigation employed a multi-locus GWAS framework (FarmCPU/BLINK) in 96 sweet potato genotypes to identify Fusarium root rot resistance loci, implementing stringent false discovery rate control (Kim et al., 2023). Cross-model validation, combined with Bonferroni-adjusted thresholds, reduced spurious associations compared to single-model approaches, demonstrating methodological triangulation as an effective strategy to enhance GWAS reliability in small cohorts (n < 100). This combinatorial strategy has parallel applications in research, where GWAS on 144 maize accessions (using the Illumina MaizeSNP50 array) identified 18 head smut resistance genes (Wang et al., 2012). To mitigate cohort size constraints, the study implemented threefold optimization: 1) strategic selection of genetically diverse germplasm; 2) rigorous quality control pipelines; 3) multi-model validation using general and mixed linear models (GLM/MLM). Such methodological layering effectively compensates for population-scale limitations while maintaining robustness in the discovery process.

Furthermore, all existing studies have demonstrated that cohort size, experimental conditions, and analytical methodologies significantly influence GWAS outcomes. While large populations enhance statistical authenticity, practical constraints—such as the minimal availability of germplasm with specific genetic backgrounds —often restrict sample sizes. Small-cohort GWAS can achieve biological relevance through methodological safeguards: multi-model association analyses, stringent significance thresholds, population structure correction, and kinship matrix integration. Notably, current ShB resistance research disproportionately focuses on *japonica* rice, leaving *indica* systems underexplored. Building upon these insights, our study utilizes a localized *indica* core collection to dissect the architecture of ShB resistance. Employing Tassel-based GWAS with minor allele frequency (MAF) filtered SNPs (>0.05), we integrate GLM/MLM to identify consensus loci. Subsequent linkage disequilibrium (LD) block delineation and candidate gene mining within critical regions refined the mechanistic hypotheses.

2 Materials and methods

2.1 Plant material and phenotypic evaluation

The plant material of 84 *indica* rice accessions was used in this research experiment. Detailed information about the breeding background and geographical distribution of these 84 accessions has been presented in Supplementary Table S1. Rice ShB resistance phenotyping of 84 accessions was performed using an optimized *in vitro* inoculation assay adapted from Li et al (Li et al., 2022). Additionally, 20 accessions were randomly selected, and the results were validated through field-based toothpick insertion inoculation. Locally isolated *R. solani* strains were cultivated on potato dextrose agar (PDA) medium, with 5-mm mycelial plugs serving as inoculum. Healthy leaves from four-leaf-stage seedlings were surface-sterilized and excised into 6 cm segments. Four segments per plate were arranged on triple-layer sterile filter paper in 9 cm Petri dishes, centrally inoculated with mycelial plugs, and maintained at 25°C under a 12 h photoperiod. To sustain humidity, a pH-adjusted 40 mg/L 6-BA solution was applied. Disease severity was recorded at 24 h post-inoculation using a 6-tier scale, as follows:

    i. 0 (Immune, I): No lesions
   ii. 1 (Resistant, R): <1/8 leaf area affected
  iii. 2 (Moderately resistant, MR): 1/8–1/4
   iv. 3 (Moderately susceptible, MS): 1/4–1/2
    v. 4 (Susceptible, S): 1/2–3/4
   vi. 5 (Highly susceptible, HS): >3/4

Three biological replicates were conducted for each accession.

2.2 DNA extraction and resequencing

Genomic DNA was isolated from 84 accessions using a magnetic bead-based extraction method. DNA concentration and purity were verified using Nano-Drop 2000 (Thermo Fisher Scientific). Sequencing libraries were prepared through fragmentation, size selection, and adapter ligation and then sequenced on an Illumina NovaSeq 6000. Raw reads were filtered

(Phred score < 20) and aligned to the reference genome (MSU v7.0) using Bowtie2 v2.3.5.1 (Langmead and Salzberg, 2012), with subsequent sorting performed via SAMtools v1.3.1 (Danecek et al., 2021). Redundant reads were filtered using Mark-Duplicates from Picard (Belay et al., 2024), followed by variant calling via GATK4 Haplotype-Caller (McKenna et al., 2010). Joint genotyping generated population-level genomic variant call formats (gVCFs), filtered through stringent quality thresholds (QUAL ≥30, QD ≥2.0, MQ ≥40, FS ≤60.0, MQRankSum ≥-12.5, ReadPosRankSum ≥ 8.0). Final SNP sets underwent MAF (≥0.05) and missing rate (≤20%) filtering using VCF-tools v0.1.6 (Danecek et al., 2011), with genetic diversity assessed via PLINK toolkit (Purcell et al., 2007).

2.3 Population analysis and structure analysis

Following variant filtering, population genetic analyses were conducted using TASSEL5.0 (Bradbury et al., 2007). Principal component analysis (PCA) and kinship estimation were employed to identify genetic relationships among the 84 *indica* accessions. A neighbor-joining (NJ) phylogenetic tree was constructed and visualized in MEGA11 (Tamura et al., 2021). Population structure analysis using ADMIXTURE (Alexander et al., 2009) identified K = 5 as the optimal clustering parameter, as determined by minimizing the cross-validation error, with corresponding ancestry proportions graphically represented.

2.4 GWAS analysis and linkage disequilibrium analysis

To enhance statistical robustness in GWAS with limited sample size, we implemented a multi-model framework integrating TASSEL's GLM, MLM, EMMAX (Kang et al., 2010), and GEMMA's (Zhou and Stephens, 2012) Bayesian sparse linear mixed models. The LD decay analysis via PopLD decay (Zhang et al., 2019) informed recombination boundaries. Significant phenotype-associated SNPs were cross-validated across models, with LD blocks delineated through LD-Block-Show visualization. Candidate genes within these blocks were prioritized through functional annotation against established databases of plant-pathogen interactions. The functions of the candidate genes were annotated using the NCBI (https://www.ncbi.nlm.nih.gov/), UniProt (https://www.uniprot.org/), and Plant-Pathogen Interactions (http://www.phi-base.org/) databases.

2.5 Quantitative real-time PCR analysis

Three ShB resistance candidate genes (*BGIOSGA000756*, *BGIOSGA036403*, and *BGIOSGA037213*) previously identified through GWAS were selected for expression analysis. The resistant line X-11 and susceptible line X-71 were inoculated with *R. solani*, and leaf samples were collected at 0 and 24 hours post-inoculation (hpi). Total RNA was extracted from leaves using the MiniBEST Plant RNA Extraction Kit (TaKaRa). RNA quality and quantity were assessed, and qualified RNA was reverse-transcribed into cDNA using the PrimeScript™ 1st Strand cDNA Synthesis Kit (TaKaRa). Gene-specific primers for each candidate gene are listed in Table 1.

Quantitative real-time PCR (qPCR) was performed using the SupRealQ Ultra Hunter SYBR qPCR Master Mix (U+) (Vazyme Biotech) on an MA-6000 Real-Time PCR System (Suzhou Yarui Biotechnology). Each 20 μL reaction contained 10 μL of 2× Master Mix, 1.0 μL of cDNA template, and 0.4 μL each of forward and reverse primers (10 μM). The thermal cycling conditions were as follows: initial denaturation at 95°C for 5 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 30 seconds. The rice Actin gene (*BGIOSGA033259*) was used as an internal reference for normalizing relative gene expression, with the following primer sequences: forward, 5'-GCCATTCTCCGTCTTGATCTTGC-3'; reverse, 5'-AGCGACAACCTTGATCTTCATGCT-3' (Baiya et al., 2014). Relative gene expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method. The statistical significance of expression differences was determined using an independent-samples t-test, with *p* < 0.05 considered significant and *p* < 0.01 considered highly significant.

3 Results

3.1 Phenotyping

Dual *in vitro* ShB resistance assays yielded statistically concordant phenotypic measurements, as determined by density plot analysis (Supplementary Figure 1). Consequently, the lesion length ratio was selected for genome-wide association studies. The *indica* panel (n = 84) exhibited a resistance spectrum stratification, with 8 resistant (R), 17 moderately resistant (MR), 51 moderately susceptible (MS), and 8 susceptible (S) accessions. The normal distribution of resistance phenotypes (Supplementary Table S2), characterized by the predominance of moderate resistance and lack of extreme phenotypes, reflects the natural population architecture and validates its suitability for GWAS. Importantly, all genotypes

TABLE 1 Primer sequences used for qPCR.

	Gene	Forward primer	Reverse primer
1	*BGIOSGA000756*	AGATCTTCCGCCACTTCGAC	TCGTAGGCGCTTATCATGG
2	*BGIOSGA036403*	GGTCAATGGAATCCCTGAG	ACAATGCTCCTGTGTGGAC
3	*BGIOSGA037213*	ATTAGTGTGAGTGCTGTGGTG	TGGCTGCTTGTGGATTTAG

displayed measurable lesions, including nominally resistant lines, confirming successful pathogen invasion. The results showed that the coincidence rate between toothpick insertion inoculation and *in vitro* leaf inoculation at the tillering stage was high, reaching 85%, indicating that seedling-stage *in vitro* leaf inoculation can accurately identify phenotypes (Supplementary Table S3, Supplementary Figure S1, Supplementary Figure S2).

3.2 Genotyping

Variant calling and stringent filtering through the Genome Analysis Tool Kit (GATK) yielded 2,933,319 genome-wide SNPs (Table 1), distributed across all 12 chromosomes and scaffolds (Figure 1). Chromosomal SNP densities ranged from 6.43/kb (Chr8) to 8.57/kb (Chr2), averaging 7.48 SNPs/kb (Supplementary Table S4). Post quality control (QC) metrics revealed minimal missing data (<15% per sample) and pronounced allelic asymmetry: 65.28% of SNPs exhibited MAF <0.05, with heterozygosity levels below 0.15 in 98.82% loci; patterns are consistent with rice's autogamous reproduction. Final GWAS-ready SNPs (904,708) were obtained through MAF (≥0.05) and missing rate (≤20%) filtration. As shown in the supplementary results, the average sequencing depth of the 85 indica rice accessions ranges from 4.53X to 13.54X (mean = 9.02X); the 1X sequencing depth coverage rate varies between 0.68 and 0.87 (mean = 0.80); and the 5X sequencing depth coverage rate is from 0.44 to 0.77 (mean = 0.67). Although minor variations exist in the sequencing coverage metrics among different accessions, the resequencing data overall exhi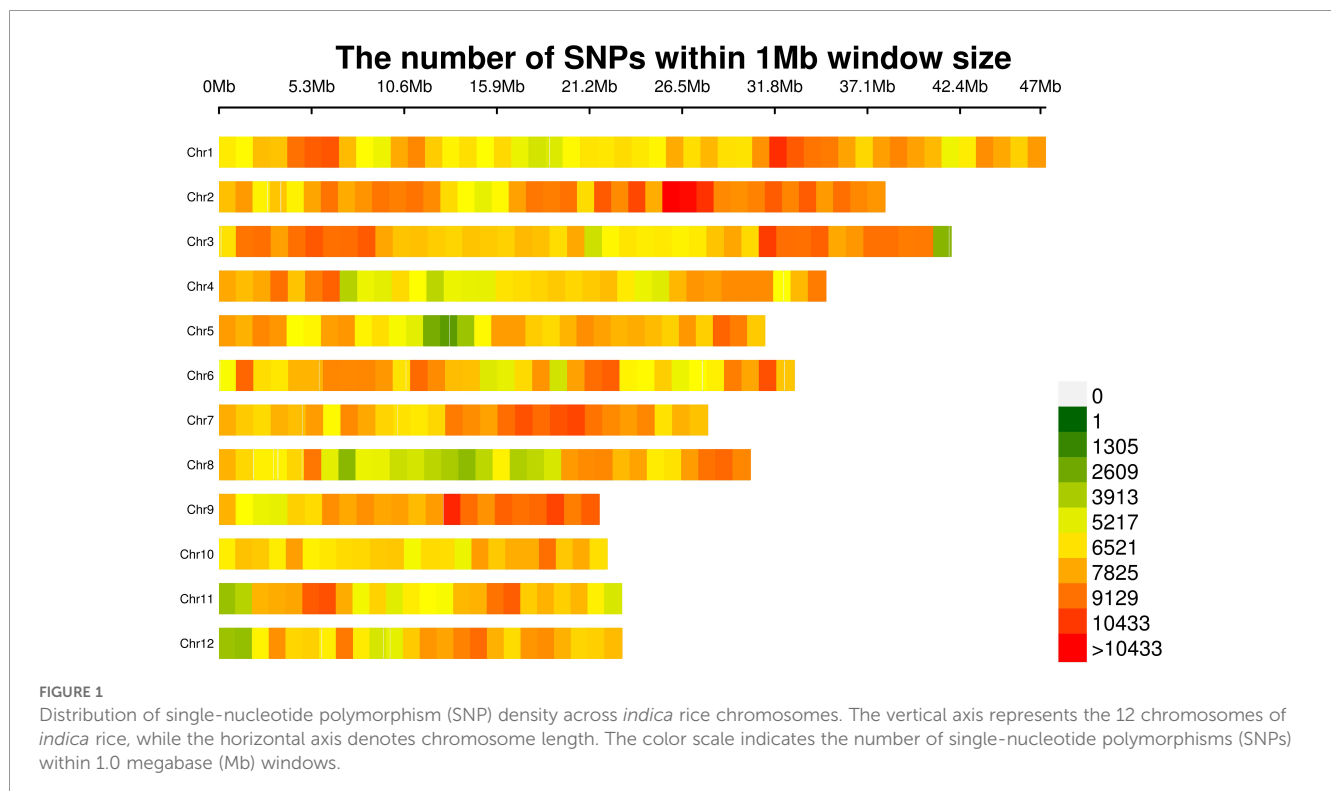bit good genome coverage completeness. This reliable dataset can provide solid support for subsequent studies, such as genetic diversity analysis and functional gene mining (Supplementary Table S5).
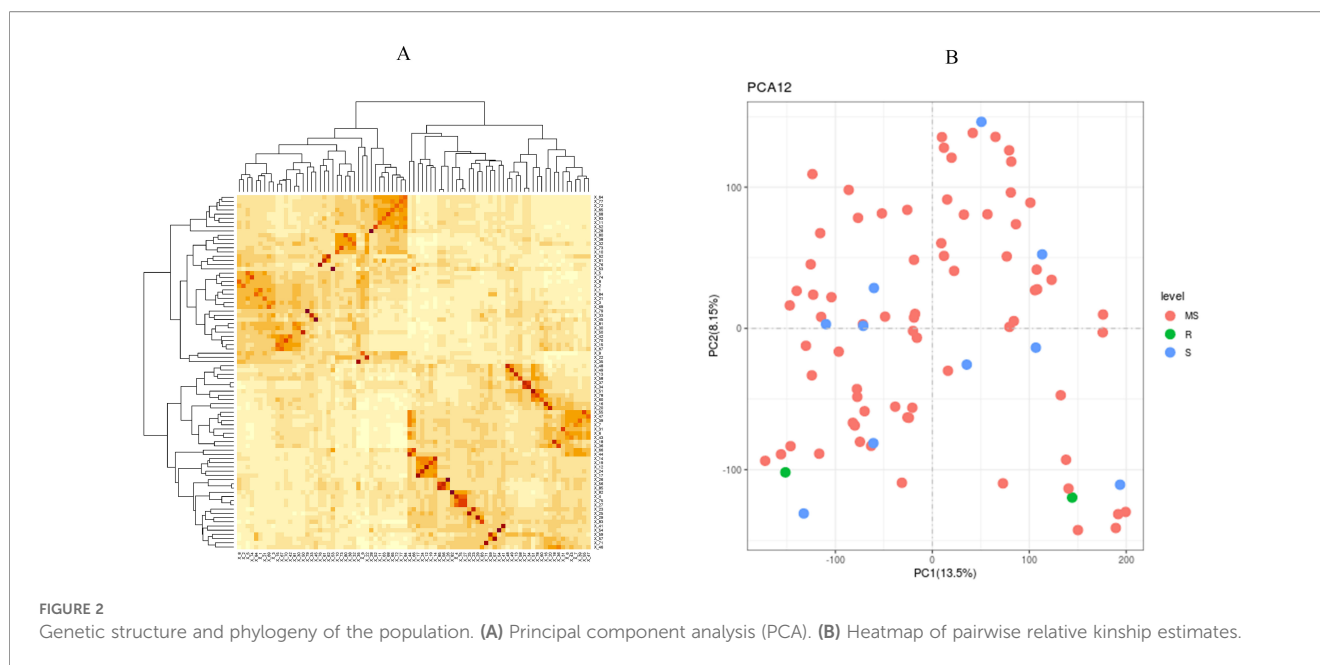
3.3 Population structure analysis

Among 84 *indica* accessions, there was a little substructure found in the population genetic analyses of 904,708 SNPs Principal component analysis (PCA) demonstrated no disease resistance-based clustering (Figure 2A). In contrast, kinship heatmap analysis identified localized regions of moderate relatedness without distinct population stratification (Figure 2B). ADMIXTURE analysis determined K = 5 as the optimal cluster number through cross-validation error minimization (Figure 3A). However, ancestral components showed extensive admixture across accessions (Figure 3B). Phylogenetic re-construction via neighbor-joining method yielded short branches with overlapping clusters across K = 5 subgroups (Figure 3C), validating PCA/kinship findings of high genetic diversity and low differentiation. The LD decay analysis revealed moderate recombination rates, with the LD decay distance at ~70 kb (Figure 3D).

3.4 GWAS of genes resistant to ShB

Using the GLM of TASSEL software, 27 SNP markers were detected genome-wide that were significantly associated with the target traits (threshold 1×10^{-4}). The maximum number (6) of these significant loci mapped to chromosome 12, followed by 4 loci on

FIGURE 1
Distribution of single-nucleotide polymorphism (SNP) density across *indica* rice chromosomes. The vertical axis represents the 12 chromosomes of *indica* rice, while the horizontal axis denotes chromosome length. The color scale indicates the number of single-nucleotide polymorphisms (SNPs) within 1.0 megabase (Mb) windows.

FIGURE 2
Genetic structure and phylogeny of the population. **(A)** Principal component analysis (PCA). **(B)** Heatmap of pairwise relative kinship estimates.

chromosome 6 (Figure 4A). By contrast, using the MLM of TASSEL software, 89 significantly associated SNP markers (threshold: 1×10^{-3}) were identified, with 63 loci predominantly located on chromosome 12 and 8 loci on chromosome 6 (Figure 4B). The EMMAX software's MLM yielded 79 significant SNP markers (threshold: 1×10^{-4}), of which 63 loci were on chromosome 12 and 14 loci on chromosome 3 (Figure 4C). The GEMMA software's MLM identified 78 significant SNP markers (threshold: 1×10^{-4}), with 67 loci primarily on chromosome 12 (Figure 4D). The detection results of all software showed significant enrichment on chromosome 12, suggesting that this chromosomal region may harbor major QTLs regulating the target trait.

Subsequently, a Venn analysis was performed on the significant SNP markers resulting from the four models. As shown in Figure 5, three SNP markers were commonly identified by all four methods, located on chromosomes 3 (position: 38632300), 3 (position: 15089660), and 4 (position: 31894573), respectively. Twenty-two SNP markers were identified by three or more methods (Figure 5). For these 22 SNP loci, the genotypes of each sample and corresponding phenotypic information were extracted, and boxplots of genotype-phenotype relationships for each SNP locus were constructed. Taking three key SNP loci as examples, Figure 6A shows the distribution of lesion length ratio to leaf length for the S1_35575127 locus (Chr1: 35575127) on chromosome 1, with two genotypes (AA and AC). For the AA genotype, the median lesion length ratio was approximately 41.78%, indicating that most samples under this genotype exhibited lesion ratios concentrated around this level. The data distribution was relatively dispersed with a large box height, reflecting high variability in lesion ratios among AA genotype samples. The AC genotype had a median lesion length ratio of approximately 31.24%, significantly lower than that of the AA genotype, suggesting that *indica* rice carrying the AC genotype generally had lower lesion ratios and relatively stronger resistance to the ShB. The annotations above the two boxes in the figure indicate

extremely significant differences in lesion length ratios between the two genotypes (P = 0.0002, P < 0.01); this demonstrates that the genotype of the SNP locus S1_35575127 is closely associated with ShB resistance in *indica* rice, with the AC genotype being favorable for enhancing resistance. The EF-hand domain-containing protein (EFh protein) regulated by this locus plays a critical role in crop disease resistance, primarily through the regulation of calcium signaling pathways (Supplementary Table S6). Figure 6B illustrates the distribution of lesion length ratio to leaf length for three genotypes (AA, CA, CC) at the S12_7238044 locus on chromosome 12. The AA genotype is represented by a single horizontal line, indicating a small sample size, with a lesion length ratio of approximately 50.15%. The CA genotype exhibited a median lesion length ratio of ~45.77%, with relatively concentrated data distribution and a small box height, suggesting low variability in lesion ratios among samples. The CC genotype had a median lesion length ratio of ~39.08%, which was lower than both the CA and AA genotypes, indicating that *indica* rice with the CC genotype generally exhibited lower lesion ratios and stronger ShB resistance. Although its data distribution was more dispersed (larger box height, higher variability), the annotations above the CA and CC genotype boxes (P = 0.0002, P < 0.01) confirmed extremely significant differences in lesion ratios between these genotypes. These results indicated that different genotypes at the S12_7238044 SNP locus are closely associated with ShB resistance in *indica* rice, with the CC genotype potentially being favorable for enhancing resistance. This locus is located in the regulatory region of the ATP-binding cassette (ABC) transporter G family member 43-like gene, which regulates ABC transporter families and participates in transmembrane transport of plant defense-related metabolites (Supplementary Table S6).

Figure 6C presents the distribution of lesion length ratios for two genotypes (CC, CT) at the S12_7393096 locus. The CC genotype had a median lesion length ratio of ~39.20%, with
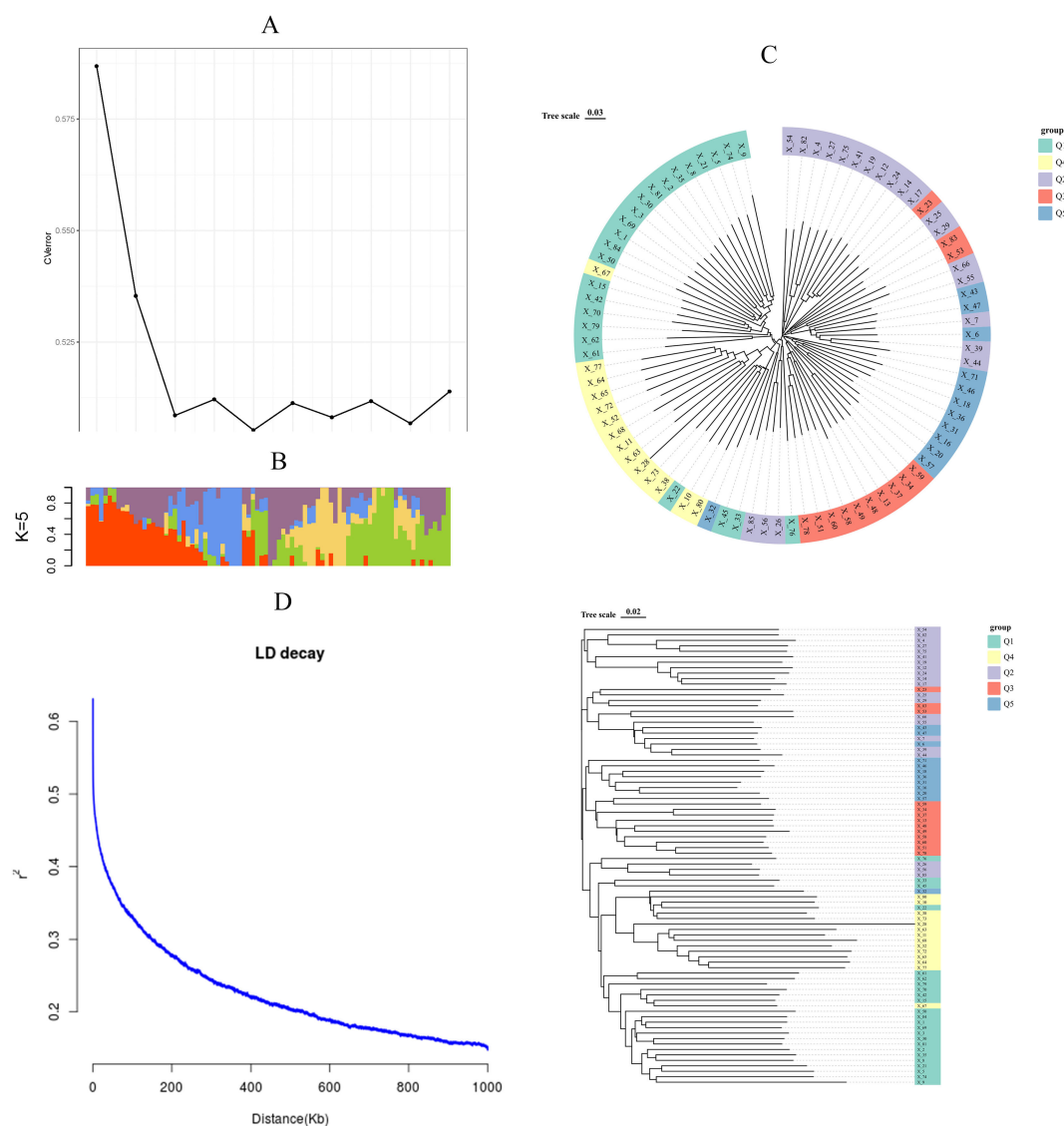
FIGURE 3
Population Genetic Architecture and LD Decay Analysis. **(A)** Cross-validation error plot for determining optimal ancestral clusters (K = 5).
(B) ADMIXTURE bar plot showing genetic structure across 84 accessions (K = 5), with colors representing distinct ancestral components.
(C) Neighbor-joining phylogenetic tree color-coded to ADMIXTURE clusters, demonstrating weak population substructure. **(D)** genome-wide linkage
disequilibrium (LD) decay profile showed that the LD of the population decayed to half its initial value at an approximate distance of 70 kb, with a
corresponding r^2 of approximately 0.35.

relatively dispersed data distribution and a large box height, reflecting high variability in lesion ratios among samples. The CT genotype exhibited a median lesion length ratio of ~46.23%, higher than that of the CC genotype, indicating that *indica* rice with the CT genotype had overall higher lesion ratios and weaker resistance. The annotations above the two boxes confirmed extremely significant differences in lesion ratios between CC and CT genotypes (P = 0.00016, P < 0.01). These findings demonstrated that the genotype at the S12_7393096 SNP locus is tightly linked to ShB resistance in *indica* rice, with the CC genotype potentially contributing to enhanced resistance. This locus is involved in regulating the expression of bi-functional inhibitor/plant lipid transfer protein/ seed storage helical domain-containing protein (BiP/LTP/SSH), which contains lipid-binding motifs and antimicrobial peptide

domains that may confer disease resistance by maintaining membrane integrity and direct action against pathogens.

3.5 Haplotype block analysis of consensus SNP-associated loci for ShB resistance

Haplotype block analysis was performed by extending 70-kb regions flanking the 22 consensus SNPs. Firstly, we extracted the 70 kb intervals upstream and downstream of the 22 consensus SNPs. For regions containing multiple consensus SNPs, these intervals were merged. Using the LD-Block-Show software and based on VCF files, we generated linkage disequilibrium (LD) heatmaps, with the haplotype ranges simultaneously displayed within the heatmaps.
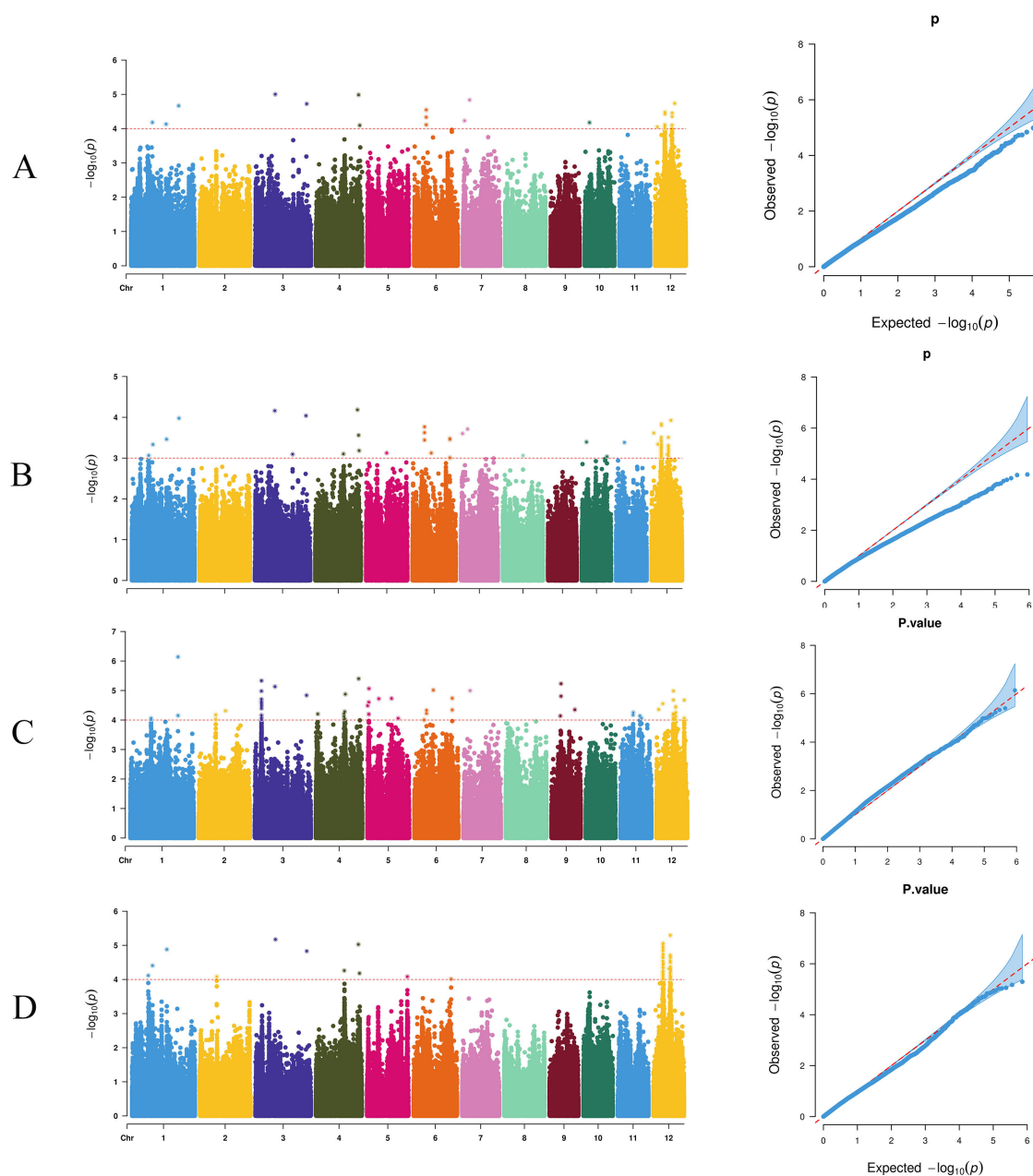
FIGURE 4
Manhattan plots and Q-Q plots for genome-wide association studies (GWAS) using four different methods. **(A)** TASSEL-GLM, **(B)** TASSEL-MLM, **(C)** EMMAX, and **(D)** GEMMA.

We further filtered the regions in LD with the consensus SNP markers (i.e., the intervals containing consensus SNPs), which were defined as the three final critical blocks. LD-Block-Show generated heatmaps revealed 12 candidate intervals (Figure 7), including three on Chr 12, one on Chr 7, and two each on Chr 1, Chr 3, Chr 4, and Chr 6. Subsequent LD network analysis prioritized three critical blocks, which contained consensus SNPs (Supplementary Table S7): block4 (Chr 4:32,681,394–32,769,963), block7 (Chr 7:4,997,038– 5,097,060), and block12 (Chr 12:1,685,784–1,685,811) (Figure 8). These regions exhibited strong LD persistence across multiple associated loci, suggesting potential regulatory hubs for ShB

resistance. The critical LD block 4 (Chr 4:32,681,394–32,769,963) spans 88.6 kb, containing 49 SNPs. Functional annotation identified three candidate genes, including a leucine-rich repeat receptor-like kinase (LRR-RLK, At1g56130 ortholog) carrying pathogen recognition domains. This kinase, containing extracellular LRR motifs, is postulated to mediate pathogen-associated molecular pattern (PAMP) perception and subsequent immune signaling activation during *R. solani* infection, consistent with its role in basal resistance pathways.

The critical LD block 7 (Chr 7:4,997,038–5,097,060), spanning 100 kb, harbors 401 SNPs and 11 candidate genes, including nine with

FIGURE 5
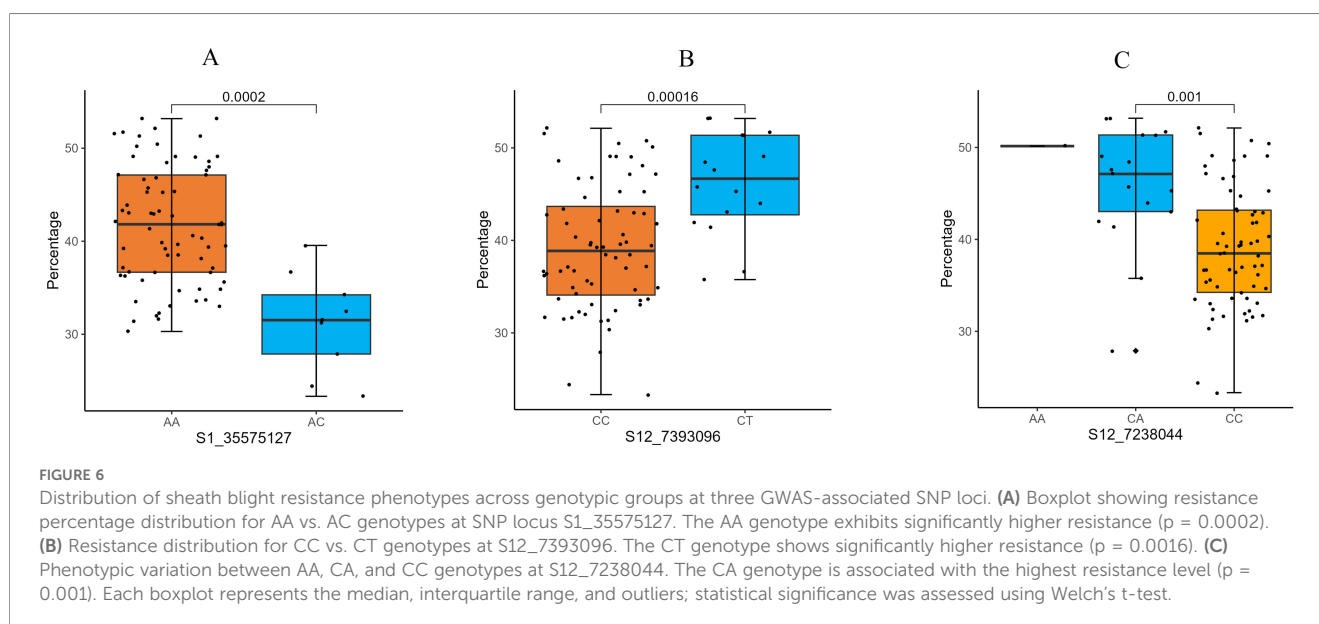Venn diagram for screening significantly associated SNP markers.
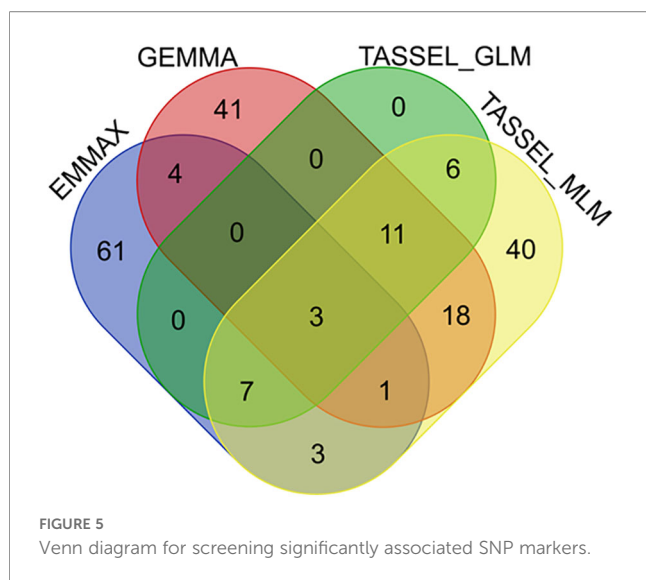
missense variants. Functional annotation highlights four mechanistically distinct candidates: an AAA+ ATPase domain protein implicated in cell wall remodeling through post-translational modification of acetylated proteins, potentially enhancing resistance against Gram-positive pathogens via lysozyme-mediated lysis (Jones et al., 2020); a glucose/sorbosone dehydrogenase hypothesized to modulate plant-pathogen interactions by altering cellular redox homeostasis, possibly through sorbosone-to-sorbosic acid conversion coupled with ROS signaling (Yang et al., 2021; Wang et al., 2023); a DDT domain-containing protein postulated to regulate immune signaling cascades via phosphorylation-dependent protein interactions (Wang et al., 2024); and a bHLH transcription factor demonstrating promoter-binding activity to defense-related genes, as evidenced by transcriptional activation assays (Supplementary Table S8) (Liao et al., 2025). These candidates collectively suggested that there are multi-layered defense mechanisms within this locus.

The critical LD block 12 (Chr 12:1,685,784–1,685,811) spans 28 bp and contains three SNPs, encompassing a single gene annotated as an uncharacterized protein. Despite limited functional evidence, this conserved genomic region is likely to harbor critical genetic determinants of ShB resistance, necessitating prioritized investigation through targeted mutagenesis and transcriptome profiling to elucidate its mechanistic role in disease defense pathways.

3.6 Expression patterns of candidate genes

BGIOSGA000756 encodes an EF-hand domain protein, a key component of the calcium signaling pathway, whose expression level directly influences the plant's early immune response to pathogens. qPCR results (Figure 9A) showed that at 0 hours post-inoculation (hpi), the relative expression level of *BGIOSGA000756* in the resistant line X-11 was 1.00 ± 0.02, higher than that in the susceptible line X-71 (0.74 ± 0.13), suggesting that regulatory sequence variations in this gene may confer a higher basal expression level in the resistant line even in the absence of pathogen stress. At 24 hpi, the expression level in X-11 was significantly upregulated to 2.19 ± 0.52, representing a more than two-fold increase compared to 0 hpi ($p < 0.05$), and was significantly higher than the level in X-71 at the same time point (0.94 ± 0.04, $p < 0.05$). In contrast, X-71 exhibited only a slight up-regulation (1.27-fold) after inoculation, which was not significant compared to its 0 hpi level ($p > 0.05$). These results indicate that variations in the regulatory region of *BGIOSGA000756* enhance its basal expression in the resistant line and facilitate further pathogen-induced activation, potentially strengthening calcium-mediated immune responses and contributing to the ShB resistance of X-11.

BGIOSGA036403 encodes an ABC transporter involved in the transmembrane transport of defense-related metabolites; thus, its expression level potentially affects the accumulation efficiency of antimicrobial compounds. The qPCR analysis (Figure 9B) revealed

FIGURE 6
Distribution of sheath blight resistance phenotypes across genotypic groups at three GWAS-associated SNP loci. **(A)** Boxplot showing resistance percentage distribution for AA vs. AC genotypes at SNP locus S1_35575127. The AA genotype exhibits significantly higher resistance (p = 0.0002). **(B)** Resistance distribution for CC vs. CT genotypes at S12_7393096. The CT genotype shows significantly higher resistance (p = 0.0016). **(C)** Phenotypic variation between AA, CA, and CC genotypes at S12_7238044. The CA genotype is associated with the highest resistance level (p = 0.001). Each boxplot represents the median, interquartile range, and outliers; statistical significance was assessed using Welch's t-test.
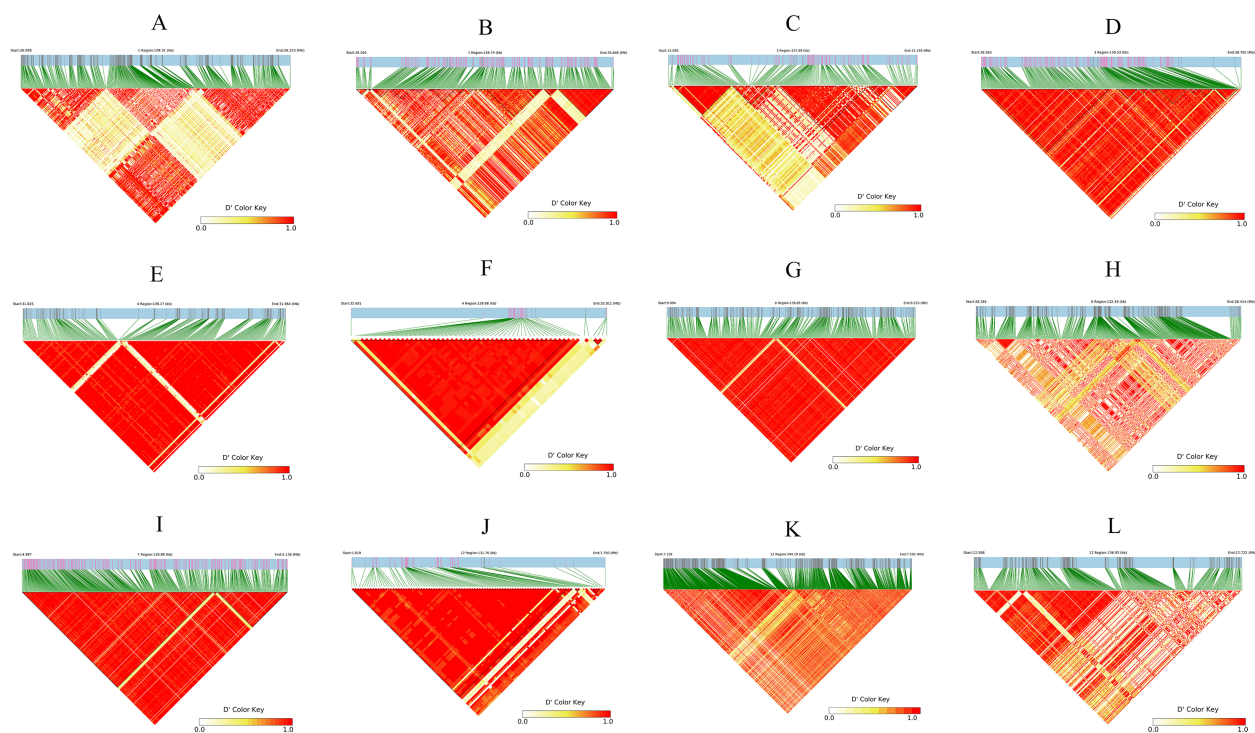
FIGURE 7
Linkage disequilibrium (LD) heatmaps of 70-kb genomic regions flanking 12 GWAS-associated SNPs. (A–L) Represent LD patterns for individual SNP loci identified in the GWAS analysis. Each heatmap displays the D' values of pairwise SNPs within a 70-kb window centered on the associated SNP (indicated at the top of each panel). The color gradient from light yellow to red reflects the degree of linkage disequilibrium (D'), ranging from 0.0 (no LD) to 1.0 (complete LD), with red indicating strong LD. Green lines denote the positions of SNPs, and red triangular regions highlight high LD blocks. Genomic coordinates (start and end positions) and region length are shown above each plot. These LD patterns help define haplotype blocks and identify candidate genes within the associated intervals.

that at 0 hpi, the relative expression in X-11 was 1.00 ± 0.13, higher than that in X-71 (0.90 ± 0.18), indicating a regulatory effect of non-coding sequence variations. At 24 hpi, the expression in X-11 increased sharply to 1.45 ± 0.08, a significant increase from its 0 hpi level (p < 0.05), and was significantly higher than the level in

X-71 (0.84 ± 0.10, p< 0.01). Notably, the expression level in X-71 showed a decreasing trend after inoculation. These findings suggest that the resistant line X-11 can rapidly upregulate the expression of this ABC transporter upon pathogen infection, likely through cis-regulatory mechanisms, thereby accelerating the transport of
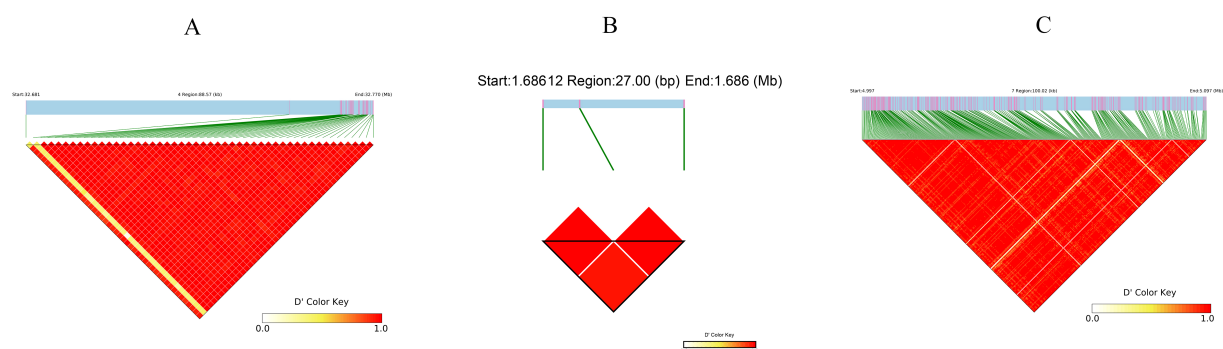

FIGURE 8
Linkage disequilibrium (LD) patterns in genomic regions harboring GWAS-associated loci. (A) LD heatmap of a 70-kb region surrounding a major resistance-associated SNP on chromosome 1 (S1_35575127), showing extensive high LD (D' > 0.8) across most SNP pairs, indicating strong haplotype block structure. (B) LD plot for a small 27-bp region near S12_7393096 on chromosome 12, where only two SNPs are present and exhibit complete LD (D' = 1.0). (C) LD heatmap of a 70-kb region flanking another significant SNP (S12_7238044), displaying moderate to high LD with several distinct blocks, suggesting potential functional haplotypes. In all panels, the color gradient from yellow to red represents D' values from 0.0 (no LD) to 1.0 (complete LD). Green lines indicate SNP positions, and diagonal bands reflect pairwise LD between SNPs.
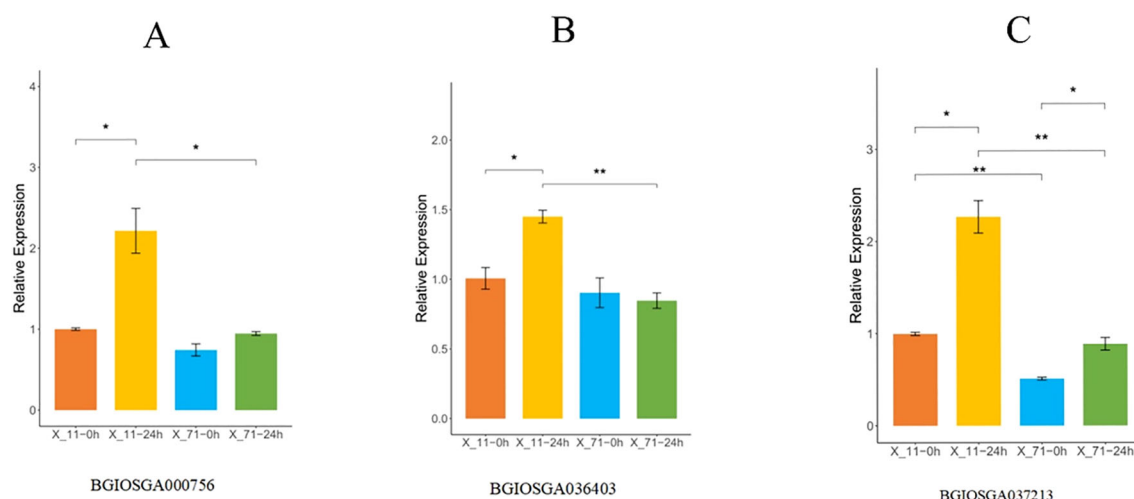
FIGURE 9
Time-course expression profiles of three candidate genes in resistant (X-11) and susceptible (X-71) rice lines following Rhizoctonia solani inoculation. **(A)** Relative expression of BGIOSGA000756 at 0 h and 24 h post-inoculation. Expression was significantly upregulated in X-11 at 24 h compared to the control (p < 0.05), while no significant change was observed in X-71. **(B)** Expression dynamics of BGIOSGA036403. The gene showed increased expression in X-11 at 24 h (p < 0.05), but remained stable in X-71. **(C)** Expression pattern of BGIOSGA037213, which exhibited strong induction in X-11 at 24 h (p < 0.01), whereas expression in X-71 remained low and unchanged. Data are presented as mean $\pm$ SD (n = 3). Statistical significance was assessed by Student's t-test: **p < 0.01, *p < 0.05 versus the respective control (or X-71 at the same time point).

antimicrobial metabolites to the infection site and/or enhancing the efflux and detoxification of fungal toxins, which aligns with the resistant phenotype of X-11.

BGIOSGA037213 encodes a BiP/LTP/SSH protein implicated in maintaining membrane integrity and exerting direct antimicrobial activity; its expression is closely associated with cellular defense capacity under pathogen stress. According to qPCR data (Figure 9C), at 0 hpi, the relative expression level in X-11 was 1.00 ± 0.03, significantly higher than that in X-71 (0.51 ± 0.03, p<0.01), demonstrating that non-coding region mutations endow the resistant material with a stronger capacity for basal synthesis of this defense protein. At 24 hpi, the expression of *BGIOSGA037213* in X-11 was further upregulated to 2.27 ± 0.31 (more than two-fold increase compared to 0 hpi) and remained significantly higher than the level in X-71 (0.89 ± 0.12, p< 0.01). Although expression increased in X-71 after inoculation, it remained lower than the basal level observed in X-11 at 0 hpi. This result indicates that the resistant line X-11 activates the expression of *BGIOSGA037213* in response to pathogen challenge, potentially enhancing membrane stability and releasing antimicrobial peptides that directly inhibit the growth of *R. solani*. In contrast, the susceptible line X-71, likely due to weaker regulatory control of gene expression, fails to mount an effective membrane defense or antimicrobial response, leading to rapid disease progression.

Pearson correlation analysis was conducted between the resistance phenotypes of X_11 and X_71, and the expression levels of three genes (BGIOSGA000756, BGIOSGA036403, and BGIOSGA037213) detected at 24 hours post-inoculation with *Rhizoctonia solani* (the causal agent of sheath blight). The results showed that the correlation coefficients between the relative expression levels of BGIOSGA000756, BGIOSGA036403, and BGIOSGA037213 and lesion length were -0.97 (p = 0.0015), -0.93

(p = 0.0064), and -0.83 (p = 0.0398), respectively. These findings indicate a significant negative correlation between the aforementioned genes and lesion length—i.e., higher gene expression induced by the pathogen corresponds to shorter lesion length and stronger sheath blight resistance.

4 Discussion

The small sample size (n = 84) imposed inherent constraints on the statistical power of the GWAS, particularly for detecting low-frequency resistance alleles. To address this limitation, we implemented a multi-model framework that integrates GLM, MLM, EMMAX, and GEMMA analyses, complemented by Venn validation—a strategy that parallels approaches in maize smut resistance GWAS with constrained populations (Wang et al., 2012). Cross-model consensus analysis enhanced reliability through methodological triangulation, mirroring successful small-sample GWAS protocols in sweet potato that combined multi-locus models with stringent multiple testing corrections (Kim et al., 2023). These optimizations effectively balanced the risks of Type I and Type II errors while maintaining the potential for discovery, demonstrating the feasibility of robust association mapping in limited germplasm panels.

Small-sample GWAS requires stringent QC to prioritize high-confidence loci. Here, MAF filtering (>0.05) and LD-guided regional heritability estimation minimized spurious associations, paralleling approaches in large-scale *indica* rice GWAS (Zhang et al., 2019). While cohort size constraints remain inherent, as demonstrated by Fu et al (Fu et al., 2022). in their multi-year validation in 150 rice accessions, integrating diverse genetic backgrounds (π = 0.0028) and kinship-adjusted models enhanced

detection accuracy, reflecting Chen et al.'s (Chen et al., 2019) subpopulation optimization strategy. Future efforts should expand germplasm diversity and incorporate multi-environment trials to dissect G × E interactions, thereby advancing translational breeding for sheath blight resistance. ABCG transporters orchestrated multi-layered defense mechanisms against *R. solani* infection. Primarily, they mediate efflux detoxification of fungal toxins (e.g., CWDEs) through substrate-specific transport, limiting cellular damage (Costanzo et al., 2011; Jia and Gealy, 2018). Alongside, these proteins regulate systemic immunity by spatially coordinating the translocation of defense hormones (SA/JA), thereby priming the expression of pathogenesis-related genes (Qin et al., 2023). Beyond their classical roles, ABCGs facilitate the apoplastic accumulation of phenylpropanoid-derived antimicrobials, including phenolics and lignin precursors, through transmembrane transport, thereby establishing dual chemical-physical barriers against pathogen invasion (Elsharkawy et al., 2022; Hunjan et al., 2022). Emerging evidence suggests non-canonical functions in pathogen-associated molecular patterns (PAMP) perception, where ABCG-mediated recognition of fungal extracellular polysaccharides (EPS) may initiate innate immune signaling, unveiling novel dimensions in plant-pathogen interface studies (PPIS) (Ma et al., 2020).

Calmodulin-like proteins (CMLs), a pivotal EF-hand domain protein family in plants, orchestrate the stress responses through calcium signaling (Zeng et al., 2023). Rice *OsCML3* (*Os01g0765600*) exemplifies this mechanism, exhibiting pathogen-induced upregulation and interacting with high-mobility group proteins (HMGB) through C-terminal extension to activate defense gene networks (Zeng et al., 2023). Calcium-dependent conformational changes in *OsCML3* modulate downstream signaling, including MAPK cascades and redox homeostasis. Specifically, *OsCML3* balances NADPH oxidase-mediated ROS production with antioxidant activity, mitigating oxidative stress during infection (Liu et al., 2022; Zeng et al., 2023). This dual regulatory capacity positions CMLs as hubs integrating calcium and redox signaling. Notably, a coding SNP (A2WVE9) in *OsCML3* alters structural dynamics and sub-cellular localization, potentially fine-tuning immune responses (Zeng et al., 2023; Yang et al., 2024), and offers targets for breeding calcium-regulated disease resistance.

Plant lipid transfer proteins (LTPs) are small, alkaline proteins widely present in higher plants, with multifunctional roles that include phospholipid transfer, cutin formation, signal transduction, and defense mechanisms against disease resistance (Collinson et al., 2024; Shang et al., 2025). Studies have shown that LTPs play essential roles in plant resistance to pathogen infection. For example, LTPs in rice enhance resistance to bacterial diseases by regulating cell membrane stability and signal transduction (Schierenbeck et al., 2021). Additionally, LTPs can interact with ROS-scavenging enzymes such as ascorbate peroxidase (APX) to regulate ROS levels and enhance plant defense against pathogens (Schierenbeck et al., 2021). In rice, overexpression of LTP significantly enhances resistance to bacterial blight, while LTP deficiency compromises this resistance, underscoring the critical role of LTPs in rice disease resistance (Schierenbeck et al., 2021). LTPs can interact with pathogenesis-related (PR) proteins, which

were produced by plants in response to pathogen attack, exhibiting antibacterial and antiviral functions; their synergistic actions more effectively inhibit pathogen invasion and reproduction (Kumar et al., 2025). Bifunctional inhibitors (BiPs) act as molecular chaperones, primarily participating in protein folding and endoplasmic reticulum stress responses. In plant disease resistance, BiPs enhance pathogen defense by regulating protein homeostasis and signal transduction. Past studies have shown that BiP expression in rice is closely associated with disease resistance (Park et al., 2010; Park et al., 2025). For example, BiP improves rice resistance to fungal diseases by regulating endoplasmic reticulum stress responses, with its expression level significantly upregulated after pathogen infection, indicating its essential role in disease resistance responses (Das et al., 2024). Furthermore, BiPs form complex defense networks through interactions with other disease resistance-related proteins, further enhancing plant disease resistance (Jain and Khurana, 2018; Zribi et al., 2024). Seed storage helical domain-containing proteins (SSHs), as proteins with specific domains, mainly participate in the accumulation and regulation of seed storage substances (Antonets et al., 2020). In disease resistance research, SSHs indirectly influence plant disease resistance by regulating storage substances in seeds (Shewry et al., 1995; Schwarzenbach, 2013). For example, the SSH expression in rice is closely correlated with the content of storage substances in seeds, which is essential for plant disease resistance (Li et al., 2018). It is reported that SSH expression levels in rice are significantly upregulated after pathogen infection, indicating their essential role in disease resistance responses (Li et al., 2018). Additionally, SSHs form complex defense networks through interactions with other disease resistance-related proteins, further enhancing plant disease resistance (Xiong et al., 2023). Future research could further explore the interactions among these proteins and their specific mechanisms in disease resistance, for example, by utilizing gene editing technologies to investigate their functions in rice disease resistance and their interactions with other proteins involved in disease resistance. Moreover, multi-omics analyses could reveal the global regulatory networks of these proteins in disease resistance.

The LRR-RLKs are transmembrane proteins comprising an extracellular LRR domain, a transmembrane domain, and an intracellular kinase domain. The LRR domain is responsible for recognizing specific ligands (such as pathogen-derived molecules), while the kinase domain activates downstream signaling pathways through phosphorylation (Xu et al., 2023; Xu et al., 2024). In rice, the LRR-RLK family comprises numerous members involved in various biological processes, including development, hormone signaling, and disease resistance responses (Block et al., 2021; Durgadevi et al., 2021). The rice LRR receptor-like protein (*OsRLP1*) and its adaptor kinase *OsSOBIR1* jointly mediate resistance to viral infections (Zhang et al., 2021). Through high-throughput sequencing and transgenic rice experiments, researchers found that *OsRLP1* and *OsSOBIR1* are significantly upregulated after viral infection and can activate plant immune responses, revealing the vital role of LRR-RLKs in plant antiviral defense (Zhang et al., 2021).

In this study, we analyzed the expression patterns of three candidate genes (*BGIOSGA000756*, *BGIOSGA036403*, and

BGIOSGA037213) associated with GWAS loci located in regulatory or intronic regions. Under non-infected conditions, the resistant line X-11 exhibited higher basal expression levels of all three genes compared to the susceptible line X-71. This suggests that non-coding variations at these GWAS loci may enhance promoter activity or transcription efficiency, thereby providing a pre-activated defense foundation in the resistant material even in the absence of pathogen stress. At 24 hours post-inoculation (hpi) with *R. solani*, all three genes were significantly induced (1.4- to 2.2-fold) in X-11, whereas their induction was markedly weaker or even decreased in X-71. This result further indicates that these GWAS loci are involved not only in regulating basal expression but also in activating the pathogen-induced expression pathway. We speculate that the regulatory variations may contain pathogen-responsive elements that are specifically functional in the resistant genotype, enabling efficient activation of defense responses upon pathogen recognition. Collectively, these non-coding variations likely contribute to the resistant phenotype by modulating gene expression levels rather than altering protein sequences.

The high expression of these genes in X-11 likely underpins its resistance mechanism through distinct roles. *BGIOSGA000756*, encoding an EF-hand domain protein, may facilitate the rapid transmission of pathogen invasion signals through calcium signaling, potentially activating downstream MAPK cascades and ROS bursts, thereby initiating subsequent defense reactions. The elevated expression of *BGIOSGA036403*, an ABC transporter, could promote the transport of antimicrobial metabolites to the infection site, forming a chemical barrier, while also effluxing fungal toxins to reduce cellular damage. Meanwhile, *BGIOSGA037213*, a BiP/LTP/SSH protein, may enhance resistance by maintaining membrane integrity against pathogen degradation and potentially directly inhibiting *R. solani* hyphal growth through the production of antimicrobial peptides. While this study advances our understanding of ShB resistance genetics, several limitations warrant consideration. The modest cohort size may limit detection of low-frequency resistance alleles, and reliance on controlled *in vitro* assays may not fully capture field resistance dynamics. Future investigations should incorporate multi-environment trials to validate ecological relevance. Although candidate genes were prioritized, functional validation through CRISPR-Cas9 editing and mechanistic studies (e.g., transcriptomics, yeast two-hybrid assays) remains imperative to delineate their roles in defense signaling. Integrated multi-omics approaches (proteomics/metabolomics) could further unravel regulated pathways and protein interaction networks, advancing resistance breeding strategies.

Although this study enhances our understanding of the genetic basis of sheath blight (ShB) resistance, several limitations should be noted. Future work should prioritize the functional characterization of the three core candidate genes (*BGIOSGA000756*, *BGIOSGA036403*, *BGIOSGA037213*) and 22 consensus SNPs identified here, to elucidate their roles and potential utility in improving ShB resistance in indica rice. To this end, knockout lines for each candidate gene could be generated in the resistant accession X-11, while overexpression lines could be created in the susceptible accession X-71. Using the *in vitro* leaf segment inoculation method developed in this study, along with whole-plant inoculation assays under greenhouse conditions, comparative phenotypic analyses can be conducted to assess the impact of gene knockout or overexpression on resistance-related traits, such as lesion length ratio and lesion height. In parallel, yeast two-hybrid assays could be employed to identify interacting partners of the candidate proteins, helping to decipher their molecular mechanisms in calcium signaling, antimicrobial metabolite transport, or maintenance of cell membrane integrity.

Furthermore, the 22 consensus SNPs could be converted into molecular markers for high-throughput genotyping across a broader collection of indica rice germplasm. By combining such marker screening with multi-environment field trials, researchers could validate the associations between these SNPs and ShB resistance. This would ultimately provide valuable genetic resources and practical tools for marker-assisted breeding of ShB-resistant indica varieties. Together, these approaches would address the current limitations in establishing gene causality and translating laboratory findings into field applications, thereby accelerating the development of ShB-resistant *indica* rice through improved breeding efficiency.

5 Conclusions

This study employed GWAS on 84 *indica* rice accessions to dissect the genetic architecture of ShB resistance. Multi-model integration (GLM/MLM/EMMAX/GEMMA) and Venn consensus analysis identified 22 robustly associated SNPs, with significant enrichment on chromosome 12, suggesting major-effecting QTLs. Haplotype-based fine mapping delineated three critical LD blocks harboring 15 candidate genes, including ABC transporters, calmodulin-like proteins, lipid transfer proteins, and LRR receptor-like kinases. Functional annotation associates these candidates with defense metabolite trafficking, calcium signaling modulation, and pathogen recognition, key mechanisms underpinning basal immunity. The findings demonstrate that non-coding variations in three GWAS-linked candidate genes enhance disease resistance in rice line X-11 by conferring strong basal and pathogen-induced expression, as validated by qPCR analysis following *R. solani* challenge. The identified loci establish a genetic framework for molecular breeding strategies. Future efforts should expand germplasm diversity, incorporate multi-environment trials, and deploy CRISPR-Cas9-mediated functional validation to accelerate the development of ShB-resistant *indica* varieties.

Data availability statement

Data generated or analyzed during this study are included in this published article and its supplementary information files. The datasets presented in this study are available in online repositories. The names of the repository and accession number(s) can be found below: NCBI https://www.ncbi.nlm.nih.gov/, Accession number: PRJNA1358769.

Author contributions

Funding

Acknowledgments

Conflict of interest

The authors declared that this work 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) declared that generative AI was not 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.2025.1718389/full#supplementary-material

References

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109

Antonets, K. S., Belousov, M. V., Sulatskaya, A. I., Belousova, M. E., Kosolapova, A. O., Sulatsky, M. I., et al. (2020). Accumulation of storage proteins in plant seeds is mediated by amyloid formation. *PloS Biol.* 18, e3000564. doi: 10.1371/journal.pbio.3000564

Baiya, S., Hua, Y., Ekkhara, W., and Ketudat Cairns, J. R. (2014). Expression and enzymatic properties of rice (*Oryza sativa* L.) monolignol beta-glucosidases. *Plant science: an Int. J. Exp. Plant Biol.* 227, 101–109. doi: 10.1016/j.plantsci.2014.07.009

Belay, S., Belay, G., Nigussie, H., Jian-Lin, H., Tijjani, A., Ahbara, A. M., et al. (2024). Whole-genome resource sequences of 57 indigenous Ethiopian goats. *Sci. Data.* 11, 1–10. doi: 10.1038/s41597-024-02973-2

Block, A. K., Tang, H. V., Hopkins, D., Mendoza, J., Solemslie, R. K., du Toit, L. J., et al. (2021). A maize leucine-rich repeat receptor-like protein kinase mediates responses to fungal attack. *Planta.* 254, 1–9. doi: 10.1007/s00425-021-03730-0

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics.* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Chen, W., Gao, Y., Xie, W., Gong, L., Lu, K., Wang, W., et al. (2014). Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat. Genet.* 46, 714–721. doi: 10.1038/ng.3007

Chen, Z., Feng, Z., Kang, H., Zhao, J., Chen, T., Li, Q., et al. (2019). Identification of new resistance loci against sheath blight disease in rice through genome-wide association study. *Rice Sci.* 26, 21–31. doi: 10.1016/j.rsci.2018.12.002

Collinson, S., Cairns, J. E., Ndlala, L., Deonovic, B., Ligeyo, D., Albertsen, M., et al. (2024). Ms44-SPT: unique genetic technology simplifies and improves hybrid maize seed production in sub-Saharan Africa. *Sci. Rep.* 14, 1–10. doi: 10.1038/s41598-024-83931-1

Costanzo, S., Jackson, A. K., and Brooks, S. A. (2011). High-resolution mapping of Rsn1, a locus controlling sensitivity of rice to a necrosis-inducing phytotoxin from *Rhizoctonia solani* AG1-IA. *Theor. Appl. Genet.* 123, 33–41. doi: 10.1007/s00122-011-1564-1

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

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience.* 10, 1–4. doi: 10.1093/gigascience/giab008

Das, D., Chowdhury, N., Sharma, M., Suma, R., Saikia, B., Velmurugan, N., et al. (2024). Screening for brown-spot disease and drought stress response and identification of dual-stress responsive genes in rice cultivars of Northeast India. *Physiol. Mol. Biol. Plants*. 30, 647–663. doi: 10.1007/s12298-024-01447-4

Durgadevi, D., Harish, S., Manikandan, R., Prabhukarthikeyan, S. R., Alice, D., and Raguchander, T. (2021). Proteomic profiling of defense/resistant genes induced during the tripartite interaction of *Oryza sativa*, *Rhizoctonia solani* AG1-1A, and Bacillus subtilis against rice sheath blight. *Physiol. Mol. Plant Pathol.* 115, 101669. doi: 10.1016/j.pmpp.2021.101669

Elsharkawy, M. M., Sakran, R. M., Ahmad, A. A., Behiry, S. I., Abdelkhalek, A., Hassan, M. M., et al. (2022). Induction of systemic resistance against sheath blight in rice by different pseudomonas isolates. *Life* 12, 349. doi: 10.3390/life12030349

Fu, D., Zhong, K., Zhong, Z., Hu, G., Zhang, P., and Tong, H. (2022). Genome-Wide Association Study of Sheath Blight Resistance within a Core Collection of Rice (*Oryza sativa* L.). *Agron.* 12, 1493. doi: 10.3390/agronomy12071493

Ganesh, G. C., and Roka, P. (2024). Sheath Blight of rice: A review of host plant interaction and disease management. *Arch. Agric. Environ. Science.* 9, 404–408. doi: 10.26832/24566632.2024.0902027

Hassan, M. A., Dahu, N., Hongning, T., Qian, Z., Yueming, Y., Yiru, L., et al. (2023). Drought stress in rice: morpho-physiological and molecular responses and marker-assisted breeding. *Front. Plant Sci.* 14, 1215371. doi: 10.3389/fpls.2023.1215371

Hunjan, M. S., Kumar, S., Lore, J. S., and Cruz, C. M. V. (2022). Efficiency of different *Rhizoctonia solani* inoculum source against sheath blight screening in rice under field conditions. *Trop. Plant Pathol.* 47, 309–313. doi: 10.1007/s40858-021-00489-3

Jain, D., and Khurana, J. P. (2018). Role of pathogenesis-related (PR) proteins in plant defense mechanism. *Mol. Aspects Plant-Pathogen Interaction*, 265–281. doi: 10.1007/978-981-10-7371-7_12

Jia, Y., and Gealy, D. (2018). Weedy red rice has novel sources of resistance to biotic stress. *Crop J.* 6, 443–450. doi: 10.1016/j.cj.2018.07.001

Jones, C. S., Sychantha, D., Lynne Howell, P., and Clarke, A. J. (2020). Structural basis for the O-acetyltransferase function of the extracytoplasmic domain of OatA from Staphylococcus aureus. *J. Biol. Chem.* 295, 8204–8213. doi: 10.1074/jbc.RA120.013108

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

Khoshkdaman, M., Mousanejad, S., Elahinia, S. A., Ebadi, A. A., Padasht-Dehkaei, F., Khoshkdaman, M., et al. (2021). Sheath blight development and yield loss on rice in different epidemiological conditions. *JPlaP.* 103, 87–96. doi: 10.1007/s42161-020-00653-9

Kim, T. H., Kim, S., Park, W., Woo, K. S., Lee, K., Chung, M. N., et al. (2023). Genome-wide association study to identify novel loci and genes for Fusarium root rot resistance in sweet potato using genotyping-by-sequencing. *Front. Plant Sci.* 14, 1251157. doi: 10.3389/fpls.2023.1251157

Kumar, P., Pandey, S., and Pati, P. K. (2025). Interaction between pathogenesis-related (PR) proteins and phytohormone signaling pathways in conferring disease tolerance in plants. *Physiol. Plant* 177, e70174. doi: 10.1111/ppl.70174

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Li, J., Pan, Y., Guo, H., Zhou, L., Yang, S., Zhang, Z., et al. (2018). Fine mapping of QTL qCTB10–2 that confers cold tolerance at the booting stage in rice. *Theor. Appl. Genet.* 131, 157–166. doi: 10.1007/s00122-017-2992-3

Li, D., Zhang, F., Pinson, S. R. M., Edwards, J. D., Jackson, A. K., Xia, X., et al. (2022). Assessment of rice sheath blight resistance including associations with plant architecture, as revealed by genome-wide association studies. *Rice* 15, 1–30. doi: 10.1186/s12284-022-00574-4

Liao, H., Fang, Y., Yin, J., He, M., Wei, Y., Zhang, J., et al. (2025). Rice transcription factor bHLH25 confers resistance to multiple diseases by sensing H2O2. *Cell Res.* 35, 205–219. doi: 10.1038/s41422-024-01058-4

Liu, Y., Chen, W., Liu, L., Su, Y., Li, Y., Jia, W., et al. (2022). Genome-wide identification and expression analysis of calmodulin and calmodulin-like genes in wheat (Triticum aestivum L.). *Plant Signal Behav.* 17, 1–11. doi: 10.1080/15592324.2021.2013646

Lu, Y., Li, M., Gao, Z., Ma, H., Chong, Y., Hong, J., et al. (2025). Advances in whole genome sequencing: methods, tools, and applications in population genomics. *Int. J. Mol. Sci.* 26, 372. doi: 10.3390/ijms26010372

Ma, H., Sheng, C., Qiao, L., Zhao, H., and Niu, D. (2020). A comparative proteomic approach to identify defence-related proteins between resistant and susceptible rice cultivars challenged with the fungal pathogen Rhizoctonia solani. *Plant Growth Regul.* 90, 73–88. doi: 10.1007/s10725-019-00551-w

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Molla, K. A., Karmakar, S., Molla, J., Bajaj, P., Varshney, R. K., Datta, S. K., et al. (2020). Understanding sheath blight resistance in rice: the road behind and the road ahead. *Plant Biotechnol. J.* 18, 895–915. doi: 10.1111/pbi.13312

Oreiro, E. G., Grimares, E. K., Atienza-Grande, G., Quibod, I. L., Roman-Reyna, V., and Oliva, R. (2020). Genome-wide associations and transcriptional profiling reveal ROS regulation as one underlying mechanism of sheath blight resistance in rice. *Mol. Plant-Microbe Interactions*. 33, 212–222. doi: 10.1094/MPMI-05-19-0141-R

Park, C. J., Bart, R., Chern, M., Canlas, P. E., Bai, W., and Ronald, P. C. (2010). Overexpression of the endoplasmic reticulum chaperone biP3 regulates XA21-mediated innate immunity in rice. *PloS One* 5, e9262. doi: 10.1371/journal.pone.0009262

Park, S. R., Jeong, Y., and Son, S. (2025). Functions of transcription factor superfamilies in rice immunity. *Crop J.* 13, 5–22. doi: 10.1016/j.cj.2024.10.006

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Qin, X., Chang, Y., Wang, Y., Yang, J., Nie, S., Guo, N., et al. (2023). Aspergillus sp. R3, a new producer for cyclopyazonic acid, inhibits rice sheath blight fungus *Rhizoctonia solani* Kühn. *Physiol. Mol. Plant Pathol.* 125, 102007. doi: 10.1016/j.pmpp.2023.102007

Richa, K., Tiwari, I. M., Kumari, M., Devanna, B. N., Sonah, H., Kumari, A., et al. (2016). Functional characterization of novel chitinase genes present in the sheath blight resistance QTL: Qsbr11–1 in rice line Tetep. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00244

Schierenbeck, M., Alqudah, A. M., Lohwasser, U., Tarawneh, R. A., Simón, M. R., and Börner, A. (2021). Genetic dissection of grain architecture-related traits in a winter wheat population. *BMC Plant Biol.* 21, 1–14. doi: 10.1186/s12870-021-03183-3

Schwarzenbach, H. (2013). "Loss of heterozygosity brenner's encyclopedia of genetics, 271–273," in *Brenner's encyclopedia of genetics* (San Diego, CA, USA: Academic Press), *2nd ed.*, 271–273.

Shang, K., Wang, C., Wang, X., Wang, Y., Xu, K., Zhou, S., et al. (2025). Non-specific lipid transfer protein stLTP6 promotes virus infection by inhibiting jasmonic acid signalling pathway in response to PVS TGB1. *Plant Cell Environ.* 48, 2343–2356. doi: 10.1111/pce.15298

Shewry, P. R., Napier, J. A., and Tatham, A. S. (1995). Seed storage proteins: structures and biosynthesis. *Plant Cell.* 7, 945. doi: 10.1105/tpc.7.7.945

Srivastava, D., Shamim, M., Ahmad, M. M., and Upadhyay, R. S. (2024). Fungal diseases of rice and their management. *Fungal Dis. Rice Their Manage.*, 1–424. doi: 10.1201/9781003332169

Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120

Turaidar, V., Reddy, M., Anantapur, R., Krupa, K. N., Dalawai, N., Deepak, C. A., et al. (2018). Rice sheath blight: major disease in rice. *Int. J. Curr. Microbiol. App. Sci.* 7, 976–988.

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., et al. (2021). Genome-wide association studies. *Nat. Rev. Methods Primers.* 1, 1–21. doi: 10.1038/s43586-021-00056-9

Uppala, S. S., and Zhou, X.-G. (2018). Rice sheath blight. *Plant Health Instructor* 18. doi: 10.1094/PHI-I-2018-0403-01

Wang, M. H., Cordell, H. J., and Van Steen, K. (2019). Statistical methods for genome-wide association studies. *Semin. Cancer Biol.* 55, 53–60. doi: 10.1016/j.semcancer.2018.04.008

Wang, D., Lu, Q., Jin, S., Fan, X., and Ling, H. (2023). Pectin, lignin and disease resistance in brassica napus L.: an update. *Horticulturae* 9, 112. doi: 10.3390/horticulturae9010112

Wang, A., Shu, X., Jing, X., Jiao, C., Chen, L., Zhang, J., et al. (2021). Identification of rice (*Oryza sativa* L.) genes involved in sheath blight resistance via a genome-wide association study. *Plant Biotechnol. J.* 19, 1553–1566. doi: 10.1111/pbi.13569

Wang, Q., Tang, J., Han, B., and Huang, X. (2020). Advances in genome-wide association studies of complex traits in rice. *Theor. Appl. Genet.* 133, 1415–1425. doi: 10.1007/s00122-019-03473-3

Wang, D., Wang, W., Zang, S., Qin, L., Liang, Y., Lin, P., et al. (2024). Sugarcane transcription factor ScWRKY4 negatively regulates resistance to pathogen infection through the JA signaling pathway. *Crop J.* 12, 164–176. doi: 10.1016/j.cj.2023.12.002

Wang, M., Yan, J., Zhao, J., Song, W., Zhang, X., Xiao, Y., et al. (2012). Genome-wide association study (GWAS) of resistance to head smut in maize. *Plant Science.* 196, 125–131. doi: 10.1016/j.plantsci.2012.08.004

Xiong, D., Xu, Q., Tian, L., Bai, J., Yang, L., Jia, J., et al. (2023). Mechanism of improving solubility and emulsifying properties of wheat gluten protein by pH cycling treatment and its application in powder oils. *Food Hydrocoll.* 135, 108132. doi: 10.1016/j.foodhyd.2022.108132

Xu, B., Cerbu, A., Tralie, C. J., Lim, D., and Krasileva, K. (2024). Structure-aware annotation of leucine-rich repeat domains. bioRxiv. *PloS Comput. Biol.* 20, e1012526. doi: 10.1371/journal.pcbi.1012526

Yang, D., Chen, T., Wu, Y., Tang, H., Yu, J., Dai, X., et al. (2024). Genome-wide analysis of the peanut CaM/CML gene family reveals that the AhCML69 gene is associated with resistance to Ralstonia solanacearum. *BMC Genomics* 25, 1–16. doi: 10.1186/s12864-024-10108-5

Yang, Z., Li, N., Kitano, T., Li, P., Spindel, J. E., Wang, L., et al. (2021). Genetic mapping identifies a rice naringenin O-glucosyltransferase that influences insect resistance. *Plant J.* 106, 1401–1413. doi: 10.1111/tpj.15244

Yu, H., Li, X., Yu, Y., Sui, Y., Zhang, J., Zhang, L., et al. (2025). A dual-branch multimodal model for early detection of rice sheath blight: Fusing spectral and physiological signatures. *Comput. Electron Agric.* 231, 110031. doi: 10.1016/j.compag.2025.110031

Zeng, H., Zhu, Q., Yuan, P., Yan, Y., Yi, K., and Du, L. (2023). Calmodulin and calmodulin-like protein-mediated plant responses to biotic stresses. *Plant Cell Environ.* 46, 3680–3703. doi: 10.1111/pce.14686

Zhai, H., Zhou, C., Zhang, Y., Wang, Y., Wang, M., Tianya, L., et al. (2025). Mechanism analysis of osbHLH34-osERF34 mediated regulation of rice resistance to sheath blight. *Int. J. Mol. Sci.* 26, 2249. doi: 10.3390/ijms26052249

Zhang, H., Chen, C., Li, L., Tan, X., Wei, Z., Li, Y., et al. (2021). A rice LRR receptor-like protein associates with its adaptor kinase *OsSOBIR1* to mediate plant immunity against viral infection. *Plant Biotechnol. J.* 19, 2319–2332. doi: 10.1111/pbi.13663

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. *Bioinformatics.* 35, 1786–1788. doi: 10.1093/bioinformatics/bty875

Zhang, F., Zeng, D., Zhang, C. S., Lu, J. L., Chen, T. J., Xie, J. P., et al. (2019). Genome-wide association analysis of the genetic basis for sheath blight resistance in rice. *Rice.* 12, 1–13. doi: 10.1186/s12284-019-0351-5

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44, 821–824. doi: 10.1038/ng.2310

Zhu, Q., Hassan, M. A., Li, Y., Fang, W., Wu, J., and Wang, S. (2024). Physiological and transcriptomic characterization of rice genotypes under drought stress. *Agronomy.* 14, 2247. doi: 10.3390/agronomy14102247

Zribi, I., Ghorbel, M., and Brini, F. (2024). Pathogenesis-related proteins and plant defense response. *Defense-Related Proteins Plants.*, 53–97. doi: 10.1016/B978-0-443-13236-0.00003-8