


RESEARCH

Open Access



Integrated meta-analysis and transcriptomics pinpoint genomic loci and novel candidate genes associated with submergence tolerance in rice

Kelvin Dodzi Aloryi^{1†}, Nnaemeka Emmanuel Okpala^{1†}, Hong Guo², Benjamin Karikari^{3,4}, Aduragbemi Amo^{5,6}, Semiu Folaniyi Bello⁷, Dinesh Kumar Saini^{8,9}, Selorm Akaba¹⁰ and Xiaohai Tian^{1*} 

Abstract

Background Due to rising costs, water shortages, and labour shortages, farmers across the globe now prefer a direct seeding approach. However, submergence stress remains a major bottleneck limiting the success of this approach in rice cultivation. The merger of accumulated rice genetic resources provides an opportunity to detect key genomic loci and candidate genes that influence the flooding tolerance of rice.

Results In the present study, a whole-genome meta-analysis was conducted on 120 quantitative trait loci (QTL) obtained from 16 independent QTL studies reported from 2004 to 2023. These QTL were confined to 18 meta-QTL (MQTL), and ten MQTL were successfully validated by independent genome-wide association studies from diverse natural populations. The mean confidence interval (CI) of the identified MQTL was 3.44 times narrower than the mean CI of the initial QTL. Moreover, four core MQTL loci with genetic distance less than 2 cM were obtained. By combining differentially expressed genes (DEG) from two transcriptome datasets with 858 candidate genes identified in the core MQTL regions, we found 38 common differentially expressed candidate genes (DECGs). *In silico* expression analysis of these DECGs led to the identification of 21 genes with high expression in embryo and coleoptile under submerged conditions. These DECGs encode proteins with known functions involved in submergence tolerance including WRKY, F-box, zinc fingers, glycosyltransferase, protein kinase, cytochrome P450, PP2C, hypoxia-responsive family, and DUF domain. By haplotype analysis, the 21 DECGs demonstrated distinct genetic differentiation and substantial genetic distance mainly between *indica* and *japonica* subspecies. Further, the MQTL7.1 was successfully validated using flanked marker S2329 on a set of genotypes with phenotypic variation.

[†]Kelvin Dodzi Aloryi and Nnaemeka Emmanuel Okpala contributed equally to this work.

*Correspondence:
Xiaohai Tian
xiaohait@sina.com

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusion This study provides a new perspective on understanding the genetic basis of submergence tolerance in rice. The identified MQTL and novel candidate genes lay the foundation for marker-assisted breeding/engineering of flooding-tolerant cultivars conducive to direct seeding.

Keywords Meta-analysis, MQTL, Submergence tolerance, Haplotype analysis, Differentially expressed candidate genes (DECGs), Marker-assisted breeding

Background

Rice (*Oryza sativa* L.) remains the most-eaten staple food for over half of the global population [1]. By 2025, the global rice demand is expected to increase by 25%, which requires a production target of 732.5 MT, with an annual increase in production of 5.9 MT [2, 3]. The irrigated rice ecosystem accounts for 55% of cultivated areas and provides 75% of the world's rice production [4]. However, the current trend in changing climates has further exacerbated the decline in rice production worldwide [5]. The practice of transplanting rice seedlings into puddled fields is widely adopted in most rice production areas in the world. Currently, direct seeding under rainfed and irrigated conditions is preferred by farmers due to the low cost of labour and convenience [6–8]. Farmers usually flood direct-seeded fields to suppress weed growth, thus reducing manual weeding and excessive herbicide use [3, 9]. However, this practice creates stress associated with submergence [10]. Rice is highly sensitive to anoxic conditions during seed germination, thus commercializing direct seeding requires breeding of rice varieties with increased submergence tolerance [11].

To withstand anoxic stress during seed germination, rice typically adopts two distinct strategies such as escape and quiescence [12]. The quiescence approach is mainly associated with the elongation of shoot by *SUB1* and survival by limiting carbohydrate consumption [13, 14]. On the other hand, the escape strategy enhanced rapid elongation of coleoptile and mesocotyl to access the water surface and transport oxygen to submerged tissues [15, 16]. Consequently, the success of seedling establishment depends on the length and the speed of coleoptile and mesocotyl elongation. However, the genetic basis of this event is less established. A considerable number of quantitative trait loci (QTL) controlling submergence germination have been identified in the past few years [17–22]. These studies detected numerous QTL in diverse mapping populations with distinct marker types [17–24]. However, there are inherent difficulties associated with the implementation of these QTL in breeding programs. The conventional QTL mapping method could lead to a wide confidence interval (CI) and undetermined QTL location. In addition, factors including mapping populations, experimental conditions, QTL mapping models, marker density, and environment affect QTL usage in marker-assisted breeding [25, 26].

To resolve these limitations, a new paradigm in QTL analysis (meta-analysis) approach was introduced [27]. This approach takes into account QTL information from multiple studies and integrates them into a consensus map enabling robust statistical power over a large dataset [28, 29]. Moreover, this method has efficiently reduced the confidence region of MQTL leading to high-resolution marker-assisted breeding and candidate gene (CG) mining [30]. Although meta-QTL analyses have been performed on abiotic-related stress and MQTL loci associated with salt, drought, cold stress, and metal ion tolerance have been identified in rice [31, 32, 33], no MQTL information is available on submergence tolerance in rice during germination. Moreover, the advent of high-throughput technology based on SNP arrays and the use of genome-wide association studies (GWAS) to dissect significant genetic loci governing complex quantitative traits is gaining popularity [34]. The available evidence has demonstrated that GWAS and MQTL could be combined to effectively explore important genomic loci and decipher genetic underlying key quantitative traits [35].

Herein, we aimed to perform a meta-QTL analysis of QTL regulating rice submergence tolerance identified in recent years and incorporate GWAS and transcriptome studies to dissect the genomic regions and potential candidate genes underpinning submergence tolerance during germination. The results from this study would aid MQTL-assisted breeding for rice cultivars suitable for direct seeding. In addition, the current study provides baseline information for understanding the metabolic processes underlying submergence tolerance or susceptibility in rice.

Methods

Screening of submergence-related traits QTL

A comprehensive search for QTL underlying submergence tolerance in rice was conducted on recently published articles from 2004 to 2023 and a total of 16 QTL studies were obtained that could furnish the initial information required for meta-analysis. The salient information about these QTL are summarized in Table 1. For each QTL study, the submergence-related traits were mainly anaerobic germination (AG), submergence (SUB), low-temperature germination (LTG), and lowest elongated internode (LEI). The initial information collected on each QTL included: (i) size of the mapping population; (ii) associated traits; (iii) logarithm of odd (LOD)

Table 1 Summary of QTL studies utilized for meta-QTL analysis

Parents	Population type ^a	Population size	No. of markers	Marker type ^b	Traits ^c	References
IR64/KHAIYAN//IR64	BC ₂	150	155	SSR	AG	[92]
IR64/Khao Hlan	BC ₂	398	136	SSR	AG	[22]
IR64 x Nanhi	F ₂	300	234	SNP	AG	[93]
IR64 x Kharsu80A	F ₂	190	384	SNP	AG	[11]
BJ1 x NSIC Rc222	BC ₂	205	102	SSR	AG	[19]
FR13A x IR42	RIL	103	232	SSR, SNP, STS	SUB	[18]
IR42 x FR13A	RIL	103	232	SNP, SSR, STS	AG	[18]
Kinmaze (japonica) x DV85 (indica)	RIL	81	137	RFLP	AG	[94]
USSR5 x N22	F ₂	148	121	SSR	LTG, AG	[95]
Tai Nguyen x Anda.	F ₂	285	1405	SNP	AG	[96]
Kinmaze (japonica)/DV85 (indica)	RIL	81	137	RFLP, SSR	AG	[97]
Indra x AC 39,416 A	RIL	184	104	SSR	AG	[21]
Swarna Sub1 x AC39416A	F ₂	188	83	SSR	AG	[20]
IR72 x Madabaru	F ₂	466	115	SSR	SUB	[98]
IR42 x Ma-Zhan Red	F ₂	175	118	SSR	AG	[99]
Habiganj x Patnai23	F ₂	192	85	RFLP, SSR	LEI	[100]

^aBC: backcross population, F₂: second filial generation population, RIL: recombinant inbred lines. ^bSSR: simple sequence repeat, SNP: single nucleotide polymorphism, STS: sequence-tagged sites, RFLP: restriction fragment length polymorphism. ^cAG: anaerobic germination, SUB: submergence, LTG: low-temperature germination, LEI: lowest elongated internode

score; (iv) phenotypic variance explained (PVE) by initial QTL; (v) closely linked molecular markers; and (vi) type of mapping population (F₂, Backcross, and recombinant inbred lines (RILs)). The submergence-associated QTL in each study with phenotypic variance explained (PVE) values less than 10% were retained for subsequent analyses. For QTL with missing LOD and PVE score, 3 and 10% were used according to Venske et al. [36] and Kahanani et al. [37]. In addition, the 95% CI of each QTL was recalculated based on the type and size of the mapping population following the standardized formula: (1) F₂ and backcross population, $CI = 530 / (\text{Mapping population} \times R^2)$; (2) RIL, $CI = 163 / (\text{Mapping population} \times R^2)$. Where R² represents the PVE by individual QTL and the mapping population was the number of lines used in the initial mapping studies [26, 38]. The details of these initial QTL are presented in Supplementary Table 1.

Construction of the consensus map and QTL projection

A high-density consensus map was constructed using the R package *LPmerge* employed in our previously developed reference map [39] and one constructed by Prakash et al. [40]. Also, markers linked to individual QTL were included in the construction of the consensus map to ensure high number of initial QTL is projected. QTL absent on the reference map were eliminated. Based on QTL projection, two approaches were adopted. First, the method proposed by Goffinet and Gerber, [27] was used when the number of QTL present on individual chromosomes is less than 10. Sequel to this, the best model with the lowest Akaike information criterion (AIC) scores in BioMercator (<https://versailles.inra.fr/Tools/BioMercator-v4>, accessed on 1, 4 2023) was considered. Second,

a two-step method proposed by Veyrieras et al. [29] was applied when the number of QTL on each chromosome exceeded 10. In the first stage, QTL on each chromosome were clustered using standard parameters. The MQTL on individual chromosomes were then estimated based on four criteria such as AIC model, corrected AIC (AICc and AIC3) model, Average Weight of Evidence (AWE) model, and Bayesian Information Criterion (BIC) model. The model with the lowest values of these criteria was chosen for the next step of analysis. In the second stage, the 95% CI and MQTL positions were ascertained following the best-selected model in the previous step. The QTL were then integrated to obtain the peak position of the initial QTL involved in MQTL CI [41]. However, QTL that did not meet the minimum AIC threshold were eliminated. MQTL were named according to their genetic positions on the rice chromosomes (i.e. MQTL1.1, MQTL2.1, MQTL3.1). The phenotypic variance explained (PVE) by MQTL was calculated as the mean PVE of initial QTL involved in each MQTL.

Validation of MQTL by GWAS-based marker-trait associations (MTAs)

MQTL identified in the current study were mapped to the rice reference genome. The markers flanking individual MQTL were manually searched, and their primer sequences were retrieved from the gramene database (<https://archive.gramene.org/>, accessed on 14, 4 2023). The flanking marker and primer sequences were blasted against the rice reference genome sequence in phytozome (<https://phytozome-next.jgi.doe.gov/>, accessed on 15, 4 2023) to obtain the physical coordinates of these markers. For markers without physical coordinates, the

physical positions were manually anchored. The marker-trait association (MTA) reported from nine independent GWAS involving submergence-associated traits published from 2015 to 2023 were collected to authenticate the accuracy of MQTL loci. The detailed information on these GWAS is presented in Supplementary Table 2. The phenotypic data of these studies included mini-core collections obtained from 4 different countries with panel sizes ranging from 153 to 498. Physical coordinates of significant SNPs associated with submergence-related traits were collected from respective studies. Finally, the physical coordinates of the MQTL identified in this study were compared with the physical locations of MTA in GWAS. MQTL co-localized with significant single nucleotide polymorphisms (SNPs) were deemed GWAS-validated/verified MQTL.

Candidate genes (CGs) excavation in core MQTL regions

The Rice Genome Annotation Project (RGAP) database (<https://rice.uga.edu/cgi-bin/gbrowser/rice>, accessed on 20, 4 2023) was used to retrieve the functional candidate genes at CI of core MQTL. The above database was then used to obtain the description of the CGs in the core MQTL intervals. To enhance the accuracy of our assumptions on CGs, the orthologous genes in the core MQTL regions were placed over the genomes of sorghum and maize. The TBtools software [42] was used to visualize the gene density and orthologue genes in sorghum and maize.

Validation of CGs in transcriptome datasets and gene expression analysis

A survey of two differentially expressed genes (DEGs) datasets involved in submergence tolerance during seed germination from independent studies [43, 44] was explored to validate candidate genes (CGs) found in core MQTL loci. A Venn diagram was employed to compare the common genes regulating submergence tolerance detected by two DEGs studies, and the genes located in MQTL regions. Venny online repository (<https://bioinformatics.psb.ugent.be/webtools/Venn/>, accessed on 1, 5 2023) was used to analyze shared candidate genes between MQTL regions, and differentially expressed genes (DEGs) datasets. The shared genes are referred to in this study as differentially expressed candidate genes (DECGs).

Interaction network analysis, GO and KEGG analysis of DECGs

STRING database (Search Tool for the Retrieval of Interacting Genes) (<https://string-db.org/>, accessed on 4, 5 2023) [45] was used to construct the protein-protein interaction of DECGs. Functional analysis of DECGs in core MQTL interval was conducted using CARMO

Comprehensive Annotation of Rice Multi-Omics tool (http://bioinfo.sibs.ac.cn/carmo/Gene_Annotation.php, accessed on 5, 5 2023). Then Gene ontology (GO) terms of DECGs were visualized using TBtools software [42]. The KOBAS v3.0 software (<http://kobas.cbi.pku.edu.cn/>, accessed on 5, 5 2023) was employed to map DECGs in Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway. The bioinformatics online tool (<http://www.bioinformatics.com.cn>, accessed on 5, 5 2023) was exploited to display the KEGG results.

Expression profiling and promoter analysis of DECGs

In silico expression analysis of DECGs was carried out using publicly available transcriptome data deposited in NCBI (<https://www.ncbi.nlm.nih.gov/gds/>, accessed on 10, 5 2023) with series accession ID GSE136885. The reported transcriptome data [46] including embryo at 2 days and coleoptile at 4 days, 8 days, and 14 days under submerged conditions was then used to study the expression patterns of DECGs. Expression levels of DECGs were assessed by transcripts per million (TPM) scores and visualized using a heatmap of log₂ in TBtools [42]. Moreover, the 2000-bp upstream of the translational start site (i.e., ATG) of each potential DECGs was obtained from the phytozome database (<https://phytozome-next.jgi.doe.gov/>, accessed on 12, 5 2023) and uploaded to PlantCARE (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>, accessed on 13, 5 2023) to predict *cis*-acting regulatory elements (CAREs) present in the potential DECGs, and the result visualized using TBtools [42].

Haplotype analysis of candidate genes

In order to explore the haplotype diversity of potential DECGs, haplotypes were evaluated in RiceVarMap software v2.0 (https://ricevarmap.ncpgr.cn/vars_in_gene/, accessed on 1, 6 2023). The variations present in the 2 kb upstream, 1 kb downstream, and coding sequence of the potential DECGs were retrieved and filtered to remove SNP markers that possessed the lowest allele frequencies (MAF < 60%). Then SNP markers in these genes were used to construct the haplotype network and distribution of the rice subpopulation.

Validation of MQTL by flanked marker

The conventional method of QTL validation involves developing near-isogenic or transgenic lines to examine QTL effect. In context, MQTL validation differs from the well-known QTL validation approach due to its inclusive property [47]. In a condensed genomic locus, MQTL identification guarantees the validation of known QTL, and therefore validating flanked markers on genotypes with extreme phenotypes has important breeding implications [27]. The peak RFLP marker linked to

the core MQTL7.1 was selected for validation on a set of six rice genotypes. The experimental genotypes were submerged for 7 days and mean coleoptile length was recorded. The genotypes with short coleoptile length were considered sensitive and those with long coleoptile length were regarded as tolerant to submergence. Pure genomic DNA was extracted from young seedlings using the standard Cetyl trimethyl ammonium bromide (CTAB) protocol [48]. The quality and quantity of genomic DNA were determined before thermal amplification of MQTL flanked marker region. The PCR cycling conditions involved an initial denaturation of 2 min at 94 °C, followed by 37 cycles of denaturation at 94 °C for 30 s, annealing at 60 °C for 30 s, extension at 72 °C for 10 s, and a final extension at 72 °C for 25 min. Amplified fragments were visualized using gel electrophoresis. The detailed information on the primers used for MQTL validation is presented in Supplementary Table 9.

Results

Characteristics of submerged-related traits in rice

A total of 16 independent studies published from 2004 to 2023 involving eight F₂ generation, five RILs, and three backcross populations (BC) were assessed to compile QTL information on submergence-related traits (Table 1). Herein, 120 QTL associated with submergence tolerance during germination were collected. Among

the studied traits, QTL involving anaerobic germination (65.83%) were most represented followed by low-temperature germination (17.5%), submergence (15.0%), and lowest elongated internode (1.67%) (Fig. 1a). The distribution of QTL on rice chromosomes was uneven, with chromosome 7 accounting for the largest number of QTL (16.67%; 20/120) followed by chromosomes 1 and 5 with 15% (15/120) and 11.67% (14/120), respectively (Fig. 1b). However, chromosome 4 contained the least QTL (3.33%; 4/120). The PVE (%) by each QTL from different studies ranged from 0.09 to 59.08% (Fig. 1c). Interestingly, majority of the initial QTL (53.33%; 64/120) had PVE above the 10% threshold. The LOD scores of initial QTL ranged from 1.35 to 18.2 with a mean of 5.03. The distribution of LOD values of the initial QTL is presented in Fig. 1d.

Consensus map of rice

The consensus map developed in this study consisted of 24,644 markers which included different types of markers such as simple sequence repeat, single nucleotide polymorphism, sequence-tagged sites, and restriction fragment length polymorphism. More importantly, genes including *OsSSII*, *OsSSI*, *Amy1*, and *ACT1* were also integrated into the consensus map. The total length of the reference map was 1,457.74 cM with a mean chromosome length of 1,341.63 cM (Supplementary Fig. 1). Moreover, individual chromosomal distance ranged from

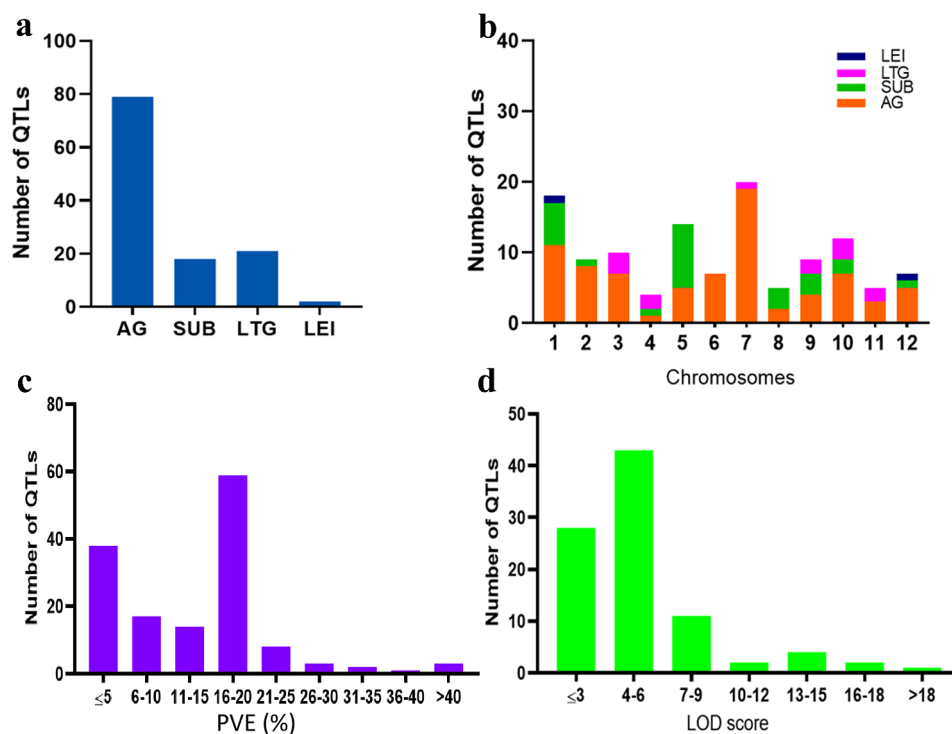


Fig. 1 Salient features of the initial quantitative trait loci (QTL) used in this study. **a** Distribution of 121 initial QTL associated with submergence-related traits; AG (anaerobic germination), SUB (submergence), LTG (low-temperature germination), LEI (lowest elongated internode), **b** Trait-wise distribution of QTL on the 12 chromosomes of rice, **c** Distribution of phenotypic variance explained (PVE) by initial QTL, **d** Logarithm of odds values of the initial QTL

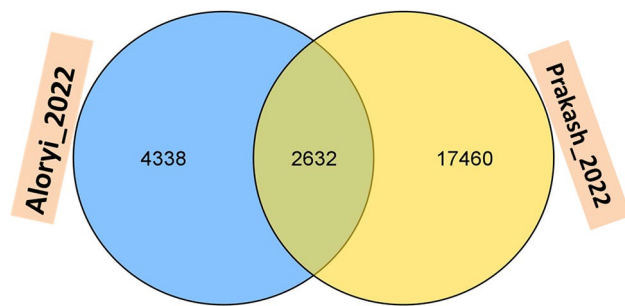


Fig. 2 Venn diagram of unique and common markers among the two maps used to construct the consensus map

45.71 (on chromosome 7) to 201.97 cM (on chromosome 8). Overall, chromosome 1 displayed the highest number of markers (3,048) and the lowest was chromosome 10 with 1354 markers. The two consensus maps adopted in this study shared 10.8% of markers; Aloryi_2022 had 4,338 unique markers (17.8%) while Prakash_2022 had 17,460 unique markers (71.5%) (Fig. 2). Markers on individual linkage groups were sparsely distributed with two ends of the chromosome displaying uneven marker density (Supplementary Fig. 1).

QTL projection and identification of MQTL

For consensus QTL prediction, the LOD score, PVE, and chromosome position were considered. As a consequence, a total of 103 QTL out of 120 QTL were successfully projected on the consensus map (Fig. 3a; Supplementary Table 1). The remaining 17 QTL were not projected due to lack of common markers between initial QTL and the reference map and low PVE scores resulting in large confidence intervals. Moreover, there was a high correlation between initial and projected QTL on rice chromosomes with Pearson correlation coefficient (r) of 0.93 (Fig. 3b). For the projected QTL, 50 QTL were identified as major QTL ($PVE \geq 10\%$), while 53 were minor QTL ($PVE < 10\%$). The QTL meta-analysis confined the initial QTL to 18 MQTL, which contained only 56 of the projected QTL (Fig. 4; Supplementary Table 2). Of the identified MQTL, 50% (9/18) were composed of two or more initial QTL and 50% (9/18) constituted only single initial QTL. The MQTL with relatively high initial QTL harbored QTL from diverse bi-parental populations indicating that they are more reliable and stable for submergence tolerance improvement in rice. Notably, chromosomes 1, 2, 3, 5, 7, 9, and 12 had MQTL with high initial QTL ranging from nine QTL (MQTL7.1) to four QTL (MQTL3.1).

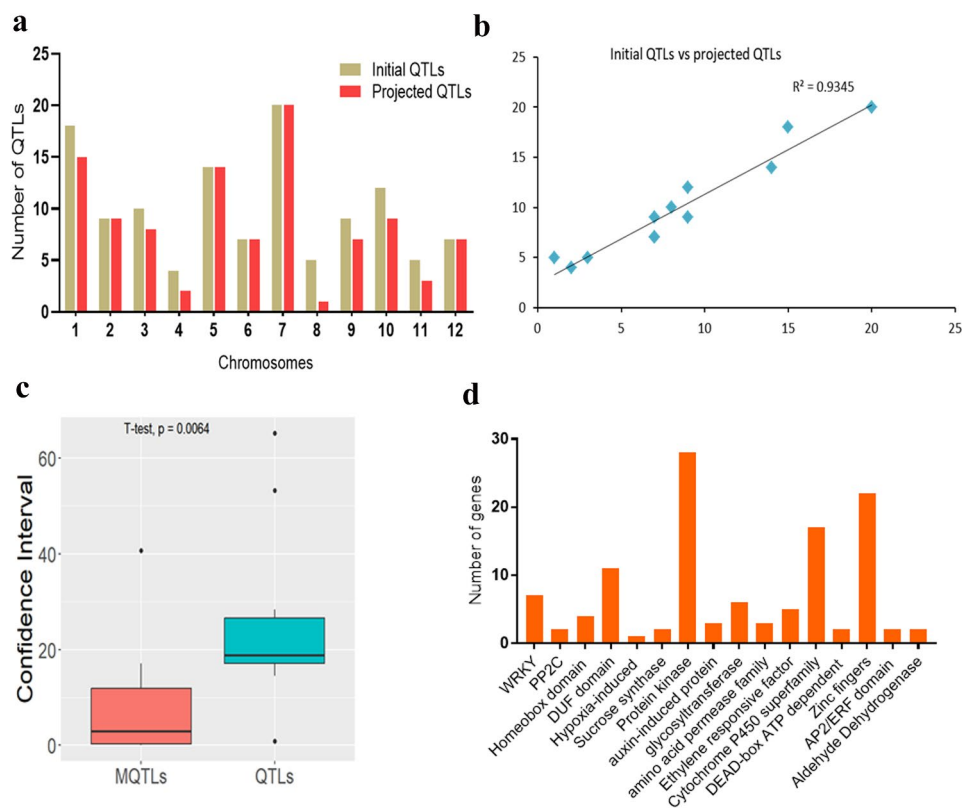


Fig. 3 Basic features of initial QTL and Meta-QTL (MQTL) in this study. **a** Frequency of initial QTL and projected QTL, **b** Correlation between initial QTL and projected QTL, **c** Comparison of mean confidence interval (CI) for initial QTL and MQTL, **d** Distribution of candidate genes encoding known protein families associated with submergence tolerance

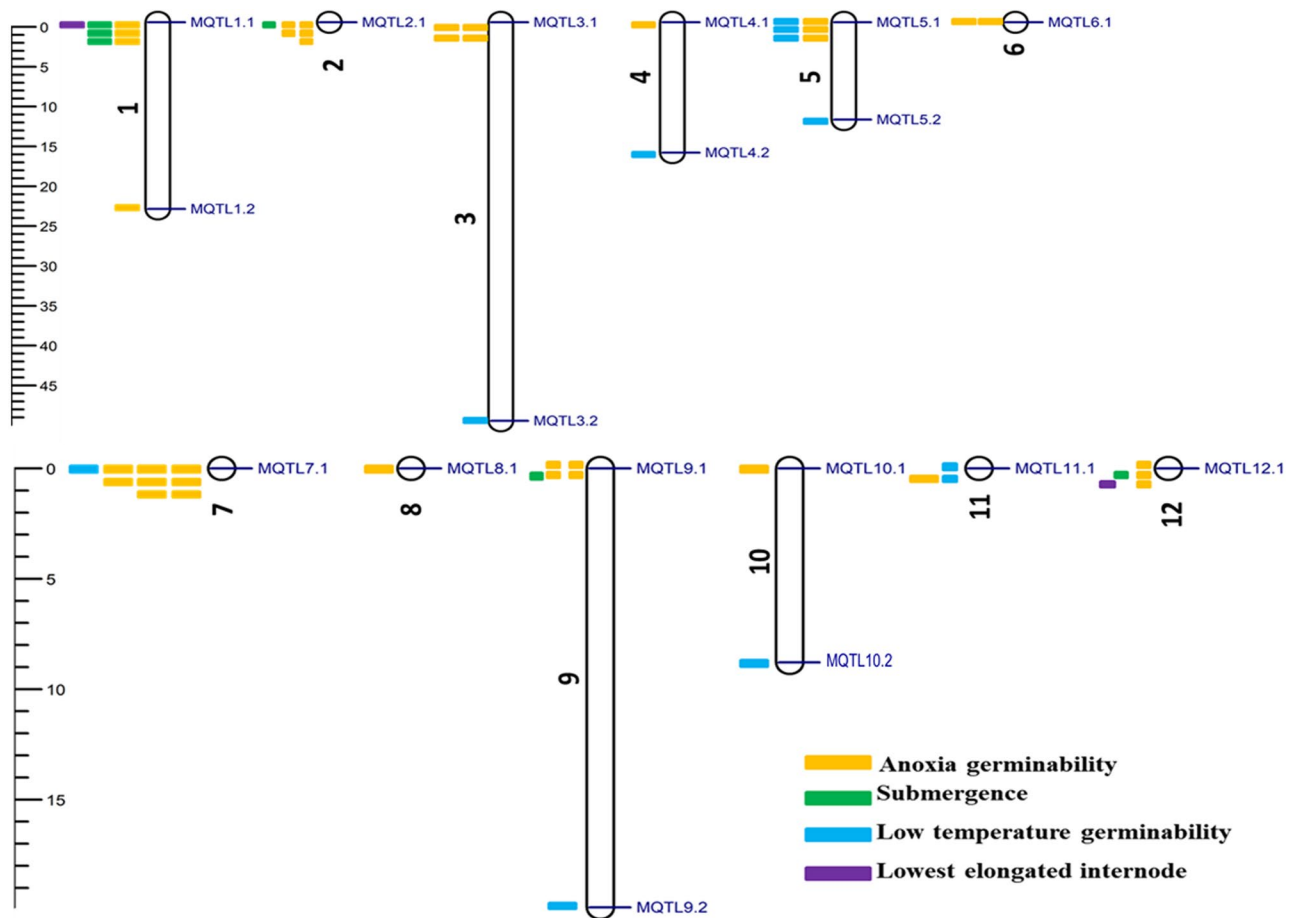


Fig. 4 The chromosome distribution of MQTL for submergence tolerance

Table 2 List of core MQTL identified in this study

MQTL ID ^a	Position (cM)	CI (cM) ^b	Genetic interval (cM)	Initial QTL	PVE (%) ^c	Flanking markers
MQTL2.1	0.25	0.12	0.14–0.26	6	12.18	RM12578 - RM12381
MQTL3.1	0.00	0.02	0.0–0.02	4	18.54	GNMS1289 - RM3265
MQTL7.1	0.00	0.12	0.11–0.23	9	12.01	S2329 - HV7-02
MQTL12.1	0.41	0.29	0.28–0.51	5	11.46	HV12-11 - TEL2A

^aMeta-QTL ^bConfidence interval ^cMean phenotypic variance explained

The CI of MQTL ranged from 0.02 to 40.7 cM with an average of 7.34 cM, while the CI of projected QTL ranged from 0.88 to 65.22 cM with an average of 25.22 cM. There was a significant difference between the mean CI of MQTL and the initial QTL that mapped on the rice chromosomes (Fig. 3c). Also, MQTL exhibited significant fold reduction in CI across rice chromosomes ranging from 0.12 for MQTL on chromosomes 2 and 7 to 20.37 cM for MQTL on chromosome 3. The PVE scores of individual MQTL ranged from 0.15 to 31.23% with a mean score of 11.85% (Supplementary Table 2). Per the criteria established by Löffler et al. [49], MQTL to be considered for marker-assisted breeding should possess a high number of initial QTL, high PVE values of initial QTL, and relatively small CI. The following criteria were then employed

to select the most important MQTL: (i) MQTL involving at least four initial QTL, (ii) mean PVE > 10%, and (iii) CI < 2 cM. Based on this, four MQTL were selected to be the most significant MQTL and thus named as core MQTL. These four MQTL were as follows: MQTL2.1, MQTL3.1, MQTL7.1, and MQTL12.1 (Table 2).

Validating MQTL by GWAS-MTAS

To confirm the reliability of the meta-analysis, MQTL were verified from GWAS studies involving submergence-related traits published in recent years (Supplementary Table 3). Intriguingly, 55.56% (10/18) of MQTL in this study co-localized with 124 SNP signals associated with submergence-related traits (Fig. 5). The number of MTA overlapping the identified MQTL also varied

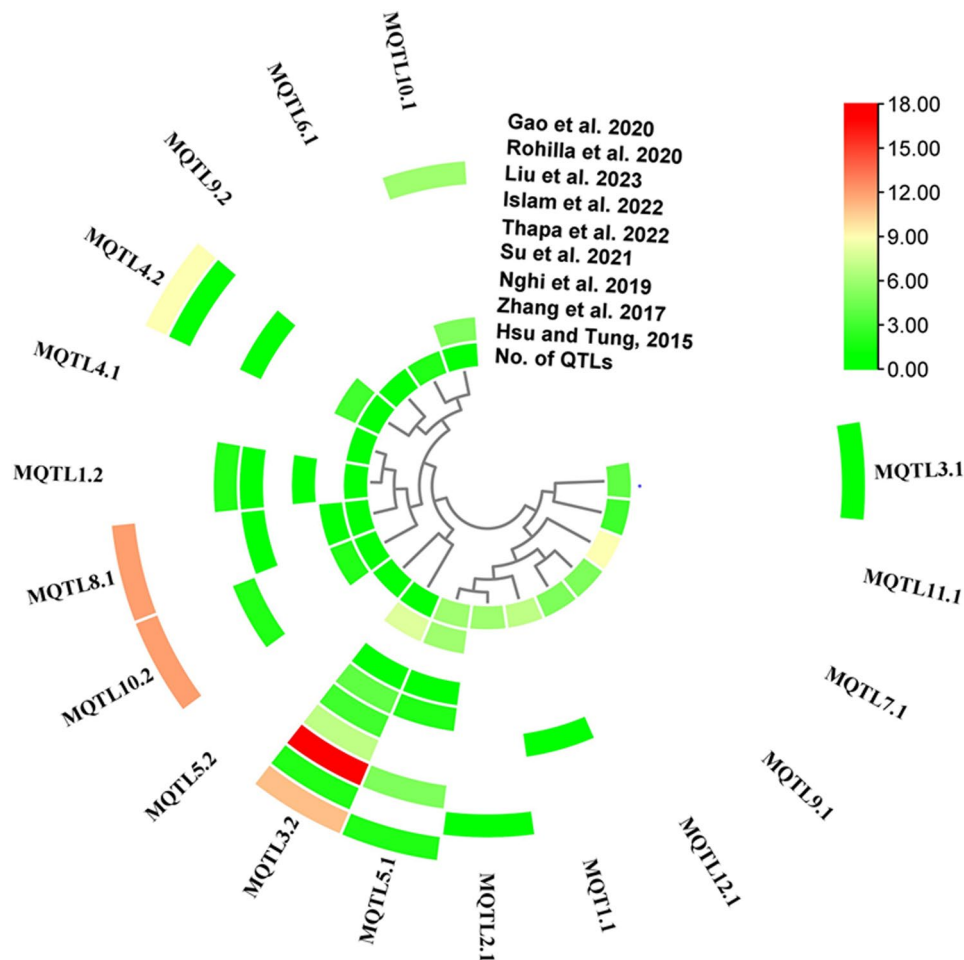


Fig. 5 Validation of MQTL by marker-trait associations (MTAs) on rice submergence-related traits from GWAS with nine diverse natural populations. Color codes showed the number of MTA overlapping the identified MQTL

from one to 17 in nine association studies. Each of the 10 MQTL overlapped at least one MTA. In particular, MQTL3.2, MQTL5.1, MQTLA.2, MQTL8.1, MQTL10.2, and MQTL10.1 overlapped 53, 16, 14, and 11 MTA. In light of the synteny among rice, maize, and sorghum, the corresponding CGs identified in the core MQTL regions were examined on the genomes of sorghum and maize. As expected, the core MQTL loci were conserved among rice relatives indicating that comparative genomics analysis could be useful in identifying candidate genes and transferring information across cereal crops (Fig. 6).

Mining of candidate genes underlying core MQTL regions and exploration of their functional classification

A total of 858 putative CGs were mined in the core MQTL regions; these candidate genes included 28 genes with protein kinase domain, 22 zinc finger family proteins, 17 cytochrome P450 proteins, 11 DUF domain proteins, 7 WRKY, and four glycosyltransferase domain-containing proteins (Fig. 3d; Supplementary Table 4). To validate CGs underlying core MQTL regions,

submergence-responsive genes were obtained from two independent transcriptome studies. Based on DEGs datasets, 5,124 and 25,614 genes were differentially expressed in response to anaerobic germination during flooding. Overall, 38 common DEGs were identified between DEGs studies and genes found in the core MQTL regions, hence they were designated as DECGs (Fig. 7a; Supplementary Table 5). Next, a protein-protein interaction (PPI) network was constructed to predict the relationship among DECGs. The PPI analysis revealed three main hub gene interactions; *LOC_Os02g03410*, *LOC_Os02g08364*, and *LOC_Os12g01140* interacted with five genes involved in calcium flux signaling, reactive oxygen species (ROS) detoxification, and cellular stress signaling (Fig. 7b; Supplementary Table 6).

Ultimately, GO and KEGG pathway enrichment of DECGs were explored to decipher genes' functional classification. As expected, three main GO terms were revealed including biological processes (six subfunctions), molecular functions (three subfunctions), and cellular components (two subfunctions) (Fig. 7c). For

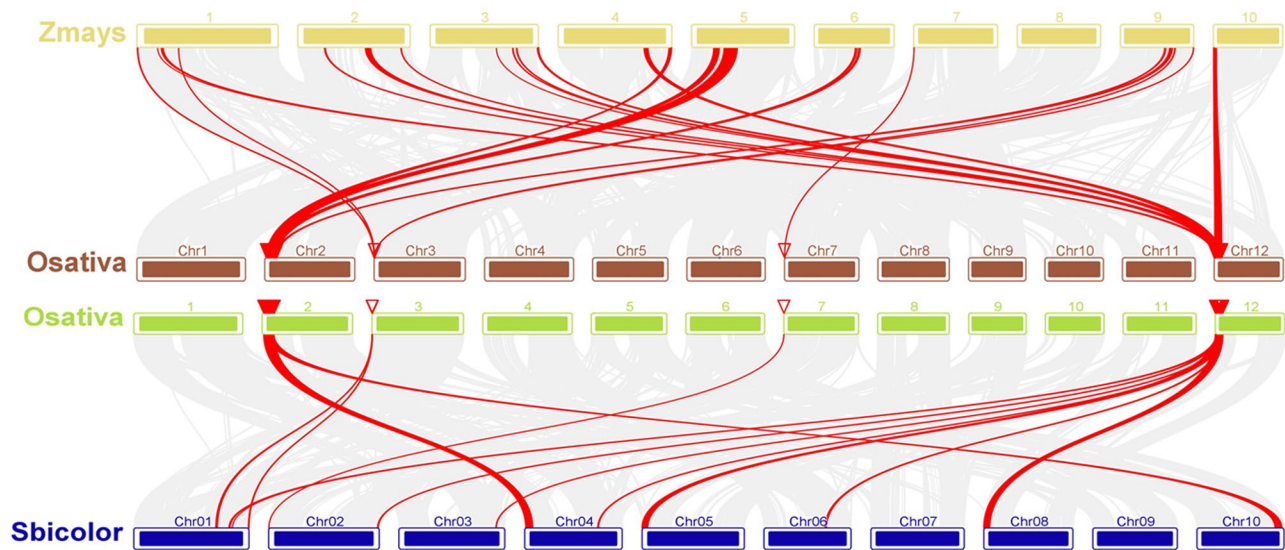


Fig. 6 Syntenic regions among rice, sorghum, and maize genomes. The gray lines in the background depict collinear blocks within pairs of rice and each other species; red lines show syntenic gene pairs

biological processes, the most enriched GO terms involved response to abiotic stimulus, lipid transport, metabolic process, and stress response. For molecular functions, the DECGs were mainly enriched in lipid binding, catalytic activity, and oxidoreductase activity, while the main enriched terms for cellular components involved cell wall and extracellular region. The KEGG pathway enrichment analysis showed several mechanisms involved in submergence tolerance, among them include glyoxylate and dicarboxylate metabolism, biosynthesis of secondary metabolite, and carbon metabolism pathways (Fig. 7d; Supplementary Table 7).

***In silico* expression analysis of DECGs and cis-acting regulatory elements in the promoters of submergence-responsive genes**

To pinpoint potential DECGs involved in submergence regulation, *in silico* expression analysis was performed. The expression patterns of these genes were assessed during developmental stages under submerged conditions. Of the 38 DECGs evaluated, only 21 genes were highly upregulated in embryo and coleoptile at 2, 4, 8, and 14 days, in response to flooding/submergence (Fig. 8). Based on their expression profiles, these putative genes could be divided into three groups (Fig. 8). The group I genes were highly expressed across all four stages (2 days in embryo, while others were expressed in coleoptile at 4, 8, and 14 days). In group II, genes were specifically highly expressed in coleoptile at 8 and 14 days, while group III genes were highly upregulated in embryo and coleoptile at 2, 4, and 8 days submergence. Therefore, we hypothesized that these genes could be promising candidate genes modulating rice submergence tolerance.

Cis-acting regulatory elements (CAREs) in gene promoters play crucial roles in gene expression regulation [50], therefore it is imperative to study CAREs located in the promoter regions of the putative DECGs identified in this study. Within the 2 kb promoter regions of the 21 potential DECGs, a total of 977 CAREs were detected; these CAREs were further divided into three groups, thus stress-responsive elements, hormone-responsive elements, and development-related elements (Fig. 9; Supplementary Table 8). Among the stress-responsive elements include those involved in anaerobic induction (ARE), stress (As-1, TC-rich repeats, STRE), and low temperature (LTR). Many CAREs, such as abscisic acid (ABRE), ethylene (ERE), Me-JA (TGACG-motif), salicylic acid (TCA-motif), gibberellin (TATC-box, GARE-motif) and auxin (TGA) are among phytohormone responsiveness elements. Interestingly ABRE was the most abundant element located in the promoter regions of 21 DECGs. Also, shoot and root meristem expression (CAT-box), light responsiveness (G-box, TCT-motif, etc.), and zein metabolism-mediated *cis*-elements were all involved in the development processes. These results suggest that a complex regulatory network modulates the transcription of genes involved in submergence tolerance.

Haplotypes of putative DECGs

Rice genes associated with abiotic stress tolerance have been reported to exhibit a specific distribution of mutation between *indica* and *japonica* groups [51, 52]. Based on this, we investigated the haplotype distribution in the 21 putative DECGs. The results revealed that the 21 DECGs contributed a total of 188 haplotypes (Fig. 10; Supplementary Fig. 2). It was found that the haplotypes in individual genes differentiated significantly in *indica*,

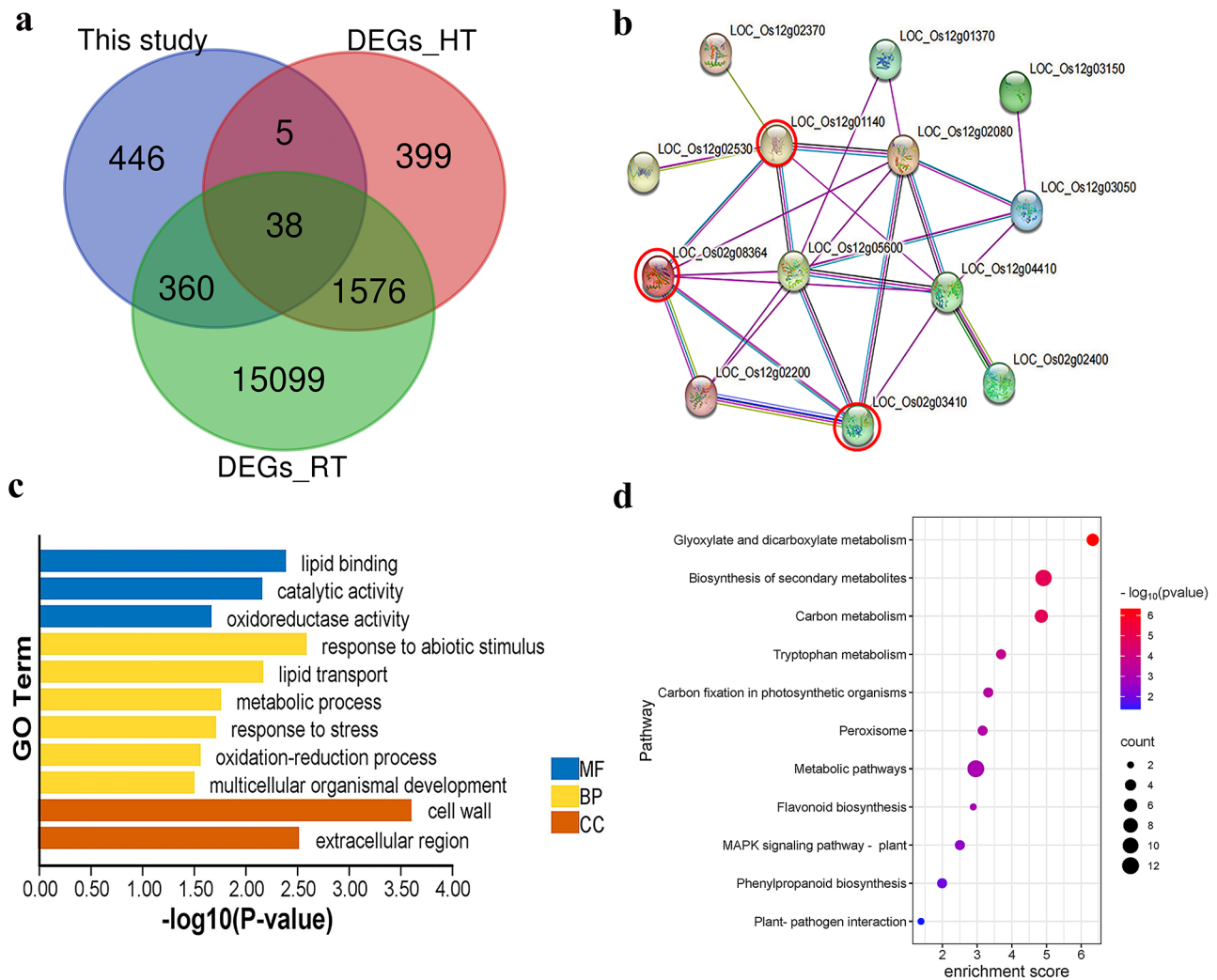


Fig. 7 Analysis of transcriptomic datasets and genes in the core MQTL regions. **a** Venn diagram showing the number of differentially expressed genes (DEGs) between transcriptome data and genes in the core MQTL loci, **b** Protein-protein interaction among the common DEGs, **c** GO terms for DEGs in core MQTL regions, **d** KEGG enrichment pathways for DEGs in core MQTL regions

japonica, *aus*, and intermediate accessions with the majority of haplotypes distributed in *indica* rice. Haplotypes in these genes ranged from 20 in *LOC_Os12g02370* to three in *LOC_Os03g01320*. The candidate genes in this study showed a great deal of genetic differentiation and genetic distance mainly between *indica* and *japonica* accessions.

Validation of meta-QTL

Based on the consensus map, one of the peak markers flanking the core MQTL was selected for MQTL validation. Marker S2329 from MQTL7.1 was used for MQTL validation. Interestingly, S2329 could differentiate the genotypes into two phenotypic groups, i.e., short coleoptile length (sensitive lines) and long coleoptile length (tolerant lines) when genotypes when submerged. The mean coleoptile length of sensitive lines ranged from 0.27 to 0.7 cm while the mean coleoptile length of tolerant lines

ranged from 2.42 to 2.70 cm (Fig. 11a and b). The allele with an amplicon size of 500 bp was found specifically related to tolerant genotypes (long coleoptile length) while an amplicon size of 250 bp was related to sensitive genotypes (short coleoptile length) (Fig. 11a; Supplementary Fig. 3).

Discussion

Due to rising costs, labour shortages, and water shortages, farmers are now practicing direct seeding in the puddled field in rainwater and irrigation systems across the globe [53, 54]. In such a case, the ability of rice to survive under submergence/flooding conditions during germination is critical. Although rice seeds can germinate under hypoxic or even anaerobic conditions and extend the coleoptile, anaerobic conditions are extremely detrimental to seedling growth and establishment at the early seedling stage [11, 22, 55]. Numerous

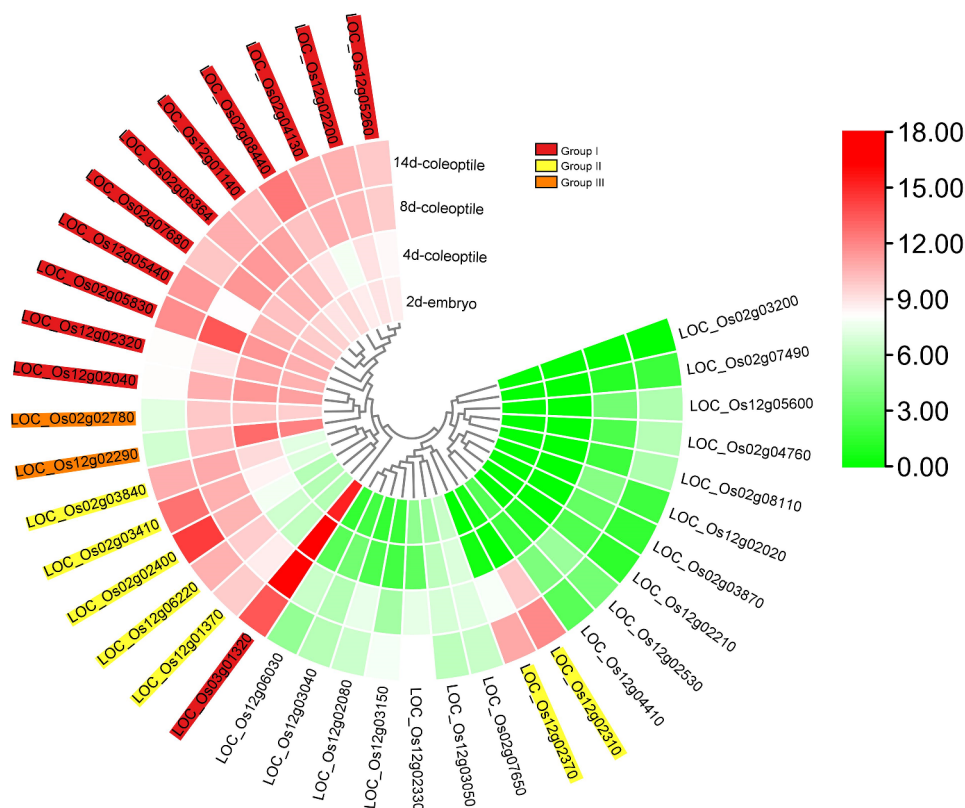


Fig. 8 Expression profiles of 38 differentially expressed candidate genes (DECGs) in two different tissues under submergence conditions. Transcriptome data were downloaded from NCBI

submergence-related QTL have been identified in the past few decades thanks to the discovery of molecular markers and breakthroughs in QTL mapping strategies [20, 21]. However, the use of these QTL in breeding programs faces major bottlenecks due, in part, to low PVE values and large confidence intervals.

In the present study, we conducted a meta-analysis of 120 previously reported QTL underlying submergence-related traits from 16 independent studies and discovered that these 120 QTL were unevenly spread across the rice chromosomes, with a large number of QTL located on chromosomes 7, 1, and 5 and the least QTL on chromosomes 4, 8, and 11 (Fig. 1) which agreed with our previous study [39] and studies conducted by Wu et al. [56] and Khahani et al. [37]. One of the interesting findings of this study is that the PVE by most of the initial QTL was low indicating that minor effect loci are at play in influencing rice submergence tolerance. The CI of 18 MQTL identified in this study was 3.44 times lower than that of the initial QTL studies indicating a possible close linkage of associated markers to the MQTL, thus enabling flexibility of introgressing MQTL in marker-assisted breeding. This finding is consistent with our previous report [39] and other studies in rice and wheat [30, 57, 58]. However, for individual MQTL, we found that MQTL3.2 had a large confidence interval (40.7 cM) although did

not exceed the CI of the initial QTL (Supplementary Table 2). This phenomenon could be ascribed to low initial QTL involvement and small CI associated with the reported QTL [59]. Moreover, Saini et al. [34] observed a similar event occurring in their meta-analysis study involving wheat and suggested a possible occurrence of low recombination frequency in the MQTL locus. Furthermore, the distribution of MQTL was uneven across the rice chromosomes with chromosomes 1, 3, 4, 5, 9, and 10 constituting the highest number of MQTL than other chromosomes with a single MQTL.

To confirm the reliability of our MQTL, we validated MQTL results with markers reported in GWAS and found that more than 50% of the 18 MQTL overlapped with loci from those studies, which corroborated with the previous report [60–62]. The MQTL colocalized with loci at least twice across independent GWAS, of which MQTL3.2 and MQTL5.1 were detected in eight and five GWAS reports, respectively (Fig. 5). This outcome indicated the reliability of our MQTL. In addition to marker-assisted selection (MAS), MQTL overlapping markers from GWAS can also be used to clone genes involved in MQTL with a higher degree of confidence.

A set of four MQTL were designated as core MQTL. Each of these MQTL was characterized by relatively small CI (CI < 2 cM), high PVE (> 10%), and involvement of at

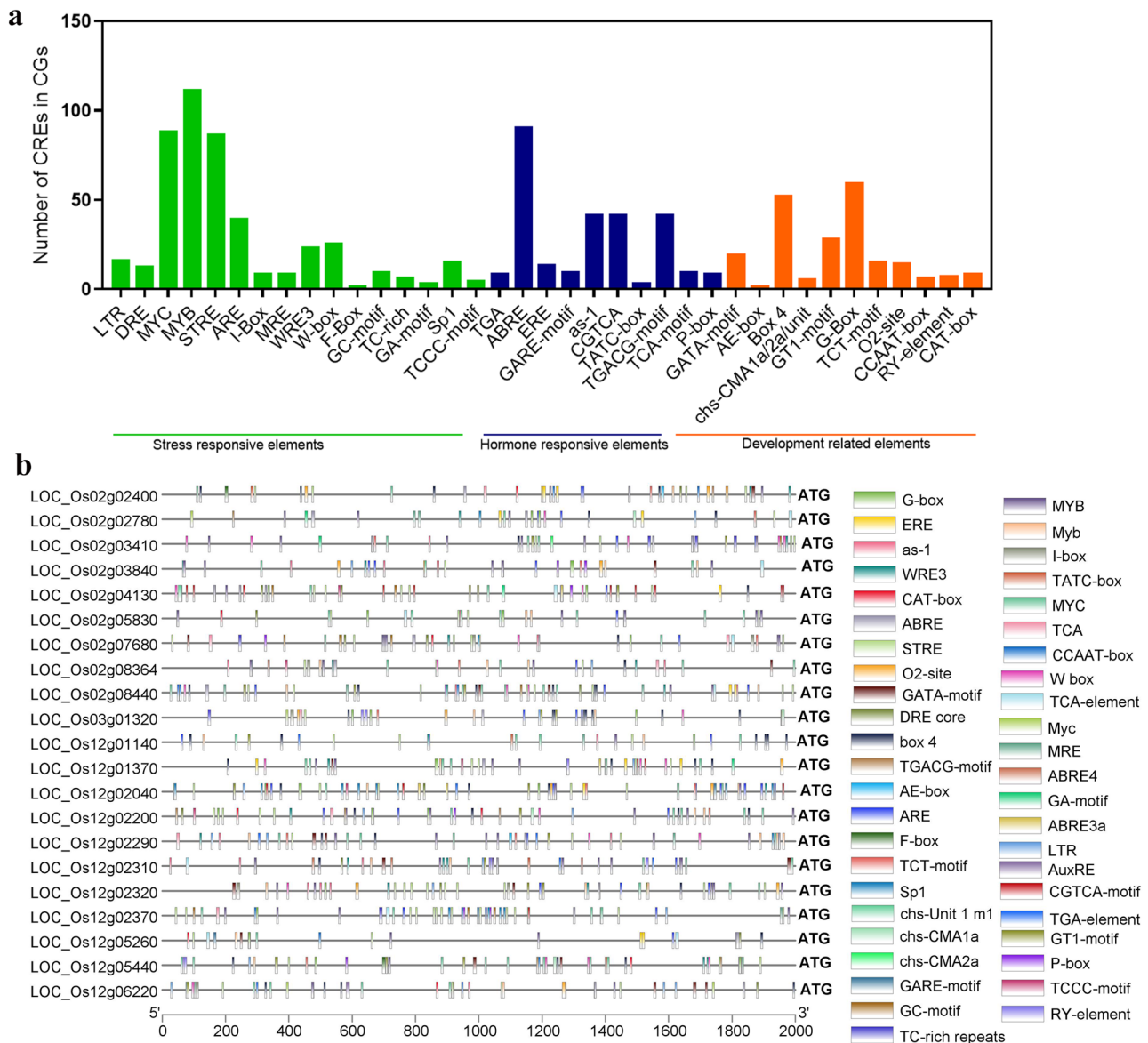


Fig. 9 Analysis of cis-regulatory elements (CREs) in promoter regions of 21 potential DECGs identified in this study

least four initial QTL. Interestingly, all four core MQTL contained initial QTL from diverse genetic backgrounds and different environments [32, 63]. By synteny analysis, we showed that these core MQTL were conserved between rice, maize, and sorghum genomes (Fig. 6) [64]. These syntenic regions constituted numerous unexploited genes in rice which can be targeted for functional validation in future research. The synteny analysis also provides the opportunity to comprehend the molecular and evolutionary basis of submergence tolerance across cereal crops. Moreover, gene-based functional marker development in these regions can be feasible in breeding for submergence tolerance in cereals.

The candidate genes present in MQTL are prioritized for allele mining [65]. In the selected core MQTL

regions, we mined 858 gene models including 173 genes with unknown functions. Some of the well-known proteins encoded genes associated with submergence stress included aldehyde dehydrogenase, F-box domain proteins, WRKYs, cytochrome P450, glycosyltransferase, hypoxia-induced protein, PP2C, homeobox domain protein, sucrose synthase, protein kinase, zinc fingers, amino acid permease family, ethylene-responsive factors/AP2 domain, and DEAD-box ATP dependent proteins (Fig. 3). In previous studies, aldehyde dehydrogenase is involved in rice submergence tolerance and related abiotic stresses [66–68]. The expression of aldehyde gene (*Aldh2a*) in rice was induced in seedlings under flooding [66]. In a recent study, a gene encoding glucosyltransferase enhanced submergence tolerance in rice during

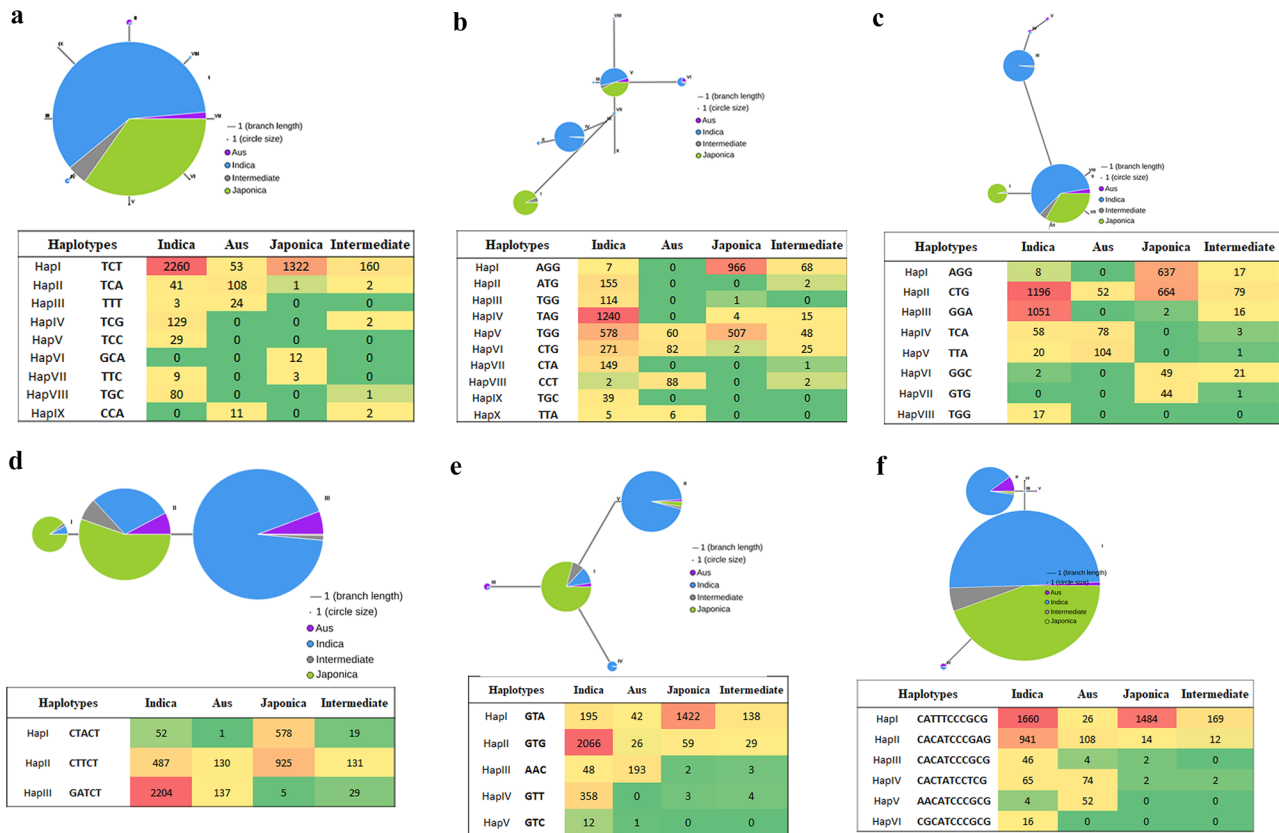


Fig. 10 Haplotype analysis of five potential DECGs. Circle size corresponds to the number of samples for a given haplotype. Lines between haplotypes indicate mutational steps between alleles. The tables give a summary of the number of haplotypes present in individual genes and their distribution in rice subpopulations. Different colors show the number and or the density of accessions corresponding to subpopulations

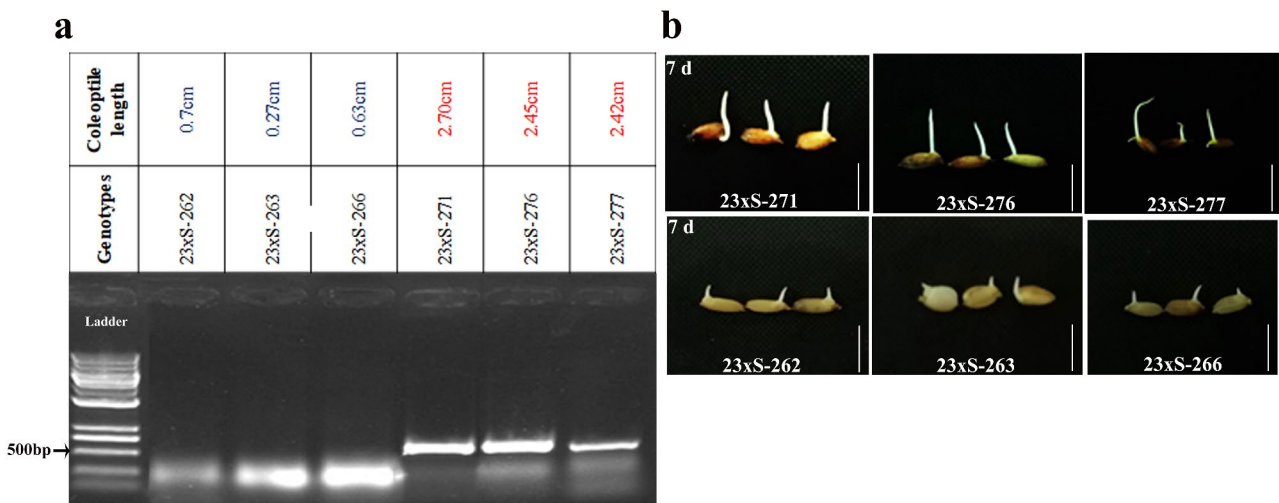


Fig. 11 Validation of marker S2329. **a** Validation of marker S2329, peak marker of MQL7.1 on a set of three long coleoptile length and three short coleoptile length genotypes under submergence condition. The allele with 250 bp related to short coleoptile length and 500 bp allele associated with long coleoptile length, values represent the mean of coleoptile length of three replicates, **b** Representative images of genotypes subjected to submergence stress for 7 days. Scale bars denote 10 mm

germination [8]. *OsUGT75A* promoted coleoptile elongation by limiting abscisic acid and jasmonic acid levels. A well-known *SUB1A* gene confers tolerance to rice submergence; *SUB1A* encoding ethylene response factor-like protein binds to the promoter of MAPK3 (Mitogen-activated protein kinase3) and controls its expression during submergence stress signaling [69]. In rice coleoptiles, a considerable number of ERFs including *ERF60*, *ERF67*, and *ERF68* genes have been implicated under anaerobic conditions and also belong to the ERF subgroup VII [70].

Table 3 List of potential candidate genes identified in this study

MQTL ID	Candidate genes	Functional description	Gene symbol
MQTL2.1	<i>LOC_Os02g02400</i>	catalase isozyme A	<i>OsCatA</i>
	<i>LOC_Os02g02780</i>	protein kinase family protein	<i>ACTPK1</i>
	<i>LOC_Os02g03840</i>	expressed protein	
	<i>LOC_Os02g04130</i>	DUF1645 domain-containing protein	<i>OsSGL</i>
	<i>LOC_Os02g05830</i>	ribulose biphosphate carboxylase small chain, chloroplast precursor	<i>OsRBCS1</i>
	<i>LOC_Os02g07680</i>	cytochrome P450	
	<i>LOC_Os02g08364</i>	protein phosphatase 2 C	
	<i>LOC_Os02g08440</i>	WRKY71, expressed	<i>Os-WRKY71</i>
	<i>LOC_Os02g03410</i>	calcium/calmodulin-dependent protein kinases	<i>OsCPK4</i>
	MQTL3.1	<i>LOC_Os03g01320</i>	LTPL116 - Protease inhibitor/seed storage
MQTL12.1	<i>LOC_Os12g01370</i>	fatty acid desaturase, putative, expressed	<i>OsFAD3</i>
	<i>LOC_Os12g02040</i>	hypoxia-responsive family protein	
	<i>LOC_Os12g02200</i>	calcium/calmodulin dependent protein kinases	<i>OsCIPK14</i>
	<i>LOC_Os12g02290</i>	LTPL23 - Protease inhibitor/seed storage	
	<i>LOC_Os12g02310</i>	LTPL11 - Protease inhibitor/seed storage	
	<i>LOC_Os12g02320</i>	LTPL12 - Protease inhibitor/seed storage	
	<i>LOC_Os12g02370</i>	chalcone-flavonone isomerase	<i>OsCHIL2</i>
	<i>LOC_Os12g05260</i>	phytosulfokines precursor	
	<i>LOC_Os12g05440</i>	cytochrome P450	<i>CYP94C2b</i>
	<i>LOC_Os12g06220</i>	harpin-induced protein 1 domain-containing protein	
	<i>LOC_Os12g01140</i>	ACG kinases include homologs to PKA, PKG and PKC	

Overall, these findings substantiate that the core MQTL regions embodied potential candidate genes involved in response to submergence and can be utilized for map-based cloning of submergence tolerance genes.

To enable robust identification of candidate genes associated with submergence, we validated genes located in the core MQTL regions by comparing them with two publicly available transcriptome datasets and found that only 38 genes (DECGs) overlapped. PPI is a useful method for understanding how genes function at the cellular level [71–73]. In the present study, we found three main hub-gene interactions among DECGs. These genes (*LOC_Os02g03410*, *LOC_Os02g08364*, and *LOC_Os12g01140*) interacted with genes involved in ROS detoxification, calcium flux signaling, and cellular stress signaling (Fig. 7), indicating that complex regulatory network of genes influenced rice seed germination under submerged conditions [74, 75]. The analysis of GO and KEGG pathway enrichment of DECGs in core MQTL regions is vital for understanding mechanisms underlying submergence tolerance in rice. Most of the genes in the core MQTL regions were significantly enriched in cell wall, abiotic stimulus, extracellular region, lipid binding, lipid transport, and catalytic activity [75]. Interestingly, we found that genes in these regions were enriched in pathways including carbon metabolism, MAPK signaling pathway-plant, tryptophan metabolism, and isoflavonoid biosynthesis pathway which have been reported to play important roles in flooding tolerance [71]. Also, pathways such as glyoxylate and dicarboxylate metabolism, biosynthesis of secondary metabolites, metabolic pathways, carbon fixation in photosynthetic organisms, and phenylpropanoid biosynthesis might be involved in submergence tolerance. To pinpoint potential DECGs, we used expression data generated under rice submergence conditions during germination deposited in NCBI database. Accordingly, we found a total of 21 genes with high and stage-specific expression profiles in embryo at two days, and coleoptile at four, eight, and 14 days (Fig. 8). These DECGs exhibited tissue and development-dependent expression profiles and are considered potential CGs directly involved in submergence tolerance (Table 3). For instance, *LOC_Os02g02780* encoding protein kinase was significantly upregulated in embryo and coleoptile. *LOC_Os02g07680* and *LOC_Os12g05440*, encoding cytochrome P450 were highly expressed in embryo and coleoptile at all stages of development under anaerobic germination. *LOC_Os12g02040* encoding expressed protein was highly induced in embryo and during coleoptile elongation under anaerobic conditions (Fig. 8). A total of 11 reported loci were found in MQTL2.1, MQTL3.1, and MQTL12.1 regions (Table 3). *OsCatA* encodes catalase isozyme A enhanced drought tolerance by modulating ROS homeostasis associated with root growth

and photorespiration [76]. The orthologue of *OsCatA* in maize *Zm00001d014848* plays an essential role in ROS scavenging by converting H_2O_2 into H_2O and O_2 [77, 78]. The locus *OsSGL* and its orthologue *ZmDUF1645* in maize increased grain yield and stress tolerance by regulating the expression of genes involved in the cytokinin signaling pathway [79–81]. The gene *OsWRKY71* (*LOC_Os02g08440*) was reported to play a positive role in cold [82] while its orthologue *ZmWRKY40* (*Zm00001d015515*) in maize was specifically up-regulated in hydrotropic maize root [83]. Similarly, salt stress and drought stress responses in rice are positively regulated by locus *OsCPK4* (*LOC_Os02g03410*) through the protection of cellular membranes from oxidative damage [84]. Although none of the reported genes within MQTL2.1 were found directly involved in submergence stress tolerance, their roles in biological processes associated with abiotic stress response and yield proved significant, hence warranting further studies using submergence stress as a case study. In the MQTL3.1 region, *qLTG3-1* (*LOC_Os03g01320*) has been proven to modulate rice seed germination under low-temperature conditions [85]. Moreover, the findings of previous studies indicated that the clone genes found in MQTL2.1 played diverse roles in plant stress response [86]. For instance, a high expression of *CYP94C2b* (*LOC_Os12g05440*) encodes cytochrome P450 inactivated jasmonic acid biosynthesis thereby enhancing salt tolerance in rice [86]. The identified candidate genes in these MQTL regions provide clues for their possible roles in submergence tolerance with similar evolutionary trajectories and conserved functions between cereal crops. Breeders can use these results to track down candidate genes and employ marker-assisted selection in breeding programs of cereals under submergence stress.

CAREs are important DNA sequences located in the promoter region of genes that allow the binding of transcription factors leading to either enhancing or silencing gene expression when unfavorable conditions prevail [48]. In our study, all 21 submergence-associated genes (DECGs) contained three distinct functional categories of cis-elements such as stress-responsiveness, phytohormone responsiveness, and development-related elements (Fig. 9). For instance, ARE and GC-motif located in the promoter region of *OsWRKY62* upregulated transcription in response to oxygen deficiency. W-box, a stress-responsive element was found in the promoter of *OsWRKY11*, *OsWRKY56*, and *OsWRKY62* and upregulated their expression under submergence stress [87, 88]. Surprisingly, phytohormone-responsive elements ABRE occurred more frequently in 21 DECGs which echoed its importance in submergence regulation (Fig. 9). Stress-responsive cis-elements DRE as well

as phytohormone-responsive elements including ABRE have all been implicated in response to submergence [89].

Haplotype strategy which utilizes genome sequence information and phenotype data has proven useful in excavating natural allelic variants present in different genotypes to develop tailored-made superior cultivars [90, 91]. By analyzing the genetic diversity of 21 DECGs, we found that the haplotypes of these genes significantly differentiated mainly in *indica* and *japonica* subspecies [31]. The differences in tolerance levels between *indica* and *japonica* rice are probably due to these haplotypes, which have important implications for breeding rice varieties suitable for direct seeding under flooding conditions. However, this requires further validation by exploring genetic variants in ideal rice cultivars to substantiate gene function.

Among the core MQTL identified in this study, we selected MQTL7.1 for validation on extreme genotypes including three sensitive lines (short coleoptile length) and three tolerant lines (long coleoptile length). Consistently, marker S2329 associated with MQTL7.1 successfully differentiated two extreme genotypes (Fig. 11). Segregation of marker amplicons among the extreme genotypes validates their close linkage with MQTL7.1. S2329 was identified as the peak marker of MQTL7.1, suggesting its close linkage with MQTL7.1. It is worth noting that the validated marker has the potential to be used in marker-assisted selection by integrating MQTL7.1 region into an elite background to improve submergence tolerance in rice.

Concluding remarks

In the present study, we integrated QTL conferring submergence tolerance in rice leading to the identification of 18 MQTL. More than 50% of the identified MQTL were validated by GWAS signals. Four MQTL were designated as core MQTL because they contained a high number of initial QTL and are potential regions for excavating submergence-associated candidate genes. Integrated DEGs from public transcriptome datasets and genes in the core MQTL regions revealed 38 genes as DECGs. Further expression analysis of 38 DECGs showed significant upregulation of 21 DECGs in various rice organs under submergence conditions. Some putative genes encoding cytochrome P450, WRKY, hypoxia-responsive protein, calcium/calmodulin-dependent protein kinases, zinc finger protein, and protein phosphatase 2 C were located in the core MQTL regions. The 21 DECGs showed a great deal of genetic differentiation and genetic distance mainly between *indica* and *japonica* accessions. The meta-QTL and candidate genes identified in this study provide valuable information for MAS and future functional validation experiments by CRISPR/Cas9 and other techniques aimed at breeding of rice varieties suitable

for direct seeding and can help decipher the molecular mechanisms underlying submergence tolerance during seed germination.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10219-z>.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

Not applicable.

Author contributions

XT, KDA, and NEO conceptualized and design the study. KDA, NEO, GH, AA, BK, DKS, SFB, and SA collected literature and tabulated QTL data for meta-analysis. KDA, NEO, and GH performed QTL projection and meta-QTL analysis. KDA wrote the manuscript. BK, SFB, AA, DKS, and XT revised and edited the manuscript with the help of KDA and NEO. All authors have read and approved the final version of the manuscript.

Funding

This study was financially supported by Hubei Provincial Natural Science Foundation of China (grant No: 2023AFA022), the Key R&D Project in Hubei Province, China (grant No: 2020BBB060), and the National Key Research and Development Program of China (grant No: 2018YFD0301306).

Data availability

All data supporting the conclusions of this article are provided within the article and its supplementary.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Hubei Collaborative Innovation Centre for Grain Industry, College of Agriculture, Yangtze University, Jingzhou, China

²University of Chinese Academy of Sciences, 100049 Beijing, China

³Département de phytologie, Université Laval, Québec, QC, Canada

⁴Department of Agricultural Biotechnology, Faculty of Agriculture, Food and Consumer Sciences, University for Development Studies, Tamale, Ghana

⁵Department of Horticultural Sciences, Texas A&M University, College Station, TX, USA

⁶Texas A&M Agrilife Research and Extension Center, Weslaco, TX, USA

⁷Department of Animal Genetics, Breeding and Reproduction, College of Animal Science, South China Agricultural University, Guangzhou, Guangdong, China

⁸Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, Punjab, India

⁹Department of Plant and Soil Science, Texas Tech University, Lubbock, TX, USA

¹⁰School of Agriculture, University of Cape Coast, Cape Coast, Ghana

Received: 13 November 2023 / Accepted: 13 March 2024

Published online: 04 April 2024

References

- Zhang P, Zhong KZ, Zhong ZZ, Tong HH. Genome-wide association study of important agronomic traits within a core collection of rice (*Oryza sativa* L). *BMC Plant Biol.* 2019;19:259.
- Jeon JS, Jung KH, Kim HB, Suh JP, Khush GS. Genetic and molecular insights into the enhancement of rice yield potential. *J Plant Biol.* 2011;54:1–9.
- Su L, Yang J, Li D, Peng Z, Xia A, Yang M, Luo L, Huang C, Wang J, Wang H, Chen Z, Guo T. Dynamic genome-wide association analysis and identification of candidate genes involved in anaerobic germination tolerance in rice. *Rice.* 2021;14:1.
- Mahender A, Anandan A, Pradhan SK. Early seedling vigour, an imperative trait for direct-seeded rice: an overview on physio-morphological parameters and molecular markers. *Planta.* 2015;241:1027–50.
- Saud S, Wang D, Fahad S, Alharby HF, Bamagoos AA, Mjrashi A, Alabdallah NM, AlZahrani SS, AbdElgawad H, Adnan M, Sayyed RZ, Ali S, Hassan S. Comprehensive impacts of climate change on rice production and adaptive strategies in China. *Front Microbiol.* 2022;13:926059.
- Sasaki T. Studies on breeding for the germinability at low temperature of rice varieties adapted to direct sowing cultivation in flooded paddy field in cool region. *Rep Hokkaido Pref Agric Exp Sta.* 1974;24:1–90.
- Ismail AM, Ella ES, Vergara GV, Mackill DJ. Mechanisms associated with tolerance to flooding during germination and early seedling growth in rice (*Oryza sativa*). *Ann Bot.* 2009;103:197–209.
- He Y, Sun S, Zhao J, Huang Z, Peng L, Huang C, Tang Z, Huang Q, Wang Z. UDP-glucosyltransferase *OsUGT75A* promotes submergence tolerance during rice seed germination. *Nat Commun.* 2023;14:2296.
- Islam MR, Naveed SA, Zhang Y, Li Z, Zhao X, Fiaz S, Zhang F, Wu Z, Hu Z, Fu B, Shi Y, Shah SM, Xu J, Wang W. Identification of candidate genes for salinity and anaerobic tolerance at the germination stage in rice by genome-wide association analyses. *Front Genet.* 2022;13:822516.
- Ismail AM, Johnson DE, Ella ES, Vergara GV, Baltazar AM. (2012) Adaptation to flooding during emergence and seedling growth in rice and weeds, and implications for crop establishment. *AoB Plants* 2012:pls019.
- Baltazar MD, Ignacio JCI, Thomson MJ, Ismail AM, Mendiolo MS, Septiningsih EM. QTL mapping for tolerance to anaerobic germination in rice from IR64 and the aus landrace Kharsu 80A. *Breed Sci.* 2019;69:227–33.
- Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ. *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature.* 2006;442:705–8.
- Bailey-Serres J, Voeselek LA. Flooding stress: Acclimations and genetic diversity. *Annu Rev Plant Biol.* 2008;59:313–39.
- Voeselek LA, Bailey-Serres J. Flood adaptive traits and processes: an overview. *New Phytol.* 2015;206:57–73.
- Zhan J, Lu X, Liu H, Zhao Q, Ye G. Mesocotyl elongation, an essential trait for dry-seeded rice (*Oryza sativa* L.): a review of physiological and genetic basis. *Planta.* 2019;251:27.
- Pucciariello C. Molecular mechanisms supporting rice germination and coleoptile elongation under low oxygen. *Plants.* 2020;9:1037.
- Angaji SA, Septiningsih EM, Mackill DJ, Ismail AM. QTL associated with tolerance of flooding during germination in rice (*Oryza sativa* L). *Euphytica.* 2009;172:159–68.
- Gonzaga ZJC, Carandang J, Sanchez DL, Mackill DJ, Septiningsih EM. Mapping additional QTL from FR13A to increase submergence tolerance in rice beyond *SUB1*. *Euphytica.* 2016;209:627–36.
- Ghosal S, Quilloy FA, Casal C, Septiningsih EM, Mendiolo MS, Dixit S. Trait-based mapping to identify the genetic factors underlying anaerobic germination of rice: phenotyping, GXE, and QTL mapping. *BMC Genet.* 2020;21:6.
- Saimohan B, Chamundeswari N, Haritha T, Veronica N. (2021) Mapping of QTL for anaerobic germination using the donor Ac39416a in the genetic background of Swarna Sub-1 (*Oryza sativa* L). *Research Square.*
- Rani GM, Satyanarayana PV, Chamundeswari N, Ramana Rao PV, Prabhakar M, Ravikumar BNVS, Nagakumari P, Kalpana K. Mapping of QTL for flood tolerance in rice using recombinant inbred lines of Indra and a new plant genetic resource AC 39416 A. *Plant Genet Res.* 2023;20(4):270–6.
- Angaji SA, Septiningsih EM, Mackill DJ, Ismail AM. QTL associated with tolerance of flooding during germination in rice (*Oryza sativa* L). *Euphytica.* 2010;172:159–68.
- Kretschmar T, Pelayo MAF, Trijatmiko KR, Gabunada LFM, Alam R, Jimenez R, Mendiolo MS, Loedin IHS, Sreenivasulu N, Serres JB, Ismail AM, Mackill DJ, Septiningsih EM. A trehalose-6-phosphate phosphatase enhances anaerobic germination tolerance in rice. *Nat Plants.* 2015;1:15124.

24. Yang J, Sun K, Li DX, Luo LX, Liu YZ, Huang M, Yang GL, Liu H, Wang H, Chen ZQ, Guo T. Identification of stable QTL and candidate genes involved in anaerobic germination tolerance in rice via high-density genetic mapping and RNA-Seq. *BMC Genomics*. 2019b;20:355.
25. Darvasi A, Weintreb A, Minke V, Weller J, Soller M. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics*. 1993;134:943–51.
26. Darvasi A, Soller M. A simple method to calculate resolving power and confidence interval of QTL map location. *Behav Genet*. 1997;27:125–32.
27. Goffinet B, Gerber S. Quantitative trait loci: a meta-analysis. *Genetics*. 2000;155:463–73.
28. Arcade A, Labourdette A, Falque M, Mangin B, Chardon F, Charcosset A, Joets J. BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. *Bioinformatics*. 2004;20:2324–6.
29. Veyrieras JB, Gofnet B, Charcosset A. MetaQTL: a package of new computational methods for the meta-analysis of QTL mapping experiments. *BMC Bioinformatics*. 2007;8:1–16.
30. Tanin MJ, Saini DK, Sandhu KS, Pal N, Gudi S, Chaudhary J, Sharma A. Consensus genomic regions associated with multiple abiotic stress tolerance in wheat and implications for wheat breeding. *Sci Rep*. 2022;12:13680.
31. Yang L, Lei L, Liu H, Wang J, Zheng H, Zou D. Whole-genome mining of abiotic stress gene loci in rice. *Planta*. 2020;252:85.
32. Khahani B, Tavakol E, Shariati V, Rossini L. Meta-QTL and ortho-MQTL analyses identified genomic regions controlling rice yield, yield-related traits and root architecture under water deficit conditions. *Sci Rep*. 2021;11:6942.
33. Selamat N, Nadarajah KK. Meta-analysis of quantitative trait loci (QTL) identified in drought response in rice (*Oryza sativa* L). *Plants*. 2021;10:716.
34. Saini DK, Srivastava P, Pal N, Gupta PK. Meta-QTL, Ortho-meta-QTL and candidate genes for grain yield and associated traits in wheat (*Triticum aestivum* L). *Theoret Appl Genet*. 2022;135:1049–81.
35. Bilgrami SS, Ramandi HD, Shariati V, Razavi K, Tavakol E, Fakheri BA, Nezhad NM, Ghaderian M. Detection of genomic regions associated with tiller number in Iranian bread wheat under different water regimes using genome-wide association study. *Sci Rep*. 2020;10:1–17.
36. Venske E, Santos RS dos, Farias Dda, Rother R, Maia V, Pegoraro LC, Costa C, de Oliveira A. (2019) Meta-analysis of the QTLome of fusarium head blight resistance in bread wheat: refining the current puzzle. *Front Plant Sci* 10:727.
37. Khahani B, Tavakol E, Shariati V, Fornara F. Genome-wide screening and comparative genome analysis for Meta-QTL, ortho-MQTL and candidate genes controlling yield and yield-related traits in rice. *BMC Genomics*. 2020;21:294.
38. Guo B, Slepner DA, Lu P, Shannon JG, Nguyen HT, Arelli PR. QTL associated with resistance to soybean cyst nematode in soybean: meta-analysis of QTL locations. *Crop Sci*. 2006;46:595–602.
39. Aloryi KD, Okpala NE, Amo A, Bello SF, Akaba S, Tian X. A meta-quantitative trait loci analysis identified consensus genomic regions and candidate genes associated with grain yield in rice. *Front Plant Sci*. 2022;13:1035851.
40. Prakash NR, Lokeshkumar BM, Rathor S, Warraich AS, Yadav S, Vinaykumar NM, Dushynthkumar BM, Krishnamurthy SL, Sharma PC. 2022 Meta-analysis and validation of genomic loci governing seedling and reproductive stage salinity tolerance in rice. *Physiol Plant* 174:e13629.
41. Daryani P, Darzi Ramandi H, Dezhsetan S, Mirdar Mansuri R, Hosseini Salekdeh G, Shobbar ZS. Pinpointing genomic regions associated with root system architecture in rice through an integrative meta-analysis approach. *Theor Appl Genet*. 2022;135:81–106.
42. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant*. 2020;13:1194–202.
43. Hsu S-K, Tung C-W. RNA-seq analysis of diverse rice genotypes to identify the genes controlling coleoptile growth during submerged germination. *Front Plant Sci*. 2017;8:762.
44. Thapa R, Tabien RE, Johnson CD, Septiningsih EM. Comparative transcriptomic analysis of germinating rice seedlings to individual and combined anaerobic and cold stress. *BMC Genomics*. 2023;24:185.
45. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47:607–13.
46. Alam R, Hummel M, Yeung E, Locke AM, Ignacio JCI, Baltazar MD, Jia Z, Ismail AM, Septiningsih EM, Bailey-Serres J. Flood resilience loci *SUBMERGENCE 1* and *ANAEROBIC GERMINATION 1* interact in seedlings established underwater. *Plant Direct*. 2020;4:e00240.
47. Islam M, Ontoy J, Subudhi PK. Meta-analysis of quantitative trait loci associated with seedling-stage salt tolerance in rice (*Oryza sativa* L). *Plants*. 2019;8:33.
48. Murray MG, Thompson WF. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res*. 1980;8:4321–6.
49. Löfer M, Schoen CC, Miedaner T. Revealing the genetic architecture of FHB resistance in hexaploid wheat (*Triticum aestivum* L.) by QTL meta-analysis. *Mol Breed*. 2009;23:473–88.
50. Kakei Y, Masuda H, Nishizawa NK, Hattori H, Aung MS. Elucidation of novel cis-regulatory elements and promoter structures involved in iron excess response mechanisms in rice using a bioinformatics approach. *Front Plant Sci*. 2021;12:303.
51. Zhang ZY, Li JJ, Pan YH, Li JL, Zhou L, Shi HL, Zeng YW, Guo HF, Yang SM, Zheng WW. Natural variation in CTB4a enhances rice adaptation to cold habitats. *Nat Commun*. 2017;8(1):1–13.
52. Li N, Zheng HL, Cui JN, Wang JG, Liu HL, Sun J, Liu TT, Zhao HW, Lai YC, Zou DT. Genome-wide association study and candidate gene analysis of alkalinity tolerance in japonica rice germplasm at the seedling stage. *Rice*. 2019;12(1):24.
53. Xu L, Li X, Wang X, Xiong D, Wang F. Comparing the grain yields of direct-seeded and transplanted rice: a meta-analysis. *Agronomy*. 2019;9(11):767.
54. Panneerselvam P, Kumar V, Banik NC, Kumar V, Parida N, Wasim I, Das A, Pattanaik S, Roul PK, Sarangi DR, Sagwal PK, Craufurd P, Balwinder-Singh, Yadav A, Malik RK, Singh S, McDonald AJ. Transforming labor requirement, crop yield, and profitability with precision dry-direct seeding of rice and integrated weed management in Eastern India. *Field Crops Res*. 2020;259:107961.
55. Ella ES, Setter TL. Importance of seed carbohydrates in rice seedling establishment under anoxia. *Acta Hort*. 1999;504:209–16.
56. Wu Y, Huang M, Tao X, Guo T, Chen Z, Xiao W. Quantitative trait loci identification and meta-analysis for rice panicle-related traits. *Mol Genet Genom*. 2016;291:1927–40.
57. Ballini E, Morel JB, Droc G, Price A, Courtois B, Notteghem JL, Tharreau D. A genome-wide meta-analysis of rice blast resistance genes and quantitative trait loci provides new insights into partial and complete resistance. *Mol Plant Microbe Interact*. 2008;21(7):859–68.
58. Courtois B, Ahmadi N, Khowaja F, Price AH, Rami JF, Frouin J, Hamelin C, Ruiz M. Rice root genetic architecture: Meta-analysis from a drought QTL database. *Rice*. 2009;2:115–28.
59. Navea IP, Maung PP, Yang S, Han JH, Jing W, Shin NH, Zhang W, Chin JH. (2023) A meta-QTL analysis highlights genomic hotspots associated with phosphorus use efficiency in rice (*Oryza sativa* L). *Front Plant Sci* 14:1226297.
60. Kumar S, Singh VP, Saini DK, Sharma H, Gautam S, Kumar S, Balyan HS, Gupta PK. Meta-QTL, ortho-MQTL and candidate genes for thermotolerance in wheat (*Triticum aestivum* L). *Mol Breed*. 2021;41:69.
61. Amo A, Soriano JM. Unravelling consensus genomic regions conferring leaf rust resistance in wheat via meta-QTL analysis. *Plant Genome*. 2022;15(1):e20185.
62. Soriano JM, Colasuonno P, Marcotulli I, Gadaleta A. Meta-QTL analysis and identification of candidate genes for quality, abiotic and biotic stress in durum wheat. *Sci Rep*. 2021;11:11877.
63. Li N, Miao Y, Ma J, Zhang P, Chen T, Liu Y, Che Z, Shahinnia F, Yang D. Consensus genomic regions for grain quality traits in wheat revealed by Meta-QTL analysis and in silico transcriptome integration. *Plant Genome*. 2023;16(2):e20336.
64. Rahmanzadeh A, Khahani B, Taghavi SM, Khojasteh M, Osdaghi E. Genome-wide meta-QTL analyses provide novel insight into disease resistance repertoires in common bean. *BMC Genomics*. 2022;23(1):680.
65. Karnatam KS, Jaganathan D, Dilip KR, Boopathi NM, Muthurajan R. Shortlisting putative candidate genes underlying qDTY1.1, a major effect drought tolerant QTL in rice (*Oryza sativa* L). *Electron J Plant Breed*. 2020;11:916–24.
66. Nakazono M, Tsuji H, Li Y, Saisho D, Arimura S, Tsutsumi N, Hirai A. Expression of a gene encoding mitochondrial aldehyde dehydrogenase in rice increases under submerged conditions. *Plant Physiol*. 2000;124(2):587–98.
67. Bartels D. Targeting detoxification pathways: an efficient approach to obtain plants with multiple stress tolerance? *Trends Plant Sci*. 2001;6:284–6.
68. Kotchoni SO, Kuhns C, Ditzer A, Kirch HH, Bartels D. Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to abiotic stress and protects plants against lipid peroxidation and oxidative stress. *Plant Cell Environ*. 2006;29:1033–48.
69. Singh P, Sinha AK. A positive feedback loop governed by SUB1A1 interaction with MITOGEN-ACTIVATED PROTEIN KINASE3 imparts submergence tolerance in rice. *Plant Cell*. 2016;28:1127–43.

70. Lin Y, Li W, Zhang Y, Xia C, Liu Y, Wang C, Xu R, Zhang L. Identification of genes/proteins related to submergence tolerance by transcriptome and proteome analyses in soybean. *Sci Rep*. 2019;9:14688.
71. Cusick ME, Klitgord N, Vidal M, Hill DE. Interactome: gateway into systems biology. *Hum Mol Genet*. 2005;14:171–81.
72. He F, Zhang Y, Chen H, Peng Y-L. The prediction of protein-protein interaction networks in rice blast fungus. *BMC Genomics*. 2008;9:519.
73. Fukao T, Xiong L. Genetic mechanisms conferring adaptation to submergence and drought in rice: simple or complex? *Curr Opin Plant Biol*. 2013;16:196–204.
74. Lin C, Lee W, Zeng C, Chou M, Lin T, Lin C, Ho M, Shih M. (2023) SUB1A-1 anchors a regulatory cascade for epigenetic and transcriptional controls of submergence tolerance in rice. *PNAS Nexus* 2(7).
75. Chen W, Yao Q, Patil GB, Agarwal G, Deshmukh RK, Lin L, Wang B, Wang Y, Prince SJ, Song L, Xu D, An YC, Valliyodan B, Varshney RK, Nguyen HT. Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. *Front Plant Sci*. 2016;7:1044.
76. Joo J, Lee YH, Song SI. Rice CatA, CatB, and CatC are involved in environmental stress response, root growth, and photorespiration, respectively. *J Plant Biol*. 2014;57:375–82.
77. Das P, Nutan KK, Singla-pareek SL, Pareek A. Oxidative environment and redox homeostasis in plants: dissecting out significant contribution of major cellular organelles. *Front Environ Sci*. 2015;2:1–11.
78. Woodhouse MR, Cannon EK, li JLP, Harper LC, Gardiner JM, Schaeffer ML, Andorf CM. (2021) A pan-genomic approach to genome databases using maize as a model system. *BMC Plant Biol*, 1–10.
79. Cui Y, Wang M, Zhou H, Li M, Huang L, Yin X, Zhao G, Lin F, Xia X, Xu G. OsSGL, a novel DUF1645 domain-containing protein, confers enhanced drought tolerance in transgenic rice and arabidopsis. *Front Plant Sci*. 2016;7:2001.
80. Wang M, Lu X, Xu G, Yin X, Cui Y, Huang L, Rocha PSCF, Xia X. OsSGL, a novel pleiotropic stress-related gene enhances grain length and yield in rice. *Sci Rep*. 2016;6:38157.
81. Li Y, Wang W, Hu C, Yang S, Ma C, Wu J, Wang Y, Xu Z, Li L, Huang Z, Zhu J, Jia X, Ye X, Yang Z, Sun Y, Liu H, Chen R. Ectopic expression of a Maize Gene ZmDUF1645 in Rice increases grain length and yield, but reduces Drought stress tolerance. *Int J Mol Sci*. 2023;24(12):9794.
82. Kim CY, Vo KTX, Nguyen CD, Jeong D-H, Lee S-K, Kumar M, Kim S-R, Park S-H, Kim J-K, Jeon J-S. Functional analysis of a cold-responsive rice WRKY gene, OsWRKY71. *Plant Biotechnol Rep*. 2016;10:13–23.
83. Wang Y, Afeworki Y, Geng S, Kanchupati P, Gu M, Martins C, Rude B, Tefera H, Kim Y, Ge X, Auger D. Hydrootropism in the primary roots of maize. *New Phytol*. 2020;226:1796–808.
84. Campo S, Baldrich P, Messeguer J, Lalanne E, Coca M, Segundo BS. Overexpression of a calcium-dependent protein kinase confers salt and drought tolerance in rice by preventing membrane lipid peroxidation. *Plant Physiol*. 2014;165:688–704.
85. Fujino K, Sekiguchi H. Origins of functional nucleotide polymorphisms in a major quantitative trait locus, qLTG3-1, controlling low-temperature germinability in rice. *Plant Mol Biol*. 2011;75:1–10.
86. Kurotani K-I, Hayashi K, Hatanaka S, Toda Y, Ogawa D, Ichikawa H, Ishimaru Y, Tashita R, Suzuki T, Ueda M, Hattori T, Takeda S. Elevated levels of CYP94 family gene expression alleviate the jasmonate response and enhance salt tolerance in rice. *Plant Cell Physiol*. 2015;56:779–89.
87. Fukushima S, Mori M, Sugano S, Takatsuji H. Transcription factor WRKY62 plays a role in pathogen defense and hypoxia-responsive gene expression in rice. *Plant Cell Physiol*. 2016;57:2541–51.
88. Viana VE, Marini N, Busanello C, Pegoraro C, Fernando JA, Da Maia LC, de Costa A. Regulation of rice responses to submergence by WRKY transcription factors. *Biol Plant*. 2018;62:551–60.
89. Sharma N, Dang TM, Singh N, Ruzicic S, Mueller-Roeber B, Baumann U, Heuer S. Allelic variants of *OsSUB1A* cause differential expression of transcription factor genes in response to submergence in rice. *Rice*. 2018;11(1):2.
90. Bevan MW, Uauy C, Wulff BB, Zhou J, Krasileva K, Clark MD. Genomic innovation for crop improvement. *Nature*. 2017;543(7645):346–54. <https://doi.org/10.1038/nature22011>.
91. Aloryi KD, Jing H-C, Dijkwel PP. Comparison of leaf senescence regulation between distantly related plant species uncovers knowledge gaps and opportunities for plant improvement strategies. *Environ Exp Bot*. 2023;214:105474.
92. Angaji SA. Mapping QTL for submergence tolerance during germination in rice. *Afr J Biotechnol*. 2008;7:2551–8.
93. Baltazar MD, Ignacio JCI, Thomson MJ, et al. QTL mapping for tolerance of anaerobic germination from IR64 and the aus landrace Nanhi using SNP genotyping. *Euphytica*. 2014;197:251–60.
94. Jiang L, Hou M, Wang C, Wan J. Quantitative trait loci and epistatic analysis of seed anoxia germinability in rice (*Oryza sativa*). *Rice Sci*. 2004;11:238–44.
95. Jiang L, Liu S, Hou M, Tang J, Chen L, Zhai H, Wan J. Analysis of QTL for seed low-temperature germinability and anoxia germinability in rice (*Oryza sativa* L). *Field Crops Res*. 2006;98:68–75.
96. Kim SM, Reinke RF. Identification of QTL for tolerance to hypoxia during germination in rice. *Euphytica*. 2018;214:160.
97. Ling J, Ming-yu HW, Ming C, Jian-min W. Quantitative trait loci and epistatic analysis of seed anoxia germinability in rice (*Oryza sativa* L). *Rice Sci*. 2004;11:238–44.
98. Septiningsih EM, Sanchez DL, Singh N, Sendon PM, Pamplona AM, Heuer S, Mackill DJ. Identifying novel QTL for submergence tolerance in rice cultivars IR72 and Madabar. *Theor Appl Genet*. 2012;124(5):867–74.
99. Septiningsih EM, Ignacio JC, Sendon PM, Sanchez DL, Ismail AM, Mackill DJ. QTL mapping and confirmation for tolerance of anaerobic conditions during germination derived from the rice landrace Ma-Zhan Red. *Theor Appl Genet*. 2013;126(5):1357–66.
100. Tang DQ, Kasai Y, Miyamoto N, Ukai Y, Nemoto K. Comparison of QTL for elongation ability between two floating rice cultivars with a different phylogenetic origin. *Breed Sci*. 2005;55:1–5.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.