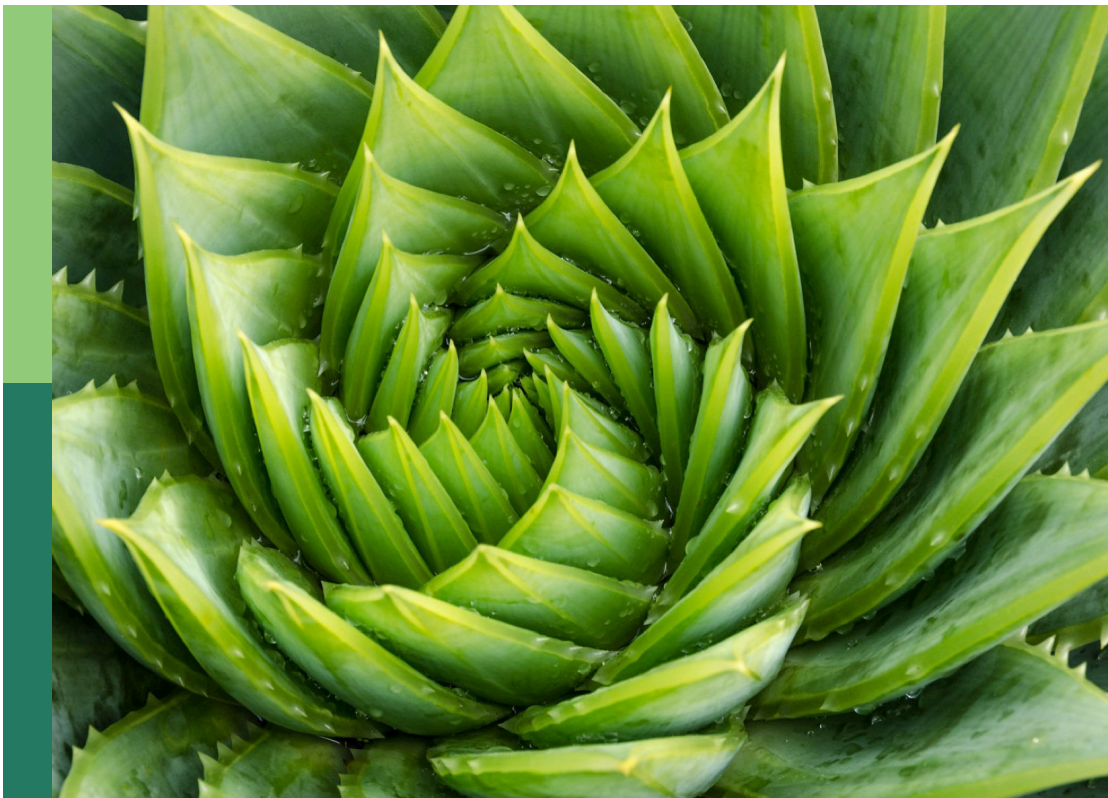
frontiers | Research Topics

Mining and utilization of favorable gene resources in rice

Edited by
Xiaojin Luo, Xueyong Li
and Ryuji Ishikawa

Coordinated by
Ying Wang

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-3608-7
DOI 10.3389/978-2-8325-3608-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Mining and utilization of favorable gene resources in rice

Topic editors
Xiaojin Luo — Fudan University, China
Xueyong Li — Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, China
Ryuji Ishikawa — Hirosaki University, Japan

Topic coordinator
Ying Wang — Fudan University, China

Citation
Luo, X., Li, X., Ishikawa, R., Wang, Y., eds. (2023). *Mining and utilization of favorable gene resources in rice*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3608-7

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 19 September 2023
DOI 10.3389/fpls.2023.1289069

OPEN ACCESS

EDITED AND REVIEWED BY
Jihong Hu,
Northwest A&F University, China

*CORRESPONDENCE
Xiaojin Luo
luoxj@fudan.edu.cn

RECEIVED 05 September 2023
ACCEPTED 13 September 2023
PUBLISHED 19 September 2023

CITATION
Wang Y, Li X, Ishikawa R and Luo X (2023) Editorial: Mining and utilization of favorable gene resources in rice.
Front. Plant Sci. 14:1289069.
doi: 10.3389/fpls.2023.1289069

Editorial: Mining and utilization of favorable gene resources in rice

Ying Wang[1,2], Xueyong Li[3], Ryuji Ishikawa[4] and Xiaojin Luo[1]*

[1]State Key Laboratory of Genetic Engineering and Engineering Research Center of Gene Technology (Ministry of Education), School of Life Sciences, Fudan University, Shanghai, China, [2]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Department of Ecology and Evolutionary Biology, School of Life Sciences, Fudan University, Shanghai, China, [3]National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, [4]Faculty of Agriculture and Life Science, Hirosaki University, Hirosaki, Aomori, Japan

KEYWORDS

gene resources, germplasm utilization, yield component traits, mapping and cloning, genome-wide association study, molecular design breeding, rice

Editorial on the Research Topic
Mining and utilization of favorable gene resources in rice

Agriculture has been evolving along with the development of human society. Agricultural germplasm resources play a crucial role in supporting the development of human civilization (Doebley et al., 2006; Tian et al., 2021). With the change of agricultural population structure, there are new requirements for germplasm resources (Yu and Li, 2021). Combined with global climate change, crops should also be able to cope with the challenges of environmental stresses and maintain a stable increase in production (Shahzad et al., 2021). Subsequently, rice has new breeding requirements in production, quality, and yield (Zeng et al., 2017). This Research Topic covers multiple approaches to exploit rice germplasm, excavate and utilize the key genes, and understand the molecular basis of yield component traits.

Mapping and cloning of new genes or revealing new functions

The classical mapping and cloning technique still play an important role in the discovery of new genetic loci. Kim et al. cloned the *OsD2* gene by map-based cloning based on the previously reported quantitative trait locus (QTL) *qLTG1* responsible for low-temperature germinability in rice. Lan et al. identified the *hfr131* gene, which affects plant height and grain length in rice, through genetic analysis of introgression line. Yin et al. detected 16 QTL ranges for salt tolerance at bud stage in rice and a new candidate gene *LOC_Os12g25200*, using recombinant inbred line populations.

Interestingly, these works revealed new functions of known genes. *OsD2* is allelic to *DWARF2*, which encodes a cytochrome P450 involved in brassinosteroid (BR) biosynthesis. Although most alleles of the *dwarf2* mutant showed severe dwarfism with undesirable traits which reduced their breeding application value, the *OsD2* allele improved

low-temperature germinability without any harmful phenotype. As a new allele of the BR receptor *OsBRI1*, *hfr131* affected rice plant height through the dn-type internode elongation pattern, which had not been reported previously in any BR biosynthesis- or *OsBRI1*-defective mutants. These findings suggest that the SNPs polymorphism may confer new functions to the known genes, and the discovery of new alleles is helpful to expand the diversity of gene function and mechanism.

Genome-wide association study facilitates mining of loci associated with complex agronomic traits

Complex agronomic traits of crops involve complex genetic loci, and genome-wide association study (GWAS) can significantly improve the identification efficiency of potential QTLs. Li et al. conducted GWAS using 211 rice accessions to determine salt tolerance germinability indices. A total of 43 QTLs were identified, 18 of which were co-localized with previous studies. According to the RNA-seq and haplotype analysis, rice varieties with elite haplotypes in *LOC_Os03g13560*, *LOC_Os03g13840* and *LOC_Os03g14180* genes had high salt tolerance germinability. Dai et al. detected high-quality loci responsible for high seed vigor from 346 diverse accessions. By GWAS, 51 significant SNPs were identified, which were further validated using chromosome segment substitution lines. Integrating gene expression, gene annotation, and haplotype analysis, 21 strong candidate genes were identified. The functions of *LOC_Os01g11270* and *LOC_Os01g55240* were further verified by CRISPR/Cas9. Xu et al. combined GWAS and linkage mapping to analyze the candidate intervals for seedling salinity tolerance of 295 japonica rice varieties. After identifying the lead SNP (Chr12_20864157), the candidate gene *LOC_Os12g34450* was obtained by haplotype analysis, qRT-PCR, and sequence analysis.

Salt tolerant germinability, high seed vigor and seedling salinity tolerance correspond to the key problems affecting the yield of rice in agricultural production. By GWAS, these works verified the known genetic loci and found new functional genes, which provided a promising resource for solving the problems. In addition, the integration of multiple methods (chromosome segment substitution, linkage mapping, RNA-seq database, haplotype analysis, CRISPR/Cas9, etc.) can validate the results of GWAS and help identify candidate genes.

Functional expansion of key gene resources via phylogenetic analysis and reverse genetics

Functional expansion of known key genes is another effective pathway for favorable gene resources. Papain-like cysteine proteases (PLCPs) play a crucial role in plant growth and development. Li et al. constructed CRISPR/Cas9 lines and showed that a PLCP, an oryzain alpha chain precursor (OCP), negatively regulated resistance to blast disease. OCP interacted with OsRACK1A or OsSNAP32 and influenced the expression of many resistance-related genes. Zheng et al.'s work was based on the cell cycle-associated transcription factors E2F (E2 promoter binding factor), and their downstream target gene the mismatch repair-related gene *OsMSHs* (Mutated S homologue). They systematically categorized six rice E2Fs, and constructed four *msh* mutants using the CRISPR-Cas9 technique. This study elucidated the mechanism of *E2F* and *MSH* for enhancing cadmium stress tolerance in rice. Zeng et al. focused on the glutamate-like receptor (GLR) genes, which play a crucial role in signal transduction and communication. An integrated approach involving phylogenetic analysis, phenotypic and functional characterization and comprehensive population genetic analyses was employed to understand the functionalities of 26 rice *GLR* genes. The results suggested that natural variations at most rice *GLR* loci had potential value in improving the productivity and tolerance to abiotic stresses. These studies expand our understanding of these known key gene functions, and genes in related biological processes and molecular networks are also worthy for the further exploration of effective gene resources.

Finally, one review article from Zhong et al. presented a comprehensive review on the recent breakthroughs in rice yield traits and molecular design breeding in China. The authors believed that the further mining of genes and gene regulatory networks, development and utilization of molecular markers, establishment of the high-throughput and low-cost genotype detection system, and rational aggregation of high-quality genes (genes related to yield traits, stress resistance traits and quality traits) to breed new varieties with high yield and superior quality, will be an inevitable trend in future rice research.

Perspectives

In summary, this Research Topic brought recent advances in the mining and utilization of rice germplasm resource. These studies indicated the new situations in the exploration of favorable gene resources. For example, crop yield breeding needs to balance the influence of environmental stress, and molecular design breeding needs to consider the pleiotropy between genes and complex regulatory networks among genes. Overcoming these problems will open the way to achieve a further breakthrough in the current yield level.

Author contributions

YW: Writing – original draft. XLi: Writing – review & editing. RI: Writing – review & editing. XLu: Writing – review & editing.

Funding

This work was supported by National Natural Science Foundation of China (31900393), Shanghai Rising-Star Program (20QB1402100), Natural Science Foundation of Shanghai

(19ZR1436600), and National Key Research and Development Program of China (2022YFD1200103).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Doebley, J. F., Gaut, B. S., and Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell* 127 (7), 1309–1321. doi: 10.1016/j.cell.2006.12.006

Shahzad, A., Ullah, S., Dar, A. A., Sardar, M. F., Mehmood, T., Tufail, M. A., et al. (2021). Nexus on climate change: agriculture and possible solution to cope future climate change stresses. *Environ. Sci. pollut. Res. Int.* 28 (12), 14211–14232. doi: 10.1007/s11356-021-12649-8

Tian, Z., Wang, J. W., Li, J., and Han, B. (2021). Designing future crops: challenges and strategies for sustainable agriculture. *Plant J.* 105 (5), 1165–1178. doi: 10.1111/tpj.15107

Yu, H., and Li, J. (2021). Short- and long-term challenges in crop breeding. *Natl. Sci. Rev.* 8 (2), nwab002. doi: 10.1093/nsr/nwab002

Zeng, D., Tian, Z., Rao, Y., Dong, G., Yang, Y., Huang, L., et al. (2017). Rational design of high-yield and superior-quality rice. *Nat. Plants* 3, 17031. doi: 10.1038/nplants.2017.31

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 15 July 2022
doi: 10.3389/fpls.2022.934515

Genome-Wide Association Study Reveals a Genetic Mechanism of Salt Tolerance Germinability in Rice (*Oryza sativa* L.)

Caijing Li[1,2†], *Changsheng Lu*[1,2†], *Baoli Zou*[1,2], *Mengmeng Yang*[1,2], *Guangliang Wu*[1,2], *Peng Wang*[1,2], *Qin Cheng*[1,2], *Yanning Wang*[1,2], *Qi Zhong*[1,2], *Shiying Huang*[1,2], *Tao Huang*[1,2], *Haohua He*[1,2*] *and Jianmin Bian*[1,2*]

[1] *Key Laboratory of Crop Physiology, Ecology and Genetic Breeding, Ministry of Education, Nanchang, China,* [2] *Key Laboratory of Crop Physiology, Ecology and Genetic Breeding, Nanchang, China*

Salt stress is one of the factors that limits rice production, and an important task for researchers is to cultivate rice with strong salt tolerance. In this study, 211 rice accessions were used to determine salt tolerance germinability (STG) indices and conduct a genome-wide association study (GWAS) using 36,727 SNPs. The relative germination energy (RGE), relative germination index (RGI), relative vigor index (RVI), relative mean germination time (RMGT), relative shoot length (RSL), and relative root length (RRL) were used to determine the STG indices in rice. A total of 43 QTLs, including 15 for the RGE, 6 for the RGI, 7 for the RVI, 3 for the RMGT, 1 for the RSL, and 11 for the RRL, were identified on nine chromosome regions under 60 and 100 mM NaCl conditions. For these STG-related QTLs, 18 QTLs were co-localized with previous studies, and some characterized salt-tolerance genes, such as *OsCOIN*, *OsHsp17.0*, and *OsDREB2A*, are located in these QTL candidates. Among the 25 novel QTLs, *qRGE60-1-2* co-localized with *qRGI60-1-1* on chromosome 1, and *qRGE60-3-1* and *qRVI60-3-1* co-localized on chromosome 3. According to the RNA-seq database, 16 genes, including nine for *qRGE60-1-2* (*qRGI60-1-1*) and seven for *qRGE60-3-1* (*qRVI60-3-1*), were found to show significant differences in their expression levels between the control and salt treatments. Furthermore, the expression patterns of these differentially expressed genes were analyzed, and nine genes (five for *qRGE60-1-2* and four for *qRGE60-3-1*) were highly expressed in embryos at the germination stage. Haplotype analysis of these nine genes showed that the rice varieties with elite haplotypes in the *LOC_Os03g13560*, *LOC_Os03g13840*, and *LOC_Os03g14180* genes had high STG. GWAS validated the known genes underlying salt tolerance and identified novel loci that could enrich the current gene pool related to salt tolerance. The resources with high STG and significant loci identified in this study are potentially useful in breeding for salt tolerance.

Keywords: salt tolerance germinability, rice accessions, GWAS, QTLs, haplotype analysis

OPEN ACCESS

Edited by:
Xiaojin Luo,
Fudan University, China

Reviewed by:
Yoshiaki Ueda,
Japan International Research Center for Agricultural Sciences (JIRCAS), Japan
Liang Chen,
Shanghai Agrobiological Gene Center, China
Hongwei Cai,
China Agricultural University, China

***Correspondence:**
Haohua He
hhhua64@163.com
Jianmin Bian
jmbian81@126.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Bioinformatics, a section of the journal Frontiers in Plant Science

Received: *02 May 2022*
Accepted: *16 June 2022*
Published: *15 July 2022*

Citation:
Li C, Lu C, Zou B, Yang M, Wu G, Wang P, Cheng Q, Wang Y, Zhong Q, Huang S, Huang T, He H and Bian J (2022) Genome-Wide Association Study Reveals a Genetic Mechanism of Salt Tolerance Germinability in Rice (Oryza sativa L.). Front. Plant Sci. 13:934515. doi: 10.3389/fpls.2022.934515

INTRODUCTION

As the world's population continues to increase, researchers predict that food production must increase by at least 70% over the next 40 years to meet demand (Tilman et al., 2011). Rice is one of the world's most important food crops and feeds more than half of the world's population. Therefore, improving rice yields under conditions of limited arable land is an urgent problem for researchers to solve. The traditional rice transplanting mode is time-consuming and labor-intensive and has been gradually replaced in recent years by a direct seeding mode in many areas (Wang et al., 2011); at the same time, this development also means higher requirements for the germination ability of rice seeds in harsh environments.

Salt stress is one of the main factors that limits direct seeding of rice. Salinity alters the osmotic potential of the germination medium, enzyme activities in cells, and destroys cell structures so that seeds cannot germinate normally and establish stable stands of seedlings, which finally affects rice yields (Koyro, 2002; Khan and Weber, 2006; Gomes-Filho et al., 2008). Therefore, unraveling the genetic architecture for salt tolerance germinability and cultivating rice varieties with high STG are very important in accomplishing direct rice seeding.

To our knowledge, only a few genes have been reported to be involved in STG. Of those, OsHsp17.0 encodes heat shock proteins (Hsps), and its overexpression in plants demonstrated higher germination abilities than those of wild-type (WT) plants when subjected to NaCl (Zou et al., 2012). The basic helix-loop-helix (bHLH) transcription factor, OsbHLH035, is a salt-induced gene, and Osbhlh035 mutants exhibit delayed seed germination, especially under salt stress conditions (Chen et al., 2018). Under salt stress, qSE3 promoted the absorption of K+ and Na+, induced abscisic acid biosynthesis and abscisic acid signaling pathway gene expression, inhibited the accumulation of reactive oxygen species in seeds, and thus improved the salt tolerance of seeds during germination (He et al., 2019). NaCl promoted the expression of *OsNAC45* in roots, and the knockout of *OsNAC45* resulted in greater ROS accumulations in roots and increased the sensitivity of rice to salt stress (Zhang et al., 2020). Therefore, to obtain more causal genes, we must use more germplasm resources, identify more QTLs, and screen more candidate genes to better understand the genetic mechanism of STG.

Salt tolerance is a complex quantitative trait that is often controlled by multiple genes and the environment. To better understand the genetic mechanism of salt tolerance in rice at the germination stage, many QTLs have been identified by using traditional QTL mapping methods in biparental populations. A total of 16 QTLs were detected during the germination stage under 100 mM NaCl conditions from a RILs that were derived from a cross between Jiucaiqing (salt-tolerant) and IR26 (salt-susceptible; Wang et al., 2011); a total of 11 QTLs were identified for their salt tolerance (1.5% NaCl) at the germination stage by using linkage mapping from a $BC_2F_{2:3}$ population including 190 accessions that were derived from Dongnong 425 and Changbai 10 (Zheng et al., 2014); 17 QTLs that were related to germination traits under salt stress (80 mM NaCl) in an F2:4 population were detected, which were derived from a cross between the salt-tolerant variety "Gharib" and salt-sensitive variety "Sepidroud" (Mardani et al., 2014); 10 QTLs for five salt stress (0.01 mol L^{-1})-related traits were detected a population of backcross inbred lines, which was derived from a cross between Changhui 891 and 02428 with 3,057 bin markers (Luo et al., 2020); 13 QTLs were identified under H_2O conditions and salt conditions (300 mM NaCl) by using a BC1F2 population that derived from a WJZ/Nip cross (Zeng et al., 2021); 23 QTLs for germination parameters were detected under 18 dS m^{-1} NaCl solution conditions by using a BIL population that was derived from a backcross of an African rice, ACC9, as the donor and *indica* cultivar Zhenshan97 (ZS97) as the recurrent parent (Nakhla et al., 2021). Biparental populations, although, have played an important role in the past to provide a resource for researchers to work with; construction of populations entails major investments in time, which has therefore limited the number of genes identified to date (Bian et al., 2010).

With the rapid development of NGS (next-generation sequencing) and the existence of high-density SNP markers obtained through resequencing, an increasing number of studies have used GWAS as an efficient strategy to replace the traditional QTL mapping method (Huang et al., 2012). In rice, some studies have applied this method to study rice STG indices. A total of nine significant markers were identified under salt conditions by using 276 *indica* accessions (Cheng et al., 2015); 11 QTLs were identified based on the stress-susceptibility indices (SSIs) of the vigor index (VI) and mean germination time (MGT) in 478 diverse rice accessions (Shi et al., 2017); six quantitative trait nucleotides (QTNs) that affect salt tolerance at the germination stage were identified by GWAS using a panel of 208 rice mini-core accessions (Naveed et al., 2018); 371 QTNs were identified by using six multilocus GWAS methods for the salt tolerance traits of 478 rice accessions at the seed germination stage (Cui et al., 2018); a total of 12 associated peaks were detected by using GWAS in 295 rice accessions during the germination stage (Yu et al., 2018); and 21 QTLs associated with salinity stress were detected in 498 highly diverse rice accessions (Islam et al., 2022). However, to our knowledge, there are still few studies on the identification of QTLs/genes, particularly for the salt tolerance in the germination stage using GWAS in rice. Here, we applied GWAS mapping and used 36,727 SNPs that covered all 12 rice chromosomes in a natural population that consisted of 211 rice accessions to identify QTLs/candidate genes that may contribute to salt tolerance during the rice germination stage with the aim of guiding the breeding of salt-tolerant rice varieties.

MATERIALS AND METHODS

Plant Materials

All materials used in this study were derived from Li et al. (2021). A set of 211 accessions of diverse germplasm accessions representing the major rice-growing regions in China was selected from the International Rice Research Institute (https://www.irri.org/) to comprise the materials used in this work. All data of 36,727 SNPs are published at https://snp-seek.irri.org/. This pool of germplasm accessions consists of two

subpopulations, including 130 *indica* and 81 *japonica*, and these 211 rice accessions were divided into two subgroups by using principal component analysis and relationship analysis, which suggested that these germplasm resources had abundant genetic diversity, which was beneficial for performing GWAS.

Evaluation of Salt Tolerance Germinability

All of the yellow, ripe seeds from each accession were dried at 45°C for 2 days to break seed dormancy. The seeds were surface-sterilized with a 15% sodium hypochlorite solution for 15 min and then rinsed three times with sterile distilled water before the germination experiment. A total of 30 sterilized seeds from each accession were placed on two filter papers soaked with 10 ml of sodium chloride in Petri plates (9 cm) during the germination stage to screen the salinity tolerances, and the concentrations of the sodium chloride solutions were 60 and 100 mM. In the control treatment, the same number of seeds per line was placed on filter papers in Petri dishes that were soaked in 10 ml of distilled water. All Petri dishes were incubated under controlled conditions in a growth chamber at a temperature of 28°C with 12 h each of light and dark. The seed germination time was 7 days, and the numbers of seeds from each accession that germinated were recorded each day. On the 7th day, 10 seedlings from each accession were selected for calculations of shoot length (SL) and root length (RL), and the sodium chloride solutions in the Petri dishes were changed every 2 days. All experiments were repeated three times. The relative germination energy (RGE), relative germination index (RGI), relative vigor index (RVI), relative mean germination time (RMGT), relative shoot length (RSL), and relative root length (RRL) were calculated and subjected to a GWAS, and the calculation formula refers to Yu et al. (2018).

GE = Number of germinated seeds at 4 days/Total number of seeds tested × 100%.
GI = Σ(*G*t/t), where *G*t is the number of seeds that germinated on day t.
VI = GI × SL

MGT = Mean germination time (MGT) = $\sum$Dn/$\sum n$, where n is the number of seeds that germinated on D day, and D is the number of days counted.

Relative value = Value under salt stress/control.

All phenotypic data are presented in **Supplementary Table 1**. Each index at the two concentrations is divided into five parts. The larger the RMGT value is, the lower the score, and the larger the values of other indices are, the higher the score. The total score of the 12 indices for each accession represents the STG of this accession, and the specific evaluation criteria are shown in **Supplementary Table 2**.

Genome-Wide Association Analysis

The Tassel 5.2.73 software and a mixed linear model (MLM) with a PCA matrix (the first five PCs were used) and kinship (K matrix) were used to determine the associations among SNP markers and the 12 phenotypic traits. To ensure data accuracy, the phenotypic data were standardized, and the SNP data (36,727 SNPs) were filtered (in Tassel 5.2.73, the genotypic data were first numerically analyzed, then SNPs with minor allele frequencies <0.05 were removed, and 33,777 SNPs were finally obtained and used for GWAS) before association mapping. Manhattan plots were generated using the Cloud platform (http://www.cloud.biomicroclass.com/).

Candidate Gene Analyses

The LD blocks were used to identify candidate gene regions using the Haploview 4.2 software (Barrett et al., 2005). The SNPs with the most significant associations in a block were identified as the leading SNPs, and LD blocks containing significantly associated SNPs were defined as candidate genomic regions. Differentially expressed genes were identified using the Plant Public RNA-seq Database (Yu et al., 2022). The expression patterns of the differentially expressed genes were analyzed using publicly available microarray data (http://www.genevestigator.com/). Haplotype analyses were performed using RiceVarMap V2.0 (Zhao et al., 2015). Information on candidate genes was collected and classified by the NCBI (https://www.ncbi.nlm.nih.gov/), China Rice Data Center (https://www.ricedata.cn/), and the Rice Genome Annotation Project (http://rice.uga.edu/index.shtml).

Quantitative Real-Time PCR Analysis

Embryos of seeds that were germinated in water and in 60 mM NaCl solutions for 2 days were collected for expression analysis. Total RNA was prepared using a MiniBEST Plant RNA Extraction kit (Takara, China). The corresponding sequences of these genes were obtained from the Rice Genome Annotation Project (http://rice.uga.edu/index.shtml). The primers (e.g., *LOC_Os03g13560*-F: CAGATTGTGATATGGTGTTCGGC; *LOC_Os03g13560*-R: GGAGACTGAGAAGCTGTCATCAT; *LOC_Os03g13840*-F: CTGGACAAGGTACTGGAGGAGTA; *LOC_Os03g13840*-R: CGTTCTTTCCAGACACCTCTACA; *LOC_Os03g14180*-F: AGGTGAGGATGCGGTTCG; and *LOC_Os03g14180*-R: CGCTCACAGGCTCACATCC) were designed based on the CDSs of the corresponding genes by using Primer3Plus (https://www.primer3plus.com/). OsActin was used as the internal control. Real-time PCR was carried out using the SYBR Green method, and the relative expression levels were calculated using the $2^{-\Delta\Delta CT}$ method.

Statistical Analysis and Mapping

The mean values and standard errors of the phenotypic data were calculated using Microsoft Excel 2010, and the correlation coefficients were calculated using the SPSS 26.0 software. Box plots were created using the OriginPro 2021 software.

RESULTS

Phenotypic Analysis of the 211 Rice Accessions and Correlations Among STG Indices

In this study, two sodium chloride concentrations (e.g., 60 and 100 mM) were selected to conduct an STG assessment. After 7 days of germination, the growth states of 211 rice varieties under the two salt solutions were inhibited compared with the control

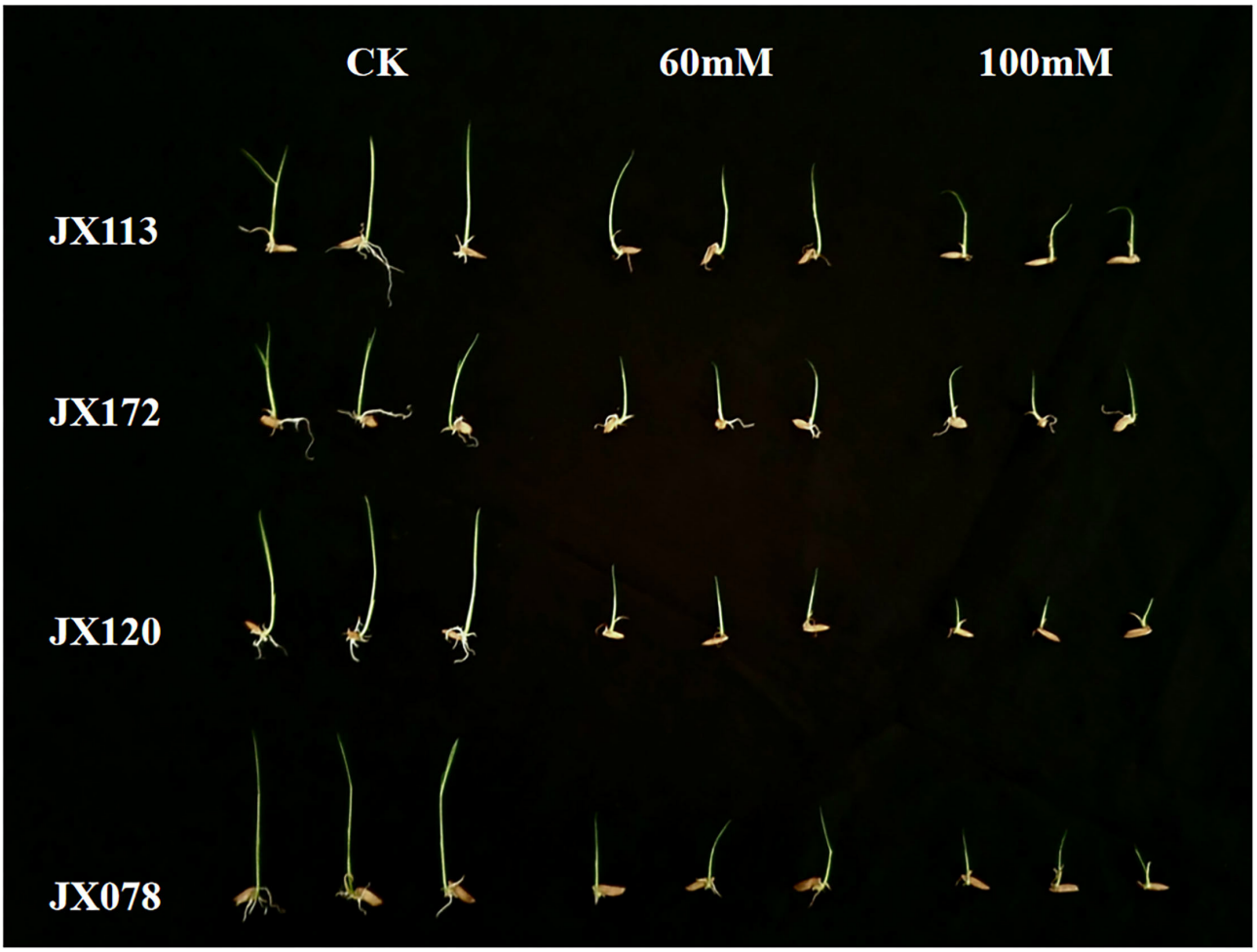


FIGURE 1 | Comparison of growth status of two salt-tolerant rice varieties (JX 113 and JX 172) and two salt-sensitive rice varieties (JX 120 and JX 078) after 7 days of control treatment and 60 and 100 mM NaCl solution.

treatment, but the distribution of the inhibition degree among different varieties was different (**Figure 1**), which indicates that both sodium chloride concentrations are suitable for screening salt-tolerant and salt-sensitive varieties. The salt-tolerant rice varieties identified in this study can be used as excellent parent resources in subsequent breeding projects.

Six indices obtained under two sodium chloride concentrations were used to evaluate the STG in this natural population, and all of the indices showed large phenotypic variations among the rice accessions (**Table 1**). In particular, except for RMGT60 and RMGT100, the values of all indices were widely distributed, and the coefficients of variation were very large. This result suggested that these STG indices may be suitable for performing a GWAS. In addition, six STG-related indices were compared among different subgroups in the population (**Figure 2**; **Supplementary Figure 1**). Under the 60 mM NaCl condition, the RSL, RGE, RGI, and RVI values of the *japonica* group were significantly lower than those of the *indica* group ($p < 0.01$), but the RRL and RMGT values exhibited no significant differences between the two subgroups. On contrary, at an NaCl concentration of 100 mM, the RSL, RGE, RGI, and RVI values in the *indica* subgroup were higher than those in the *japonica* subgroup, but the RRL values in the *japonica* subgroup were higher than those in the *indica* subgroup, and there were no significant differences in the RMGT values between the two subgroups. We used these six indices to define a scoring standard to screen varieties with high STG, and the results showed that most *indica* rice varieties generally had greater STG than the *japonica* rice varieties (**Figure 3**).

To determine how the mean values of the different STG indices for each accession compared to each other, pairwise Pearson's correlation analysis was conducted (**Figure 4**). Under the 60 mM NaCl condition, the RSL, RRL, RGE, RGI, and RVI values were all positively correlated to varying degrees, and they were all significantly negatively correlated with the RMGT. However, the RSL had no correlation with RRL at a concentration of 100 mM, but both were significantly positively correlated with the RGE, RGI, and RVI values, and the RMGT was significantly negatively correlated with the other five indices. These results suggest that the RGE, RGI, RVI, and RMGT indices

might share genetic pathways under both the 60 and 100 mM salt treatments. In addition, a correlation analysis between the two treatments was conducted. The results showed that there were significant positive correlations between the two, and the correlation coefficient was 0.568.

Identification of QTLs for Salt Tolerance Germinability by GWAS

To investigate the possible genetic architecture of STG, a GWAS was conducted, and Manhattan plots were generated to illustrate the significant SNP associations with STG (**Figure 5**; **Supplementary Figure 2**). A total of 43 QTLs, including 15 for RGE, 6 for RGI, 7 for RVI, 3 for RMGT, 1 for RSL, and 11 for RRL, which were associated with STG, were identified in rice (**Table 2**). On chromosome 1, 10 QTLs were discovered for the first time, among which *qRGE60-1-2* overlapped with *qRGI60-1-1* and was located at 1.2 Mb, and *qRVI60-1-2* and *qRGI60-1-4* were co-localized at 3.5 Mb. In addition, nine QTLs were co-localized with those described in previous studies: *qRRL-100-1* was located at 0.2 Mb and overlapped with *OsCOIN* (Liu et al., 2007); *qRGE60-1-2*, *qRGI60-1-2* and *qRMGT60-1-1* overlapped with each other, and a previously cloned gene related to salt stress, *OsHsp17.0* (Zou et al., 2012), was located in this region; at ∼3 Mb, three QTLs (e.g., *qRGI60-1-3*, *qRVI60-1-1*, and *qRGE60-1-4*) were co-located with each other, and *OsDREB2A* (Mallikarjuna et al., 2011) was located in this region; and two QTLs related to RRL, namely, *qRRL60-1-2* and *qRRL60-1-3*, overlapped with *OsPLDα1* (Shen et al., 2011) and *OrbHLH001* (Chen et al., 2013), respectively. These QTLs had explained phenotypic variances (R^2) ranging from 5.48% for *qRGE60-1-2* to 9.76% for *qRGE60-1-3*. Only three QTLs were identified on chromosome 2, among which *qRRL60-2-1* overlapped with a previously reported STG QTL, *qGR-3d2* (Nakhla et al., 2021), and the peak SNP was chrA02_4380604. Seven QTLs mapped to chromosome 3 and *qRGE60-3-1* and *qRVI60-3-1* shared the same SNP peak at 7.66 Mb, three QTLs from two indices (e.g., *qRGE60-3-3*, *qRGE100-3*, and *qRVI60-3-2*) shared the same SNP peak at 9.85 Mb, while this QTL region contains *OsMAPK5* (Xiong and Yang, 2003), a salt stress-related gene. These QTLs had explained phenotypic variances (R^2) ranging from 7.46% for *qRVI60-3-1* to 12.27% for *qRGE60-3-3*. In the remaining chromosomes,

TABLE 1 | The distribution of STG indices in 211 rice accessions.

Indices	Range (%)	Mean (%)	SD	Coefficient of variation (%)
RSL60	35–161	81	0.15	18.34
RRL60	1–176	86	0.25	28.52
RGE60	8–167	87	0.21	23.94
RGI60	26–137	82	0.14	17.44
RVI60	9–130	68	0.19	27.39
RMGT60	95–118	104	0.03	3.31
RSL100	21–112	64	0.15	23.10
RRL100	1–130	60	0.20	33.89
RGE100	1–112	73	0.26	36.11
RGI100	20–112	68	0.17	24.40
RVI100	5–98	45	0.17	37.48
RMGT100	98–123	108	0.05	4.28

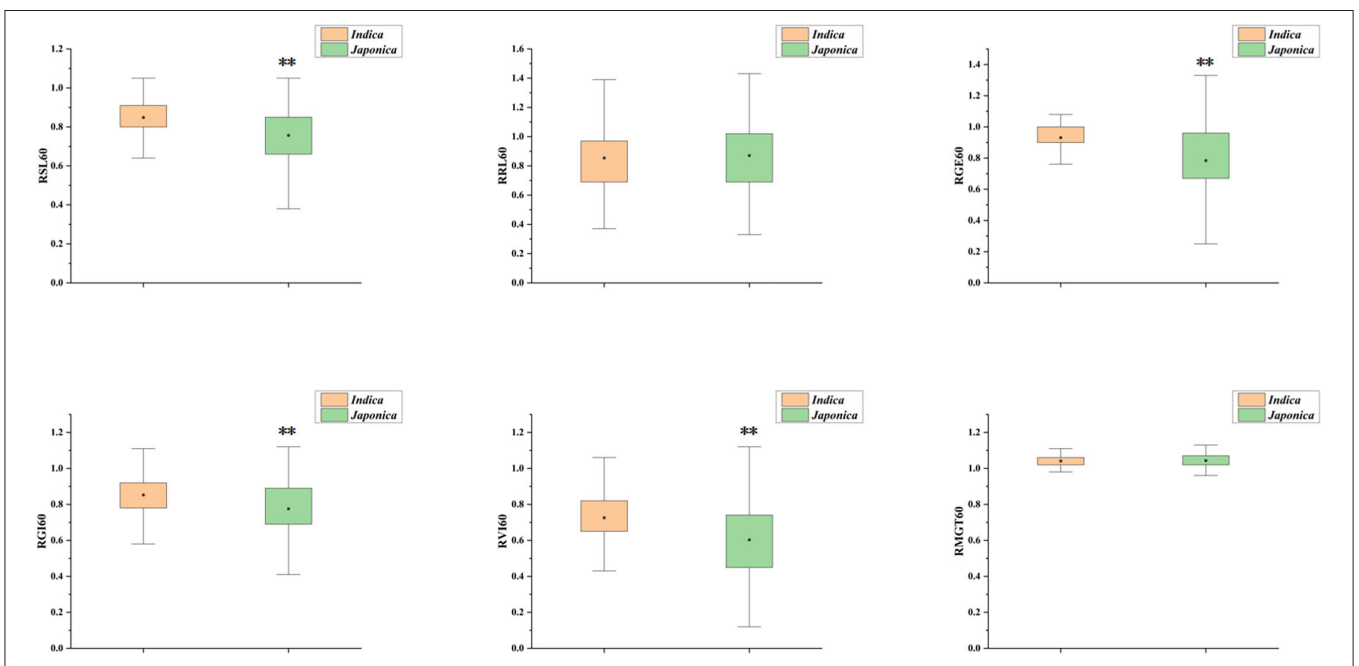


FIGURE 2 | Box plot for comparison of six indices between *indica* and *japonica* subgroups under 60 mM NaCl. The yellow box represents *indica*, the green box represents *japonica*, the black dot in the box represents the median, and the value range of the box is 25–75%, **Indicates significance at the 1% level.

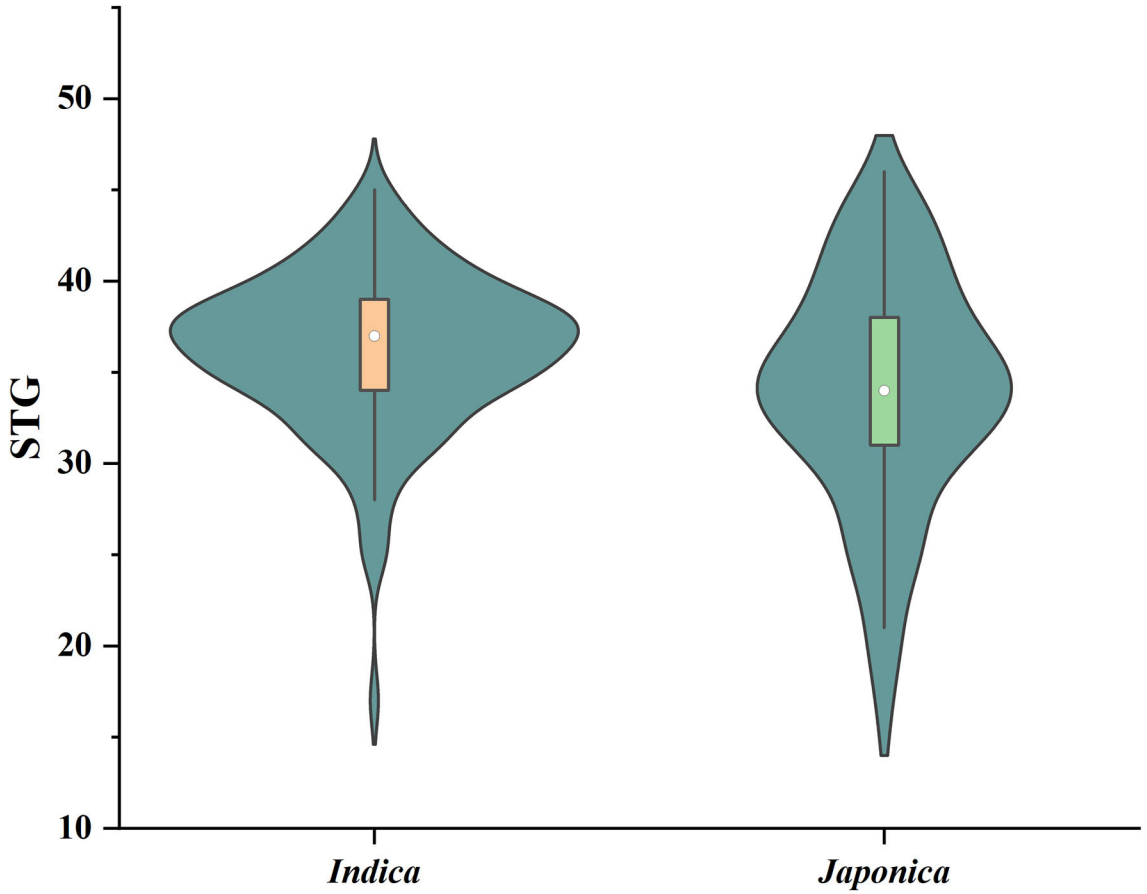


FIGURE 3 | Violin diagram of the distribution of STG values of *indica* and *japonica* subgroups. Boxes in violins range in value from 25 to 75%, with white dots representing the median.

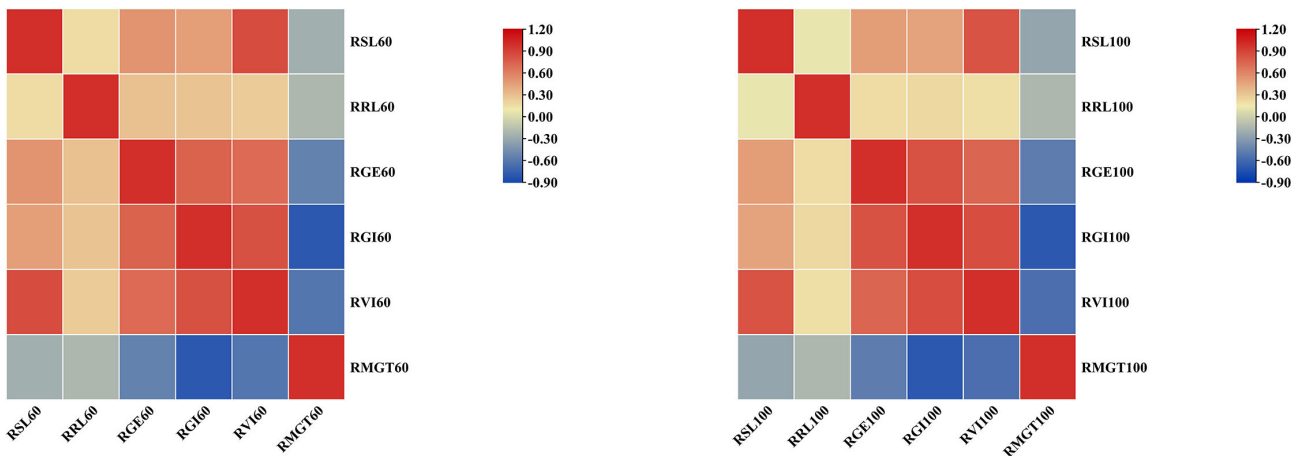


FIGURE 4 | Heat map of correlation analysis between six indices under two different concentrations of salt solution. Blocks that tend to be red or blue indicate greater absolute values of the correlation coefficients.

5, 6, 7, 9, 10, and 11, five QTL regions were co-localized with previously reported salt-tolerant QTLs or genes (Park et al., 2010; Li et al., 2015; Nakhla et al., 2021), and the explained phenotypic variances (R^2) ranged from 5.41% for *qRVI60-5* to 10.14% for *qRGE60-11*. We noted that many RGE, RGI, RVI, and RMGT QTLs mapped to identical locations with the same peak SNPs.

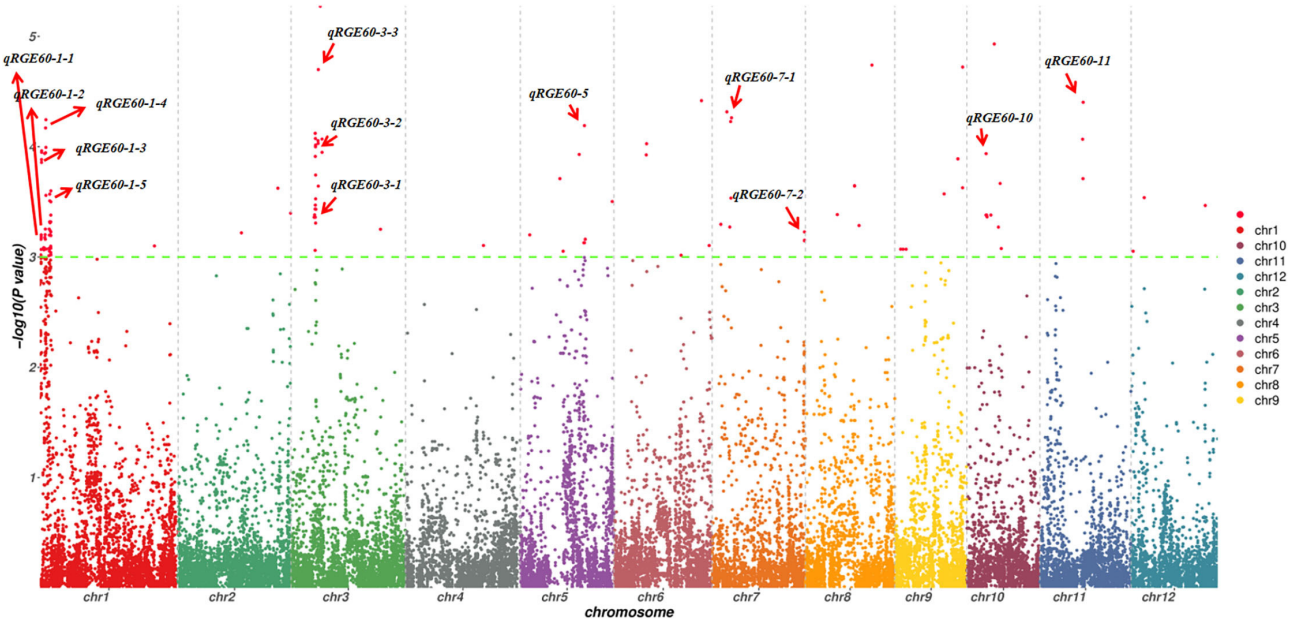


FIGURE 5 | Manhattan plots of GWAS for RGE under 60 mM NaCl. The red arrow indicates QTLs detected from six indices.

This supports our previous speculation that these four indices might share genetic pathways.

Candidate Gene Prediction and Expression Profiling

One QTL, *qRGE60-1-2* (*qRGI60-1-1*), which controls both RGE and RGI on chromosome 1, was identified by GWAS. According to the LD decay analysis, a 321-kb region was identified as the candidate region for *qRGE60-1-2* (**Figure 6A**), which contained 52 genes, including 25 functionally annotated genes, 18 expressed proteins with unknown functions, four retrotransposon proteins, three transposon proteins, and two hypothetical proteins (**Supplementary Table 3**). To decrease the number of candidate genes, genes categorized as expressed proteins, hypothetical proteins, retrotransposons, and transposons were discarded. Among the 25 functionally annotated genes, nine genes were found to exhibit significant differences in their expression levels between the control and salt treatments according to the Plant Public RNA-seq Database (|logFoldChange| $\geq$ 1.5), among which seven genes were upregulated under salt stress, one gene was downregulated, and the other gene had different expression trends under the different projects (**Supplementary Table 4**).

Another QTL, *qRGE60-3-1* (*qRVI60-3-1*), is also considered a candidate, which controls both RGE and RVI on chromosome 3. According to the LD decay analysis, a 489-kb region was identified as the candidate region (**Figure 6B**), and it contained 42 functionally annotated genes (**Supplementary Table 3**). According to the RNA-seq database (|logFoldChange| $\geq$ 1.5), the expression levels of seven genes were changed under salt stress, among which four genes were upregulated and three genes were downregulated (**Supplementary Table 4**).

To further decrease the number of candidate genes, the expression patterns of these 16 differentially expressed genes were identified by using Genevestigator. Fortunately, nine genes (e.g., *LOC_Os01g02930*, *LOC_Os01g02940*, *LOC_Os01g03330*, *LOC_Os01g03340*, *LOC_Os01g03360*, *LOC_Os03g13560*, *LOC_Os03g13840*, *LOC_Os03g14050*, and *LOC_Os03g14180*) were determined to be highly expressed in embryos (**Figure 7**; **Supplementary Figure 3**), which suggested that these nine genes are involved in rice seed germination under salt stress and could be candidates for STG.

Natural Allelic Variations of Candidate Genes Contribute to Salt Tolerance Germinability

As noted above, most *indica* rice varieties in our population had higher STG levels than *japonica* rice varieties, and we therefore investigated whether variations in the candidate gene alleles contributed to these differences. Haplotype analyses were performed using all non-synonymous SNPs within their ORFs by RiceVarMap v2.0. Finally, it was found that three genes had different haplotypes between *indica* and *japonica*. The *LOC_Os03g13560* gene harbored a total of four non-synonymous SNPs, namely, vg0307328618 (G/T), vg0307330319 (A/G), vg0307331295 (C/T), and vg0307334743 (C/T). Four distinct haplotypes, including two major haplotypes, namely, Hap I and Hap II, were identified based on the four non-synonymous SNPs in cultivated rice and exhibited large genetic differences between *indica* and *japonica* (**Figure 8A**). There are five non-synonymous SNPs in *LOC_Os03g13840*, namely, vg0307499320 (C/T), vg0307500449 (G/C), vg0307501727 (G/A), vg0307503107 (C/A), and vg0307503164 (C/T). Six major

TABLE 2 | Summary of the significant SNPs detected by GWAS and the overlapping QTLs reported previously.

QTL ID	Trait	Chr.	Peak SNPs	*P*-value	R^2	Previous QTL/genes	References
qRRL-100-1	RRL	1	201819	7.85E-04	0.070961901	*OsCOIN*	(Liu et al., 2007)
qRGE60-1-1	RGE	1	455562	6.31E-04	0.073714432		
qRGE60-1-2	RGE	1	1243681	8.25E-04	0.054838078		
qRGI60-1-1	RGI	1	1243681	1.97E-04	0.089011191		
qRGE60-1-3	RGE	1	1957926	5.70E-05	0.097587946	*OsHsp17.0*	(Zou et al., 2012)
qRGI60-1-2	RGI	1	1957926	4.53E-04	0.076102635	*OsHsp17.0*	(Zou et al., 2012)
qRMGT60-1-1	RMGT	1	1957926	7.99E-05	0.094079583	*OsHsp17.0*	(Zou et al., 2012)
qRGI60-1-3	RGI	1	3046380	6.16E-04	0.072982743	*OsDREB2A*	(Mallikarjuna et al., 2011)
qRVI60-1-1	RVI	1	3046380	7.17E-04	0.071413901	*OsDREB2A*	(Mallikarjuna et al., 2011)
qRGE60-1-4	RGE	1	3046380	2.70E-04	0.08148945	*OsDREB2A*	(Mallikarjuna et al., 2011)
qRMGT60-1-2	RMGT	1	3478635	3.18E-04	0.07974077		
qRVI60-1-2	RVI	1	3493294	3.16E-04	0.079821455		
qRGI60-1-4	RGI	1	3493294	4.43E-04	0.076340862		
qRGE60-1-5	RGE	1	3511964	6.34E-04	0.082964523		
qRRL60-1-1	RRL	1	3545015	2.15E-04	0.086175194		
qRMGT60-1-2	RMGT	1	3478635	3.18E-04	0.07974077		
qRRL60-1-2	RRL	1	3939680	2.16E-04	0.086169671	*OsPLDα1*	(Shen et al., 2011)
qRGI60-1-5	RGI	1	15977602	3.83E-04	0.077840812		
qRRL60-1-3	RRL	1	42421352	1.45E-04	0.090155653	*OrbHLH001*	(Chen et al., 2013)
qRRL60-2-1	RRL	2	4380604	6.69E-04	0.075583000	*qGR-3d2*	(Nakhla et al., 2021)
qRRL60-2-2	RRL	2	18813593	5.33E-04	0.077148185		
qRSL100-2	RSL	2	19661213	1.88E-04	0.086493361		
qRGE60-3-1	RGE	3	7668652	7.57E-05	0.094665597		
qRVI60-3-1	RVI	3	7668652	5.28E-04	0.074581176		
qRGE60-3-2	RGE	3	8660029	2.00E-05	0.108593017		
qRGE60-3-3	RGE	3	9852079	5.32E-06	0.122680821	*OsMAPK5*	(Xiong and Yang, 2003)
qRGE100-3	RGE	3	9852079	4.98E-04	0.075150154	*OsMAPK5*	(Xiong and Yang, 2003)
qRVI60-3-2	RVI	3	9852079	2.83E-04	0.080963707	*OsMAPK5*	(Xiong and Yang, 2003)
qRMGT100-3	RMGT	3	33227077	9.80E-04	0.085821366		
qRVI60-5	RVI	5	14178828	8.95E-04	0.054078023		
qRGE60-5	RGE	5	20392181	6.45E-05	0.079537025		
qRRL60-6	RRL	6	5117960	2.28E-04	0.085628982	*qGR-7d6*	(Nakhla et al., 2021)
qRGE60-7-1	RGE	7	5888791	5.92E-05	0.097180181		
qRGE60-7-2	RGE	7	29343026	5.93E-04	0.077278782		
qRRL60-9	RRL	9	16654472	1.73E-04	0.088530434	*OsDSG1*	(Park et al., 2010)
qRGI60-9	RGI	9	17191317	9.93E-05	0.100093083	*OsbHLH120*	(Li et al., 2015)
qRVI60-9	RVI	9	17191317	3.30E-04	0.084884769	*OsbHLH120*	(Li et al., 2015)
qRVI100-9	RVI	9	17191317	8.91E-05	0.095534518	*OsbHLH120*	(Li et al., 2015)
qRRL60-10-1	RRL	10	4285687	7.40E-05	0.097070419		
qRRL100-10	RRL	10	4665777	3.95E-04	0.079672322		
qRGE60-10	RGE	10	6107128	1.16E-04	0.091425897		
qRRL60-10-2	RRL	10	7248029	2.45E-04	0.084358531		
qRGE60-11	RGE	11	13756114	3.96E-05	0.101443433		
qRGE100-11	RGE	11	19317926	3.80E-04	0.079063947		

haplotypes were observed at *LOC_Os03g13840*, which encodes a senescence-associated protein. Hap I was present in *japonica*, and Hap II, III, and IV were present primarily in *indica* (**Figure 8B**). Interestingly, Hap II was carried in ∼16% of *japonica* accessions, which indicated that elite *indica* alleles had been introgressed into *japonica* accessions through breeding to the extent that a small percentage of the *japonica* accessions (such as JX172 and JX224) included in our population had high STG levels. A total of three non-synonymous SNPs [e.g., vg0307697206 (C/G), vg0307697461 (G/T), and vg0307697508 (C/T)] in *LOC_Os03g14180* identified three haplotypes (**Figure 8C**). The vast majority of *indica* accessions were Hap III, while the *japonica*

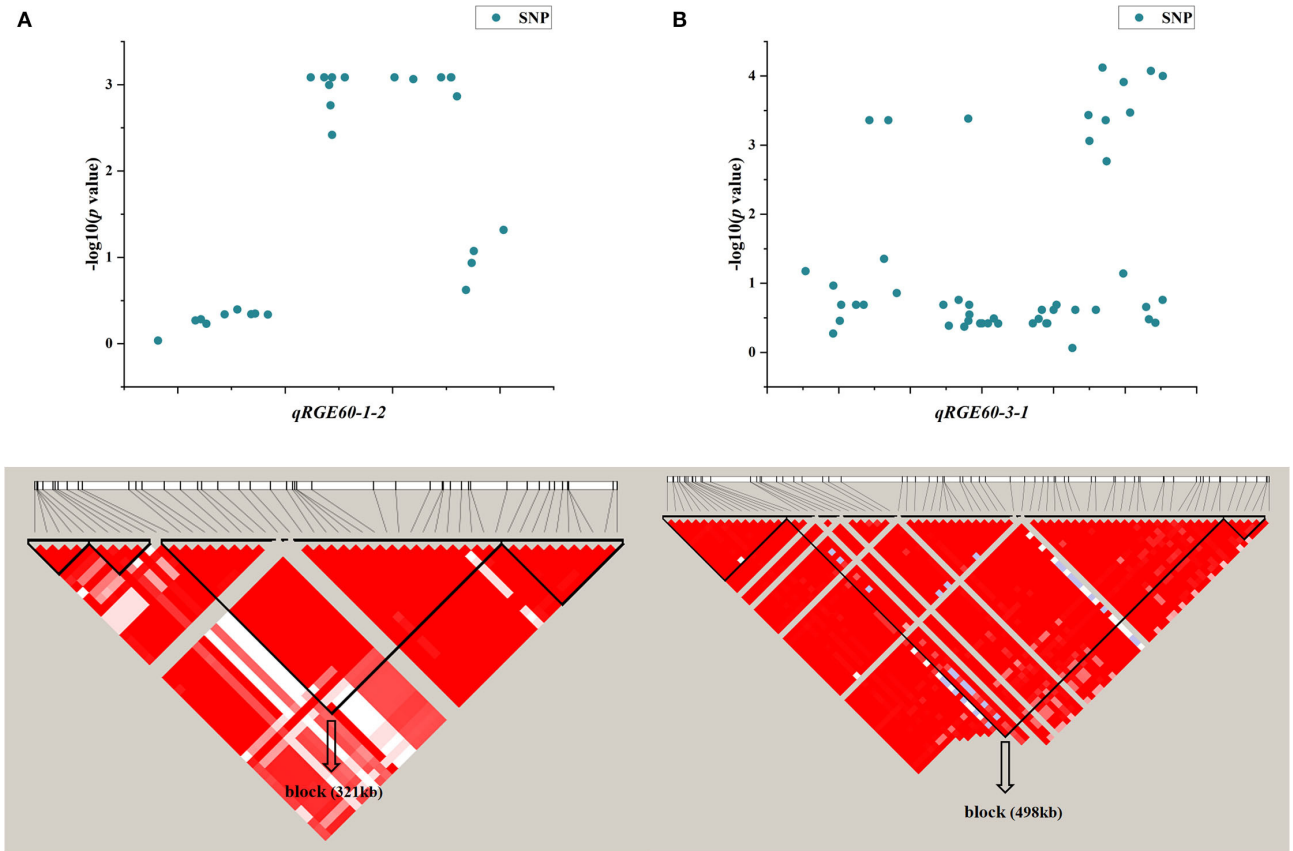


FIGURE 6 | LD block of candidate region estimation of two major QTLs, namely, *qRGE60-1-2* **(A)** and *qRGE60-3-1* **(B)**.

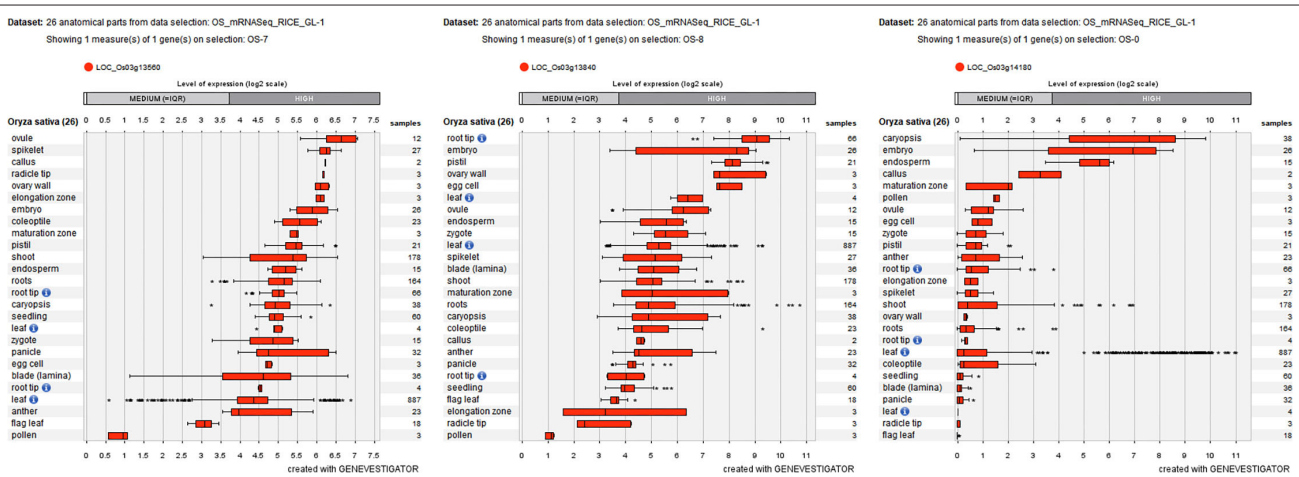


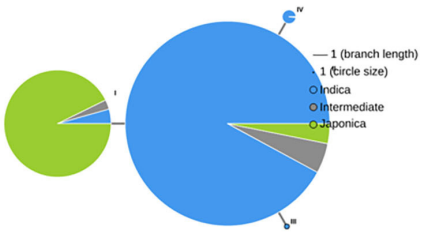
FIGURE 7 | Three differentially expressed genes (*LOC_Os03g13560*, *LOC_Os03g13840*, and *LOC_Os03g14180*) are highly expressed in embryos labeled by Geneinvestigator. The * represent discrete point.

accessions were Hap I or Hap II. Taken together, these results indicate that the natural allelic variations in these three candidate genes appeared to be associated with the different STG levels observed among the 211 rice accessions.

To further verify whether the three genes respond to salt stress and have *indica–japonica* specificity in the individuals examined in our population, quantitative real-time PCR (qRT-PCR) was performed with two salt-sensitive varieties (e.g.,

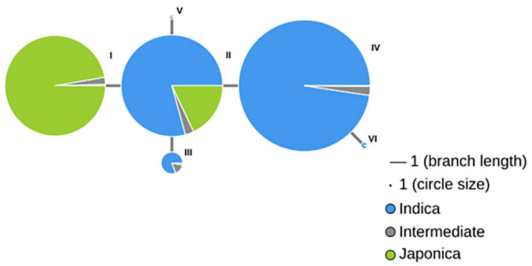
A

Haplotype	SNP physical location			
	7328618	7330319	7331295	7334743
I	G	G	C	C
II	G	A	C	C
III	G	A	T	C
IV	T	A	C	C



B

Haplotype	SNP physical location				
	7499320	7500449	7501727	7503107	7503164
I	C	G	A	C	C
II	C	G	G	C	C
III	C	G	G	C	T
IV	T	G	G	C	C
V	C	C	G	C	C
VI	T	G	G	A	C



C

Haplotype	SNP physical location		
	7697206	7697461	7697508
I	G	T	T
II	G	G	T
III	C	G	C

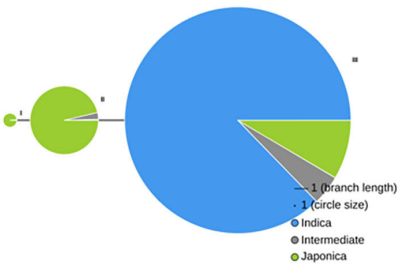


FIGURE 8 | Haplotypes of *LOC_Os03g13560* **(A)**, *LOC_Os03g13840* **(B)**, and *LOC_Os03g14180* **(C)** are associated with STG in rice.

S1 and S2) and two salt-tolerant varieties (e.g., T1 and T2) under control and 60 mM NaCl conditions. The results showed that these three genes were differentially expressed under salt stress (**Figure 9**). For *LOC_Os03g13560*, the expressions in salt-sensitive cultivars were downregulated, while the expressions in salt-tolerant cultivars were upregulated under salt stress (**Figure 9A**). In addition, the downregulation times of S1 (*indica*) were much higher than those of S2 (*japonica*) in the two salt-sensitive varieties, and the upregulation times of T2 (*japonica*) were much higher than those of T1 (*indica*) in the two salt-tolerant varieties. The *LOC_Os03g13840* gene was upregulated under salt stress in both salt-sensitive and salt-tolerant varieties, and the upregulated *LOC_Os03g13840* expressions in the two *indica* rice varieties were higher than those in *japonica* rice (**Figure 9B**). The *LOC_Os03g14180* gene responded to salt stress in all four rice varieties and was upregulated under 60 mM NaCl conditions (**Figure 9C**). Similarly, the changes in the expression levels of *LOC_Os03g14180* were *indica–japonica* specific. After salt stress, the change times of S2 were higher than those of S1, and the change times of T1 were higher than those of T2.

DISCUSSION

Direct seeding of rice has become popular in many Asian countries because of its time-saving, labor-saving, and cost-effective advantages. However, to develop high-yield and high-quality rice varieties for use in the direct seeding mode, rice seeds must have strong germination abilities under various abiotic stresses and strong growth abilities after abiotic stresses. Therefore, cultivating rice varieties with high STG has become an inevitable trend. In our study, six indices containing four indices (e.g., RGE, RGI, RVI, and RMGT) associated with seed germination and two indices (e.g., RSL and RRL) associated with seed growth capacity after salt stress were used to evaluate STG levels. These indices have been used successfully in previous GWAS studies on salt tolerance at the rice germination stage (Shi et al., 2017; Yu et al., 2018). Our study showed that additions of salt solutions inhibited seed germination compared with the control treatment; that is, the GE, GI, VI, and MGT values of seeds decreased, and the inhibition degrees also increased with increasing salt solution concentrations (**Table 1**). Despite salt stress, nearly all rice accessions germinated successfully in the

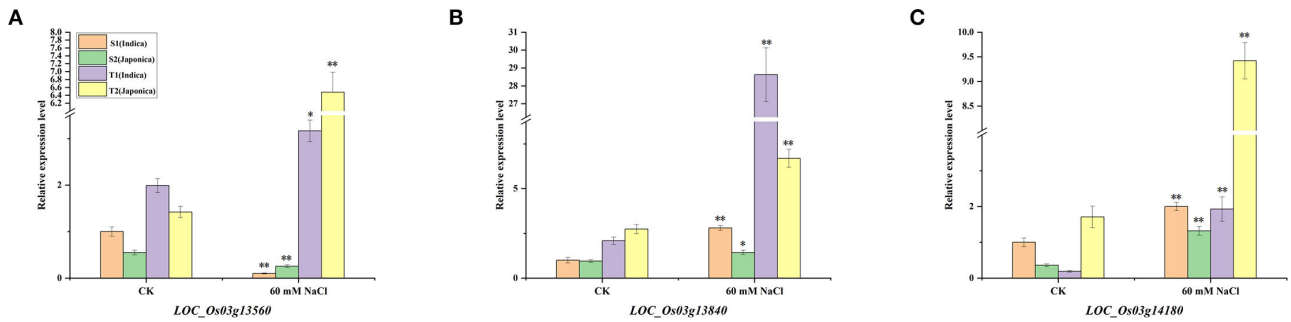


FIGURE 9 | Expression patterns of *LOC_Os03g13560* **(A)**, *LOC_Os03g13840* **(B)**, and *LOC_Os03g14180* **(C)** under normal growth conditions and salt stress conditions. *Indicates significance at the 5% level, **Indicates significance at the 1% level.

following days, but this does not mean that salt stress prolonged only the germination time and did not have other effects because after 7 days of salt stress, we found that the SL and RL values of rice were reduced compared to the control group, which suggested that salinity can influence the germination quality of seeds.

Different conclusions have been reached in different studies on the salt tolerance of different rice subgroups in the germination stage. According to Shi et al. (2017), *japonica* varieties had higher tolerance to salt stress at the germination stage than did the other subgroups, which was determined by measuring the SSIs of VI and GR at 10 days. However, Yu et al. (2018) and Islam et al. (2022) suggested that the salt tolerance levels during the seed germination stage are not well-correlated with rice subgroups. In our study, a scoring rule was developed based on six indices under two different salt solution concentrations to determine the STG level of each rice accession, and the results showed that the STG of *indica* rice was generally significantly higher than that of *japonica* rice. No significant differences in RMGT values were observed between the two subgroups at either salt solution concentration, but the RSL, RGE, RGI, and RVI values of *indica* for the two salt solution concentrations were significantly greater than those of *japonica*. There were no significant differences between the two subgroups under the 60 mM condition, but the RRL values of *japonica* were greater than those of *indica* under the 100 mM condition, which was probably because with the increased salt solution concentration, salt-sensitive rice varieties needed to extend their roots longer to increase their water uptake areas to survive.

The purpose of using a variety of indices to evaluate rice STG is to obtain a more comprehensive understanding of the genetic mechanisms of rice varieties subjected to salt stress in the germination stage. The correlations among the six indices were analyzed, and the results showed that the RGE, RGI, and RVI values were significantly positively correlated with each other and were negatively correlated with RMGT regardless of whether concentrations of 60 or 100 mM were used, which may indicate that the four indices are regulated by a consistent genetic pathway. This inference was confirmed by the subsequent GWAS analysis.

A total of 33 distinct QTL regions were identified by GWAS using six indices and two concentrations, among which 10 QTLs were found to be co-localized in the same or overlapping regions of previously reported salt stress QTLs/genes. The QTL *qRRL-100-1* overlapped with *OsCOIN*, a gene encoding ring zinc finger protein that is located in the nucleus and plasma membrane, and *OsCOIN* is expressed in all organs of rice and is strongly induced by low temperatures, ABA, salt and drought. Overexpression of *OsCOIN* upregulated the expression of *OsP5CS* and increased intracellular proline contents, which ultimately significantly enhanced the resistance of transgenic rice to cold, salt, and drought (Liu et al., 2007). A small heat shock protein, *OsHSP17.0*, appeared in a QTL region that was repeatedly detected (e.g., *qRGE60-1-3*/*qRGI60-1-2*/*qRMGT60-1-1*) in our study. The germination capacities of *OsHsp17.0*-overexpressing transgenic lines in NaCl was stronger than that of wild-type plants, and phenotypic analysis showed that transgenic rice lines were more tolerant to salt stress than WT rice lines (Zou et al., 2012). An AP2/EREBP transcription factor, *OsDREB2A*, was located in a candidate region that contains three co-localized QTLs (e.g., *qRGI60-1-3*, *qRVI60-1-1*, and *qRGE60-1-4*). *OsDREB2A* expression was induced by high salinity stress (Dubouzet et al., 2003), and its overexpressed lines were significantly tolerant to osmotic stress, drought stress, salt stress, and dehydration stress, and their growth performance was enhanced (Cui et al., 2011; Mallikarjuna et al., 2011). Five RRL-related QTLs (e.g., *qRRL60-1-2*, *qRRL60-1-3*, *qRRL60-2-1*, *qRRL60-6*, and *qRRL60-9*) that were evaluated in our study were co-localized with previously reported salt stress QTLs/genes (e.g., *OsPLDα1*, *OrbHLH001*, *qGR-3d2*, *qGR-7d6*, and *OsDSG1*). *OsPLDα1* is involved in the salt tolerance of rice by regulating the activity and expression of H+-ATPase (Shen et al., 2011). Overexpression of *OrbHLH001*, a putative helix-loop-helix transcription factor, causes increased expression of AKT1 and maintains ionic balance under salt stress in rice (Chen et al., 2013). *qGR-3d2* and *qGR-7d6* are two QTLs associated with germination rates that were detected in the germination stage (Nakhla et al., 2021). *OsDSG1* is a gene that affects seed germination, and loss of this gene results in delayed seed germination, shorter plants, and greater tolerance to high salt and drought stresses (Park et al., 2010). A QTL containing three co-located QTLs (e.g., *qRGE60-3-3*, *qRGE100-3*, and *qRVI60-3-2*) on chromosome 3 explained the largest phenotypic variations in our study, and this QTL region contained a characterized salt-tolerant gene, *OsMAPK5*, which

is a mitogen-activated protein that can positively regulate the tolerance of rice to drought, salt, and cold stresses (Xiong and Yang, 2003). Similarly, a QTL (e.g., *qRGI60-9*/*qRVI60-9*/*qRVI100-9*) is present in 17.19 Mb of chromosome 9 that explains the large phenotypic variations. This region contains a basic helix-loop-helix transcription factor called *OsbHLH120*, and the expression of *OsbHLH120* in IL392 (an introgression line) roots was higher than that of Yuefu (parent) under PEG and salt stresses (Li et al., 2015). In addition, 23 new loci were identified in this study, among which three loci contained more than one QTL. These new QTLs detected by multiple indices will provide important genetic resources for identifying candidate salt genes.

Hence, the exact candidate intervals were determined by LD blocks for two new QTLs, *qRGE60-1-2* and *qRGE60-3-1*, with high explained phenotypic variations, which were detected in multiple indices. According to a previous study (Neang et al., 2020), only 67 functionally annotated genes were retained in these two QTLs. We used multiple databases to determine the expression patterns of these functional annotation genes and whether they responded to abiotic stress. Finally, nine genes that respond to salt stress and are highly expressed in embryos were selected as candidate genes. Interestingly, the haplotype analysis suggested that the haplotypes of three candidate genes (e.g., *LOC_Os03g13560*, *LOC_Os03g13840*, and *LOC_Os03g14180*) exhibited some *indica–japonica* specificity, which was consistent with our phenotypic results that the STG levels of *indica* and *japonica* rice were significantly different. *LOC_Os03g13560* encodes a hydroxyproline-rich glycoprotein family protein. A previous study revealed that hydroxyproline-rich glycoprotein family proteins may be involved in the chilling response of soybean (Xu et al., 2016). *LOC_Os03g13840* encodes senescence-associated protein (SAP). A gene ontology classification was used to predict its function, and the results showed that there were two major GO accessions for this candidate gene, namely, GO:0009628 (response to abiotic stimulus) and GO:0006950 (response to stress). Ubaidillah et al. (2013) revealed that *OsSAP* was both highly and rapidly expressed in response to drought stress. *LOC_Os03g14180* is a rice heat shock protein gene named *OsHSP26.7*. The transcript levels of *OsHSP26.7* were also high in the imbibed seed embryos but were very low in other organs and were significantly induced by heat treatment and high salt stress (Zou et al., 2009). In conclusion, the identification of candidate genes and their elite haplotypes for QTLs provides important information to identify the causal genes and facilitate further validation and application of the identified QTLs in trait improvements.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://ngdc.cncb.ac.cn/gvm/, GVM000333.

AUTHOR CONTRIBUTIONS

CLi designed experiments, analyzed data, and wrote the manuscript. CLi, CLu, and BZ performed all the phenotypic evaluations. MY, GW, PW, QC, YW, QZ, SH, TH, and HH participated in performing experiments. JB conceived and supervised the experiments. All authors have read and agreed to the published version of the manuscript.

FUNDING

This study was supported by a grant (32160485) from the National Natural Science Foundation of China, a grant (YC2021-S341) from the Innovation Fund Designated for Graduate Students of Jiangxi Province, and College Students Innovative Entrepreneurial Training Plan Program (Grant No. 202210410002). These funding institutions provided financial support for material collection, high-throughput sequencing, and data analysis in this study.

ACKNOWLEDGMENTS

We thank the referees for their critical comments on this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.934515/full#supplementary-material

Supplementary Figure 1 | Box plot for comparison of six indices between *indica* and *japonica* subgroups under 100 mM NaCl. The yellow box represents *indica*, the green box represents *japonica*, the black dot in the box represents the median, and the value range of the box is 25–75%, **Indicates significance at the 1% level.

Supplementary Figure 2 | Manhattan plots of GWAS for RGI **(A)**, RVI **(B)**, RMGT **(C)**, RRL **(D)**, under 60 mM NaCl and RGE **(E)**, RVI **(F)**, RMGT **(G)**, RSL **(H)**, and RRL **(I)** under 100 mM NaCl. The red arrow indicates QTLs detected from six indices.

Supplementary Figure 3 | Five differentially expressed genes **(A–E)** for *qRGE60-1-2* and a gene *LOC_Os03g14050* **(F)** for *qRGE60-3-1* are highly expressed in embryos labeled by Geneinvestigator.

Supplementary Table 1 | Phenotypic data and information of the 211 rice varieties.

Supplementary Table 2 | Criteria for scoring the phenotype of each individual in the population.

Supplementary Table 3 | The candidate genes for *qRGE60-1-2* and *qRGE60-3-1*.

Supplementary Table 4 | The differentially expressed genes for *qRGE-1-2* and *qRGE-3-1* from the RNA-seq database.

REFERENCES

Barrett, J. C., Fry, B., Maller, J., and Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21, 263–265. doi: 10.1093/bioinformatics/bth457

Bian, J. M., Jiang, L., Liu, L. L., Wei, X. J., Xiao, Y. H., Zhang, L. J., et al. (2010). Construction of a new set of rice chromosome segment substitution lines and identification of grain weight and related traits QTLs. *Breed Sci.* 60, 305–313 doi: 10.1270/jsbbs.60.305

Chen, H. C., Cheng, W. H., Hong, C. Y., Chang, Y. S., and Chang, M. C. (2018). The transcription factor *OsbHLH035* mediates seed germination and enables seedling recovery from salt stress through ABA-dependent and ABA-independent pathways, respectively. *Rice* 11:50. doi: 10.1186/s12284-018-0244-z

Chen, Y., Li, F., Ma, Y., Chong, K., and Xu, Y. (2013). Overexpression of *OrbHLH001*, a putative helix–loop–helix transcription factor, causes increased expression of AKT1 and maintains ionic balance under salt stress in rice. *J. Plant Physiol.* 170, 93–100. doi: 10.1016/j.jplph.2012.08.019

Cheng, J., He, Y., Yang, B., Lai, Y., Wang, Z., and Zhang, H. (2015). Association mapping of seed germination and seedling growth at three conditions in *indica* rice (*Oryza sativa* L.). *Euphytica* 206, 103–115. doi: 10.1007/s10681-015-1477-1

Cui, M., Zhang, W., Zhang, Q., Xu, Z., Zhu, Z., Duan, F., et al. (2011). Induced over-expression of the transcription factor *OsDREB2A* improves drought tolerance in rice. *Plant Physiol. Biochem.* 49, 1384–1391. doi: 10.1016/j.plaphy.2011.09.012

Cui, Y., Zhang, F., and Zhou, Y. (2018). The application of Multi-Locus GWAS for the detection of Salt-Tolerance loci in rice. *Front. Plant Sci.* 9:1464. doi: 10.3389/fpls.2018.01464

Dubouzet, J. G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E. G., Miura, S., et al. (2003). *OsDREB* genes in rice, *Oryza sativa* L., Encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J.* 33, 751–763. doi: 10.1046/j.1365-313X.2003.01661.x

Gomes-Filho, E., Lima, C. R., Costa, J. H., Da, S. A., Da, G. S. L. M., De, L. C., et al. (2008). Cowpea ribonuclease: properties and effect of NaCl-salinity on its activation during seed germination and seedling establishment. *Plant Cell Rep.* 27, 147–157. doi: 10.1007/s00299-007-0433-5

He, Y., Yang, B., He, Y., Zhan, C., Cheng, Y., Zhang, J., et al. (2019). A quantitative trait locus, *qSE3*, promotes seed germination and seedling establishment under salinity stress in rice. *Plant J.* 97, 1089–1104. doi: 10.1111/tpj.14181

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., et al. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* 44, 32–39 doi: 10.1038/ng.1018

Islam, M. R., Naveed, S. A., Zhang, Y., Li, Z., Zhao, X., Fiaz, S., et al. (2022). Identification of candidate genes for salinity and anaerobic tolerance at the germination stage in rice by genome wide association analyses. *Front. Genet.* 13:822516. doi: 10.3389/fgene.2022.822516

Khan, M., and Weber, D. (2006). *Ecophysiology of High Salinity Tolerant Plants*. New York, NY: Springer Science & Business Media. doi: 10.1007/1-4020-4018-0

Koyro, H. (2002). *Ultrastructural Effects of Salinity in Higher Plants. In Salinity, Environment-Plants-Molecules*. New York, NY: Springer.

Li, C., Liu, J., Bian, J., Jin, T., Zou, B., Liu, S., et al. (2021). Identification of cold tolerance QTLs at the bud burst stage in 211 rice landraces by GWAS. *BMC Plant Biol.* 21:7. doi: 10.1186/s12870-021-03317-7

Li, J., Han, Y., Liu, L., Chen, Y., Du, Y., Zhang, J., et al. (2015). *qRT9*, a quantitative trait locus controlling root thickness and root length in upland rice. *J. Exp. Bot.* 66, 2723–2732. doi: 10.1093/jxb/erv076

Liu, K., Wang, L., Xu, Y., Chen, N., Ma, Q., Li, F., et al. (2007). Overexpression of OsCOIN, a putative cold inducible zinc finger protein, increased tolerance to chilling, salt and drought, and enhanced proline level in rice. *Planta* 226, 1007–1016. doi: 10.1007/s00425-007-0548-5

Luo, X., Deng, H., Wang, P., Zhang, X., Li, C., Tan, J., et al. (2020). Genetic analysis of germinating ability under alkaline and neutral salt stress by a high-density bin genetic map in rice. *Euphytica* 216, 107. doi: 10.1007/s10681-020-02623-9

Mallikarjuna, G., Mallikarjuna, K., Reddy, M. K., and Kaul, T. (2011). Expression of *OsDREB2A* transcription factor confers enhanced dehydration and salt stress tolerance in rice (*Oryza sativa* L.). *Biotechnol. Lett.* 33, 1689–1697. doi: 10.1007/s10529-011-0620-x

Mardani, Z., Rabiei, B., Sabouri, H., and Sabouri, A. (2014). Identification of molecular markers linked to salt-tolerant genes at germination stage of rice. *Plant Breed* 133, 196–202. doi: 10.1111/pbr.12136

Nakhla, W. R., Sun, W., Fan, K., Yang, K., Zhang, C., and Yu, B. (2021). Identification of QTLs for salt tolerance at the germination and seedling stages in rice. *Plants* 10:30428. doi: 10.3390/plants10030428

Naveed, S. A., Zhang, F., Zhang, J., Zheng, T. Q., Meng, L. J., Pang, Y. L., et al. (2018). Identification of QTN and candidate genes for salinity tolerance at the germination and seedling stages in rice by Genome-Wide association analyses. *Sci. Rep* 8:6505. doi: 10.1038/s41598-018-24946-3

Neang, S., de Ocampo, M., Egdane, J. A., Platten, J. D., Ismail, A. M., Seki, M., et al. (2020). A GWAS approach to find SNPs associated with salt removal in rice leaf sheath. *Ann. Bot. London* 126, 1193–1202. doi: 10.1093/aob/mcaa139

Park, G., Park, J., Yoon, J., Yu, S., and An, G. (2010). A RING finger E3 ligase gene, *Oryza sativa* Delayed Seed Germination 1 (OsDSG1), controls seed germination and stress responses in rice. *Plant Mol. Biol.* 74, 467–478. doi: 10.1007/s11103-010-9687-3

Shen, P., Wang, R., Jing, W., and Zhang, W. (2011). Rice phospholipase Dalpha is involved in salt tolerance by the mediation of H^+-ATPase activity and transcription. *J. Integr. Plant Biol.* 53, 289–299. doi: 10.1111/j.1744-7909.2010.01021.x

Shi, Y., Gao, L., Wu, Z., Zhang, X., Wang, M., Zhang, C., et al. (2017). Genome-wide association study of salt tolerance at the seed germination stage in rice. *BMC Plant Biol.* 17:92. doi: 10.1186/s12870-017-1044-0

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20260–20264. doi: 10.1073/pnas.1116437108

Ubaidillah, M., Kim, K. A., Kim, Y. H., Lee, I-J., Yun, B. W., Kim, D. H., et al. (2013). Identification of a drought-induced rice gene, OsSAP, that suppresses bax-induced cell death in yeast. *Mol. Biol. Rep.* 40, 6113–6121. doi: 10.1007/s11033-013-2723-z

Wang, Z. F., Wang, J. F., Bao, Y. M., Wu, Y. Y., and Zhang, H. S. (2011). Quantitative trait loci controlling rice seed germination under salt stress. *Euphytica* 178, 297–307 doi: 10.1007/s10681-010-0287-8

Xiong, L., and Yang, Y. (2003). Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acid–inducible Mitogen-Activated protein kinase. *Plant Cell* 15, 745–759. doi: 10.1105/tpc.008714

Xu, S., Liu, N., Mao, W., Hu, Q., Wang, G., and Gong, Y. (2016). Identification of chilling-responsive microRNAs and their targets in vegetable soybean (*Glycine max* L.). *Sci. Rep.* 6:26619. doi: 10.1038/srep26619

Yu, J., Zhao, W., Tong, W., He, Q., Yoon, M. Y., Li, F. P., et al. (2018). A Genome-Wide association study reveals candidate genes related to salt tolerance in rice (*Oryza sativa*) at the germination stage. *Int. J. Mol. Sci.* 19:103145. doi: 10.3390/ijms19103145

Yu, Y., Zhang, H., Long, Y., Shu, Y., and Zhai, J. (2022). Plant Public RNA-seq Database: a comprehensive online database for expression analysis of ∼45,000 plant public RNA-Seq libraries. *Plant Biotechnol. J.* 2022:pbi.13798. doi: 10.1111/pbi.13798

Zeng, P., Zhu, P., Qian, L., Qian, X., Mi, Y., Li, Z., et al. (2021). Identification and fine mapping of *qGR6.2*, a novel locus controlling rice seed germination under salt stress. *BMC Plant Biol.* 21:36. doi: 10.1186/s12870-020-02820-7

Zhang, X., Long, Y., Huang, J., and Xia, J. (2020). *OsNAC45* is involved in ABA response and salt tolerance in rice. *Rice* 13:79. doi: 10.1186/s12284-020-00440-1

Zhao, H., Yao, W., Ouyang, Y., Yang, W., Wang, G., Lian, X., et al. (2015). RiceVarMap: a comprehensive database of rice genomic variations. *Nucl. Acids Res.* 43, D1018–D1022. doi: 10.1093/nar/gku894

Zheng, H., Liu, B., Zhao, H., Wang, J., Liu, H., Sun, J., et al. (2014). Identification of QTLs for salt tolerance at the germination and early seedling stage using linkage and association analysis in japonica rice. *Chin. J. Rice Sci.* 28, 358–366. doi: 10.3969/j.issn.1001-7216.2014.04.004

Zou, J., Liu, A., Chen, X., Zhou, X., Gao, G., Wang, W., et al. (2009). Expression analysis of nine rice heat shock protein genes under abiotic stresses and ABA treatment. *J. Plant Physiol.* 166, 851–861. doi: 10.1016/j.jplph.2008.11.007

Zou, J., Liu, C., Liu, A., Zou, D., and Chen, X. (2012). Overexpression of *OsHsp17.0* and *OsHsp23.7* enhances drought and salt tolerance in rice. *J. Plant Physiol.* 169, 628–635. doi: 10.1016/j.jplph.2011.12.014

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fpls.2022.985559

OPEN ACCESS

EDITED BY
Xueyong Li,
Institute of Crop Sciences (CAAS), China

REVIEWED BY
Guiquan Zhang,
South China Agricultural University, China
Sung-Ryul Kim,
International Rice Research Institute (IRRI), Philippines

*CORRESPONDENCE
Sang-Nag Ahn
ahnsn@cnu.ac.kr

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Bioinformatics, a section of the journal Frontiers in Plant Science

RECEIVED 04 July 2022
ACCEPTED 29 August 2022
PUBLISHED 20 September 2022

CITATION
Kim SH, Shim K-C, Lee H-S, Jeon Y-A, Adeva C, Luong NH and Ahn S-N (2022) Brassinosteroid biosynthesis gene *OsD2* is associated with low-temperature germinability in rice. *Front. Plant Sci.* 13:985559. doi: 10.3389/fpls.2022.985559

Brassinosteroid biosynthesis gene *OsD2* is associated with low-temperature germinability in rice

Sun Ha Kim[1†], Kyu-Chan Shim[1†], Hyun-Sook Lee[2], Yun-A Jeon[1], Cheryl Adeva[1], Ngoc Ha Luong[1] and Sang-Nag Ahn[1*]

[1]Department of Agronomy, College of Agriculture and Life Sciences, Chungnam National University, Daejeon, South Korea, [2]Crop Breeding Division, National Institute of Crop Science, Wanju-Gun, South Korea

In rice, low-temperature germinability (LTG) is essential for stable stand establishment using the direct seeding method in temperate and high-altitude areas. Previously, we reported that the quantitative trait locus *qLTG1* is associated with LTG. *qLTG1* is also associated with tolerance to several abiotic stresses, such as salt and osmotic conditions. In this study, map-based cloning and sequence analysis indicated that *qLTG1* is allelic to *DWARF2* (*OsD2*), which encodes cytochrome P450 D2 (LOC_Os01g10040) involved in brassinosteroid (BR) biosynthesis. Sequence comparison of the two parental lines, Hwaseong and *Oryza rufipogon* (IRGC 105491), revealed five single nucleotide polymorphisms (SNPs) in the coding region. Three of these SNPs led to missense mutations in *OsD2*, whereas the other two SNPs were synonymous. We evaluated two T-DNA insertion mutants, *viz.*, overexpression (*OsD2*-OE) and knockdown (*OsD2*-KD) mutants of *OsD2*, with the Dongjin genetic background. *OsD2*-KD plants showed a decrease in LTG and grain size. In contrast, *OsD2*-OE plants showed an increase in grain size and LTG. We also examined the expression levels of several BR signaling and biosynthetic genes using the T-DNA insertion mutants. Gene expression analysis and BR application experiments demonstrated that BR enhanced the seed germination rate under low-temperature conditions. These results suggest that *OsD2* is associated with the regulation of LTG and improving grain size. Thus, *OsD2* may be a suitable target for rice breeding programs to improve rice yield and LTG.

KEYWORDS

rice, *Oryza sativa* L., *Oryza rufipogon*, low-temperature germination, *OsD2*, brassinosteroids, cytochrome P450

Introduction

Rice is one of the most important cereal crops in the world, and its production needs to be increased to meet the demands of an increasing population. However, resource limitations, such as decreased agricultural land, water supply, and poor soils, are major challenges in many rice producing countries (Kumar and Ladha, 2011).

Strong seedling vigor under low-temperature conditions is an important objective of rice breeding programs using direct seeding cultivation methods. The direct seeding method is widely used because it requires less labor, water, and energy than transplanted cultivation systems (Hyun et al., 2015). However, low-temperature stress retards seed germination in direct seeding cultivation, especially in temperate regions, tropical and subtropical areas at high altitudes, and areas that use cold irrigation water (Fujino et al., 2008). Germinability and early seedling growth are major components of seedling vigor, and improving germinability under low-temperature conditions can lead to improved seedling vigor (Iwata et al., 2010). A wide range of phenotypic variations of low-temperature germinability (LTG) has been observed in rice accessions (Wang et al., 2018).

It has been demonstrated that LTG is controlled by multiple genes, and over 30 LTG quantitative trait loci (QTL) have been identified using biparental mapping populations (Miura et al., 2001; Fujino et al., 2004; Chen et al., 2006; Li et al., 2013b; Jiang et al., 2017). Among these QTLs, the major QTL, *qLTG3-1*, was cloned using a population derived from a cross between Italica Livorno and Hayamasari (Fujino et al., 2008). The expression pattern of *qLTG3-1* suggested that the qLTG3-1 protein may function to weaken the tissues covering the embryo during germination. *OsSAP16* (*stress-associated protein 16*), which encodes a zinc finger domain protein, was cloned from a genome-wide association study (GWAS; Wang et al., 2018). Loss of function of *OsSAP16* led to decreased LTG levels, and the expression level of *OsSAP16* was correlated with LTG in the rice core collection. Although many other QTLs for LTG have been identified, knowledge of the molecular functions of QTL and gene characterization is still poor.

OsD2 (LOC_Os01g10040) has been previously cloned as *ebisu dwarf 2* (*d2*) gene (Hong et al., 2003; Sakamoto et al., 2012; Li et al., 2013a; Fang et al., 2016). This gene encodes cytochrome P450 and functions in the brassinosteroid (BR) biosynthesis pathway. Two mutants (*d2-1* and *d2-2*) with single nucleotide polymorphism (SNP) in the *OsD2* exon region have been characterized, and these two *d2* mutants showed reduced plant height with erect leaves (Hong et al., 2003). Many *OsD2* mutants have been reported, but most of the mutant alleles cause unfavorable phenotypes, such as severe dwarf and small grain size, leading to a decrease in yield. The *OsD2* mutants *chromosome segment deleted dwarf 1* (*csdd1*) and *small grain 11* (*smg11*) showed severe dwarf phenotype and small grain with mild dwarf phenotype, respectively, in a *japonica* background (Li et al., 2013a; Fang et al., 2016). An investigation of two independent T-DNA insertion mutants of *OsD2* showed that T-DNA homozygous insertion led to the loss of function of *OsD2*, resulting in a severe dwarf phenotype (Li et al., 2013a).

Brassinosteroids are plant hormones that have pleiotropic effects; they regulate seed germination, plant growth and development, rhizogenesis, senescence, and leaf abscission (Anuradha and Rao, 2001). The positive effects of BR under stress conditions have been reported in monocot and dicot plant species. BR application ameliorated osmotic stress in germinating seeds and promoted seedling growth in three sorghum varieties (Vardhini and Rao, 2003). Similarly, exogenous BR treatment improved the germination rate of cucumber seeds under salinity stress (Wang et al., 2011). Treatment with 24-epibrassinolide and 28-homobrassinolide, increased the germinability of *indica* rice IR64 under saline conditions (150 mM NaCl; Anuradha and Rao, 2001). Although stress tolerance at the germination stage has been reported in many plant species, studies on the effect of BR on LTG are limited.

Previously, we identified five QTLs for LTG using a population derived from the interspecific cross between the Korean *japonica* cultivar "Hwaseong" and *Oryza rufipogon* (IRGC 105491; Nguyen et al., 2012). *Oryza rufipogon* alleles of the five QTLs increased germination rate under low-temperature conditions. Among these QTLs, *O. rufipogon* alleles at *qLTG1* and *qLTG3* additively increased LTG (Shim et al., 2020). *qLTG1* was narrowed to a 167-kb region with 18 candidate genes for LTG (Shim et al., 2019). In the present study, we identified *OsD2* (LOC_Os01g10040) as a candidate gene for *qLTG1*. Two T-DNA insertion mutants were used to confirm the association between *OsD2* and LTG expression. *OsD2* encodes a cytochrome P450 protein involved in BR biosynthesis. To determine the relationship between BR and LTG, BR-feeding experiments were conducted at the germination stage under low-temperature conditions. The expression levels of BR signaling and biosynthesis genes were examined to understand the molecular mechanisms of the BR pathway. Gene expression analysis and BR application experiments demonstrated that BR enhanced the seed germination rate under low-temperature conditions. These data suggest that *OsD2* is associated with the regulation of LTG and that the *O. rufipogon OsD2* allele is a suitable option for improving LTG in rice.

Materials and methods

Plant materials and growth condition

In a previous study, 96 introgression lines (ILs; TR1–96) were developed from an interspecific cross between the Korean *japonica* line Hwaseong and *O. rufipogon* (Yun et al., 2016). QTL analysis for LTG using these 96 ILs revealed five QTLs, including *qLTG1* (Nguyen et al., 2012). To fine-map *qLTG1*, two introgression lines (TR5 and TR20) harboring the *O. rufipogon qLTG1* allele in the Hwaseong background were crossed with Hwaseong, and 971 F_2 plants were developed. The $F_{2:3}$ line was used for substitution mapping of *qLTG1*. Two T-DNA insertion lines (PFG-3A-07238 and PFG-3A-07287) in *OsD2* genic region were obtained from Kyung Hee University, Yongin, South Korea (Jeon et al., 2000; Jeong et al., 2006), and the T-DNA insertion was confirmed using two sets of PCR primers (Supplementary Table 1). From these T-DNA lines, we selected plants heterozygous at the T-DNA locus because plants homozygous for the T-DNA insertion showed a severe dwarf phenotype with no seed setting (Supplementary Figure 1). We tested

the genotypes of 40 seeds from heterozygous plants and observed 1:2:1 ratio of no T-DNA insertion, heterozygous and homozygous insertion, respectively suggesting that there is no segregation distortion at this locus ($\chi^2 = 0.105$, $0.95 \geq p \geq 0.90$). To investigate the sequence variations in *OsD2*, 96 rice accessions from the KRICE_CORE set were examined (Kim et al., 2016; Supplementary Table 2). Plant materials for fine-mapping and T-DNA insertion lines were sown in mid-April, and 30-day-old seedlings were transplanted into a paddy field at Chungnam National University, Daejeon, South Korea. Furthermore, 96 rice accessions from the KRICE_CORE set were grown in an experimental field at Chungcheongnam-do Agricultural Research and Extension Services (CNARES), Yesan, South Korea. Data of agronomic traits of Hwaseong, TR5, and TR20 were used from our previous study (Yun et al., 2016).

DNA extraction and genotype analysis

DNA was extracted from fresh leaf tissues using the method described by Causse et al. (1994). PCR was carried out as described by Jeon et al. (2018), with minor modifications: 95°C for 5 min, followed by 35 cycles at 95°C for 30 s, 55–58°C for 30 s, and 72°C for 30 s, and 5 min at 72°C for the final extension. PCR products were separated on a 2–3% agarose gel stained with StaySafe Nucleic Acid Gel Stain (RBC, New Taipei City, Taiwan). For *qLTG1* fine-mapping, InDel markers were designed (Supplementary Table 1). T-DNA insertion genotyping was conducted using two sets of primers (Supplementary Table 1).

Evaluation of low-temperature germination

Germination tests were conducted as described by Nguyen et al. (2012) and Shim et al. (2019, 2020), with minor modifications. Seeds were harvested 45 days after flowering, and dried in a greenhouse for 2 weeks. Seeds from two T-DNA insertional heterozygous lines were harvested and used for germination test. To break dormancy, seeds were incubated at 55°C for 3 days. To confirm the breakage of seed dormancy, 20 seeds were germinated in duplicates at optimal germination temperature (28°C) and dark conditions. For the LTG test, 20 seeds were placed in a 60 mm petri dish with filter paper and filled with 5 ml of distilled water. Seeds were incubated in triplicates at 13°C and dark conditions. Seeds were considered germinated when the seed coat was broken and the white embryo emerged. Germinated seeds were counted, and the germination percentage (%) was calculated. All germination experiments were conducted three times.

OsD2 haplotype analysis

Haplotype analysis was carried out using KRICE_CORE set genotype data from the Kongju National University, Yesan, South Korea (Kim et al., 2016, Supplementary Table 2). LTG phenotype data were obtained from our previous study (Shim et al., 2020).

Feeding experiment

To examine germination under stress conditions, seed germination tests were conducted using Hwaseong, *O. rufipogon,* and TR5 under salinity (250 mM NaCl) and high-osmotic (500 mM mannitol) conditions. Seeds were incubated at 28°C for 7 days under each stress condition.

To determine whether BR is associated with LTG, low-temperature germination tests were carried out under 1 μM 24-epi-brassinolide (eBL; Sigma-Aldrich Co., MO, United States) and 1 μM brassinazole (BRZ; Sigma-Aldrich Co., MO, United States) treatment conditions. T-DNA insertion lines (*OsD2*-KD and *OsD2*-OE) were used for the BR treatment test, and 25 seeds were incubated in triplicates at 13°C for 7 days. Germinated seeds were counted, and the germination rate (%) was calculated daily after incubation. All germination experiments were repeated three times.

RNA isolation and real time quantitative reverse transcription PCR

Total RNA was isolated from seeds at the germination stage or from 7-day-old seedlings using a NucleoSpin RNA kit (Macherey Nagel, Deuren, Germany), according to the manufacturer's instructions. Following reverse-transcription into the first-strand cDNA using SmartGene Mixed cDNA synthesis kit (SJ Bioscience, Daejeon, Korea), real-time PCR was performed using a CFX Connect Real-Time System (Bio-Rad, CA, United States). The real-time PCR protocol and conditions were as described by Jeon et al. (2021), with minor modifications: 15 min at 95°C to denature and activate the enzyme, followed by 40 cycles at 95°C for 20 s, 60–55°C for 40 s (depending on primer annealing temperature), and 72°C for 30 s. The quantitative reverse transcription PCR (qRT-PCR) data were analyzed using the $2^{-\Delta\Delta Ct}$ method. *Rice tumor-expressed protein* (*OsTMP*) was used for normalization, and relative expression levels were compared using Student's *t*-test. The primers used in this study are listed in Supplementary Table 1.

Statistical analysis

One-way ANOVA and Tukey's test were carried out using Minitab19 software. Student's *t*-test was performed using Microsoft Excel software.

Results

Comparison of germinability under stress conditions

Three accessions, Hwaseong, *O. rufipogon*, and the introgression line TR5 harboring *qLTG1*, were tested for germinability under three stress conditions. Initially, the germination rate was evaluated at 28°C to check seed dormancy. All lines showed a 100% germination rate 4 days after incubation (DAI), indicating that seed dormancy was broken (Figure 1A). Then, we compared LTG at 13°C, and *O. rufipogon* showed the highest LTG, followed by TR5 and Hwaseong (Figure 1B). Subsequently, germinability was evaluated under osmotic and salinity stress conditions at 28°C. In the 500 mM mannitol treatment, *O. rufipogon* and TR5 plants showed similar germination patterns, and their germination rates were higher than that of Hwaseong plants (Figure 1C). The highest germination rate was observed in *O. rufipogon* under the 250 mM NaCl treatment (Figure 1D). The germinability of TR5 was higher than that of Hwaseong but lower than that of *O. rufipogon*. These results indicate that the *O. rufipogon qLTG1* allele increased germinability at low temperatures, as well as under osmotic and salinity stress conditions.

Fine-mapping of *qLTG1*

To narrow down the *qLTG1* locus, we fine-mapped the *qLTG1* region using recombinant plants. Four recombinants with breakpoints between two markers, RM220 and RM6277, were observed in 971 F_2 plants; these plants were then self-pollinated to produce F_3 homozygous lines (Figure 2). To determine the exact breakpoint in the four recombinants, newly designed InDel markers (CRM1–40) were used. The line 4 was informative in determining the target region of the candidate genes because the LOC_Os01g10040 allele of line 4 was a hybrid between those of Hwaseong and *O. rufipogon*. This line showed significantly lower germination rate than TR5. Sequencing LOC_Os01g10040 of the line 4 revealed that the chromosome recombination point was

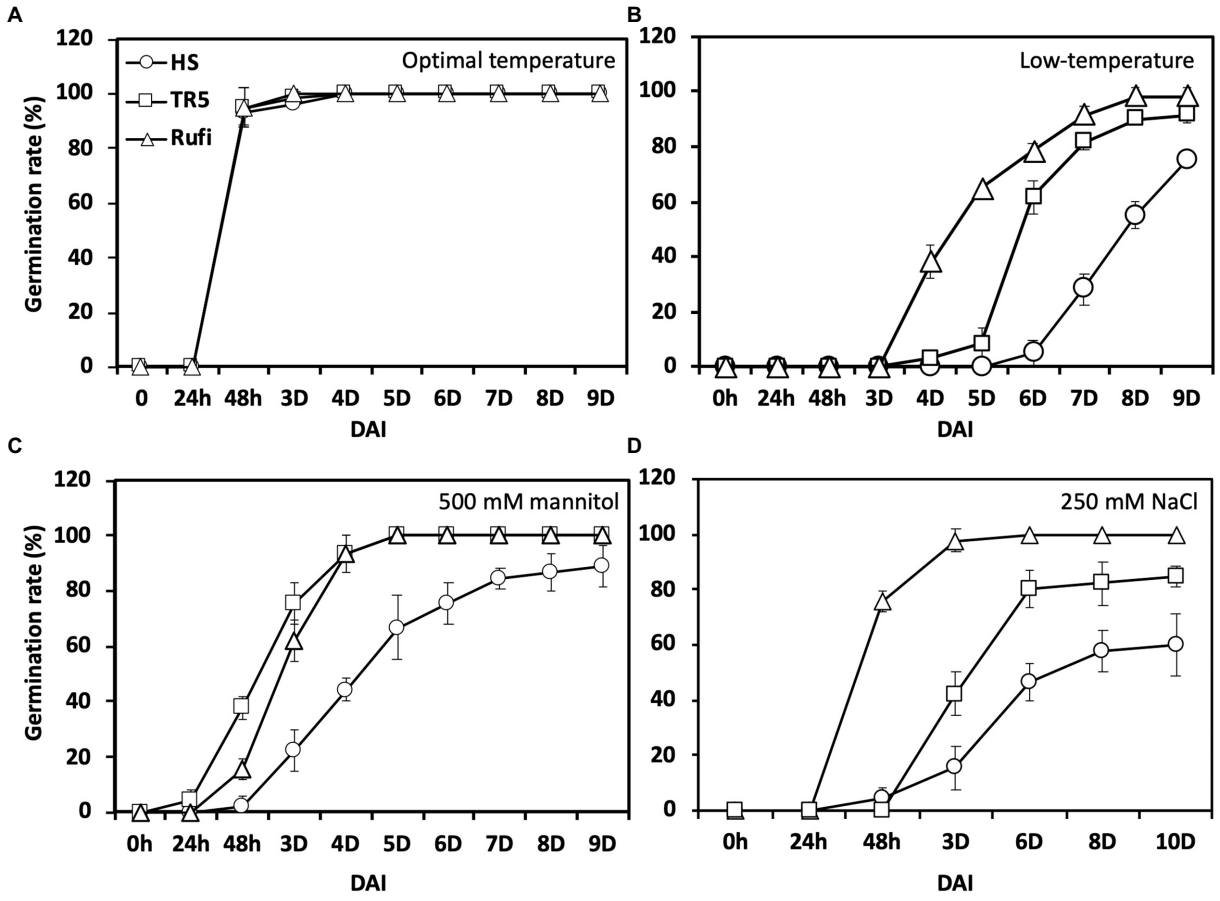


FIGURE 1
Seed germination under low-temperature and abiotic stress conditions using Hwaseong, *Oryza rufipogon*, and TR5. **(A)** Germination rate under optimal temperature (28°C); **(B)** Germination rate under low-temperature (13°C); **(C)** Germination rate under high osmotic stress (500mM mannitol) at 28°C; and **(D)** Germination rate under high salinity stress (250mM NaCl) at 28°C. Triangle, square, and circle indicate *O. rufipogon*, TR5, and Hwaseong, respectively. Data are presented as mean±SD (n=3). DAI, days after incubation.

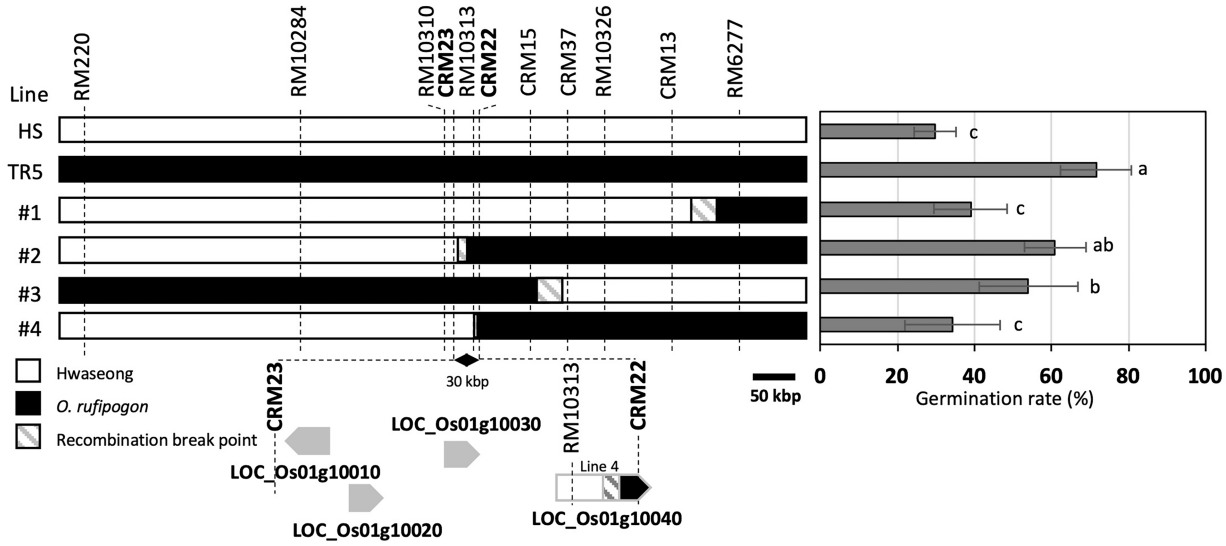


FIGURE 2
Substitution mapping of *qLTG1*. Germination percentage was measured at 7 days after incubation. Error bars indicate the SD of the mean. Different letters to the right of the bar graph indicate significant differences at $p<0.05$, based on Tukey's test.

TABLE 1 List of candidate genes in the *qLTG1* region.

Gene ID	Description
LOC_Os01g10010	Prenylated rab acceptor
LOC_Os01g10020	Ectonucleotide pyrophosphatase/phosphodiesterase family member 1
LOC_Os01g10030	Hypothetical protein
LOC_Os01g10040	Cytochrome P450, Brassinosteroids biosynthesis, Regulation of plant architecture

between the fourth intron and the seventh intron suggesting that LOC_Os01g10040 of the line 4 allele was not as effective as *O. rufipogon* in LTG and this gene should be included as candidate gene for *qLTG1*. Finally, *qLTG1* was delimited to a 30-kb region between CRM23 and CRM22. Based on the Nipponbare reference genome, this region contains four genes (Figure 2; Table 1).

Candidate gene analysis

The fine-mapping results indicated the existence of four genes (LOC_Os01g10010, LOC_Os01g10020, LOC_Os01g10030, and LOC_Os01g10040) at the *qLTG1* locus. Based on release 7 of the MSU Rice Genome Annotation Project on rice IRGSP-1.0 genome,[1] annotations of these four genes were as follows: the prenylated Rab acceptor for LOC_Os01g10010, ectonucleotide pyrophosphatase/phosphodiesterase family member 1 for LOC_Os01g10020, hypothetical protein for LOC_Os01g10030, and cytochrome P450, BR biosynthesis, and regulation of plant architecture for LOC_Os01g10040 (Table 1). Four candidate genes have sequence variation between Hwaseong and *O. rufipogon* including synonymous and non-synonymous variations (Supplementary Tables 3, 4).

To select the candidate gene for *qLTG1*, gene expression levels were examined at four different time points during seed germination (0, 24, 48, and 72 h after incubation) at 13°C in Hwaseong, *O. rufipogon*, and TR5 (Supplementary Figure 2). Among the four genes, LOC_Os01g10040 displayed a consistent expression pattern during germination in the three lines. The expression of LOC_Os01g10040 in *O. rufipogon* and TR5 was higher than that in Hwaseong 24 and 48 h after incubation, whereas no significant difference was observed at 72 h. Compared with LOC_Os01g10040, the expression patterns of the other three genes were inconsistent. These results strongly suggest an association between LOC_Os01g10040 and LTG. Several studies have reported that LOC_Os01g10040 (*OsD2*) regulates several key agronomic traits, including plant height, leaf angle, and grain size (Hong et al., 2003; Li et al., 2013a; Fang et al., 2016). Based on these findings, *OsD2* was chosen as the candidate gene for further analysis.

Sequence comparison of *OsD2* between Hwaseong and *Oryza rufipogon*

The sequence of *OsD2* was compared between Hwaseong and *O. rufipogon* (Supplementary Figure 3; Supplementary Table 4). The size of the *OsD2* genomic sequence and coding sequence were 7,783 and 1,473 bp, respectively, and several sequence differences were observed between Hwaseong and *O. rufipogon*. Although

1 http://rice.uga.edu

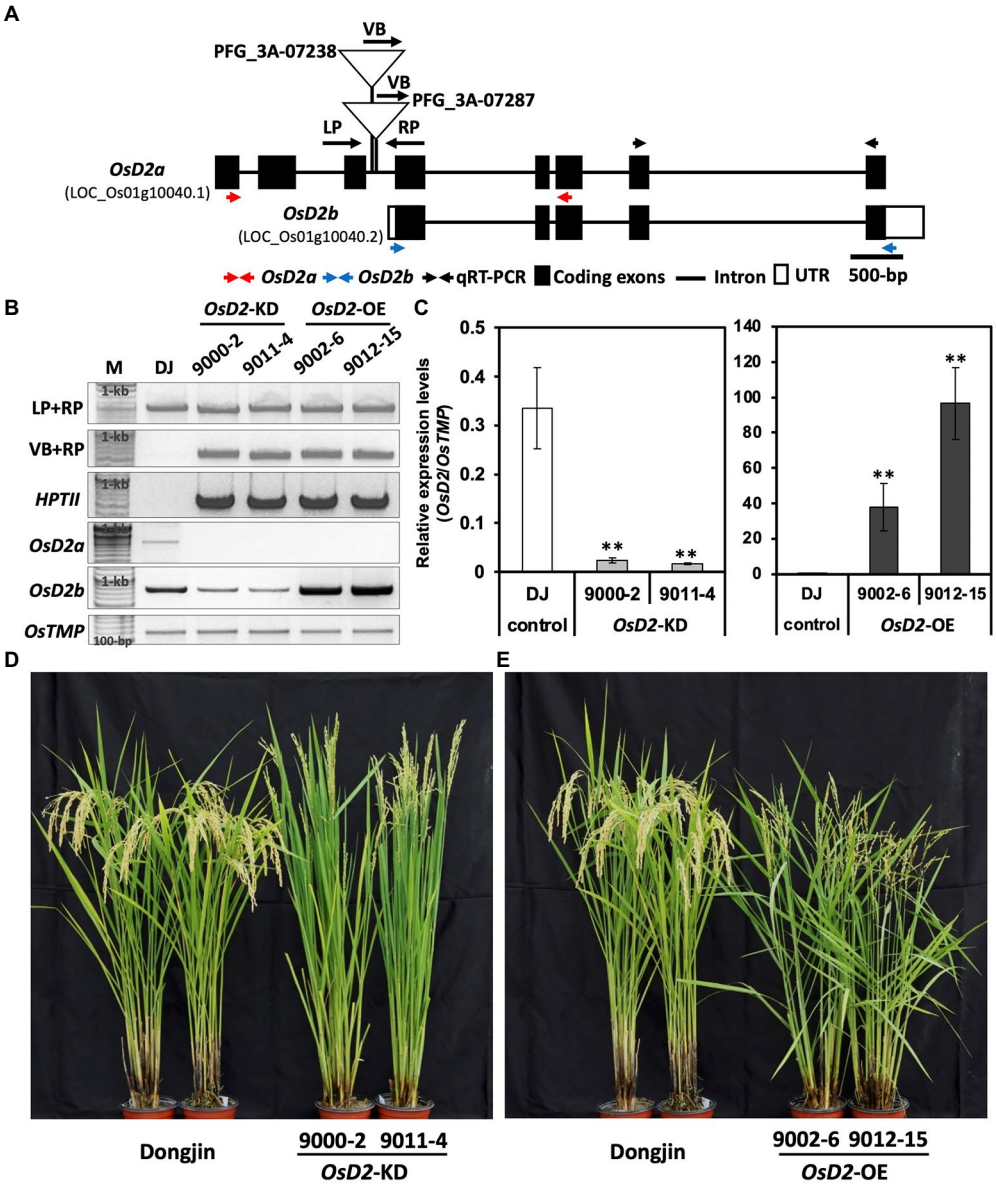


FIGURE 3
Characterization of T-DNA insertion lines. **(A)** Gene structure of *OsD2* and T-DNA insertion sites in two lines. VB, vector border; LP, left primer; and RP, right primer. **(B)** PCR amplicons to confirm T-DNA insertion together with hygromycin resistant gene (*HPTII*) and PCR amplicons for transcript form confirmation of *OsD2a* and *OsD2b* using cDNA. M: 100-bp size marker, DJ: Dongjin. **(C)** Gene expression level of *OsD2* in the T-DNA insertion lines. *OsTMP* was used for gene normalization. Error bars indicated SD. ** indicates a significant difference between each T-DNA insertion line and Dongjin based on the Student's *t*-test at $p<0.01$. **(D,E)** Phenotypic difference between Dongjin and T-DNA insertion lines.

many differences were detected in the genomic sequences, we focused only on the exon region. Five SNPs were found in the exon region; the SNPs were located on the first, fourth, and seventh exons. Among the five SNPs, three were non-synonymous; these SNPs (+76_A/G, +1706_C/T, and + 1726_G/A) changed amino acids from Gly, Leu, and Met (Hwaseong) to Ser, Pro, and Val (*O. rufipogon*), respectively (Supplementary Figure 3A). The first non-synonymous SNP (+76_A/G) was located on the membrane anchor domain, whereas the other two were not located on known protein domains (Supplementary Figure 3B). Additionally, 14 sequence variants were observed in the promoter region (Supplementary Table 4).

OsD2 controls LTG

To determine whether *OsD2* regulates LTG, we obtained two T-DNA insertion lines (PFG_3A-07238 and PFG_3A-07287) from the Kyung Hee University. PCR analysis confirmed the T-DNA insertions in the third intron (Figures 3A,B). Homozygous plants

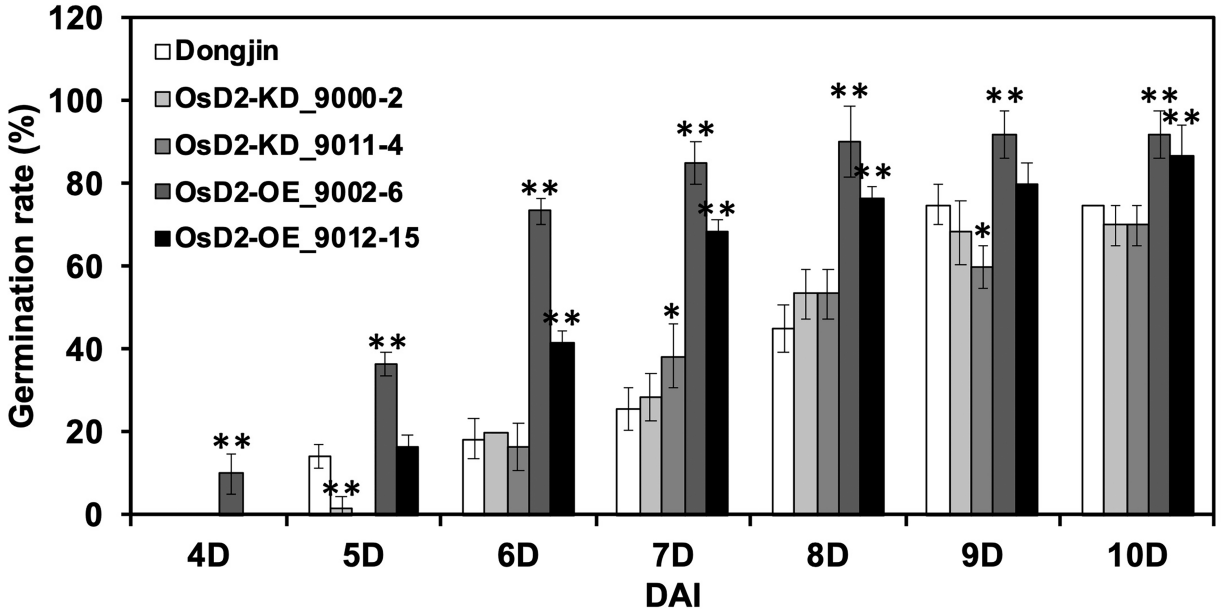


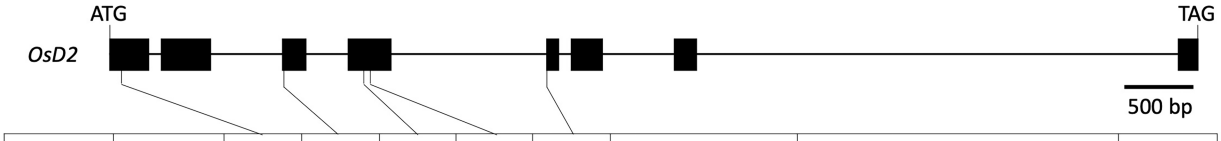
FIGURE 4
Low-temperature germinability of *OsD2* T-DNA insertion lines and Dongjin (DJ). Seeds of *OsD2*-KD (9000-2, 9011-4), *OsD2*-OE (9002-6, 9012-15), and Dongjin were incubated at 13°C for 10 days. Data are presented by mean±SD (n=3). * and ** indicate significant difference between each T-DNA insertion line and Dongjin at p<0.05 and p< 0.01, respectively, based on the Student's t-test.

with T-DNA insertion in PFG_3A-07238 showed severe dwarfism with no seed setting, whereas heterozygous plants displayed a non-dwarf phenotype (Figure 3D, Supplementary Figure 1). The qRT-PCR results showed that *OsD2* expression in PFG_3A-07238 heterozygous plants was significantly lower than that in Dongjin plants (Figure 3C). On the other hand, PFG_3A-07287 heterozygous plants showed higher *OsD2* expression levels than Dongjin plants (Figure 3C). *OsD2* has two types of transcript form, *OsD2a* (LOC_Os01g10040.1) and *OsD2b* (LOC_Os01g10040.2). PFG_3A-07287 includes pGA2715 vector which harbors CaMV 35S enhancers which might have induced the overexpression of the *OsD2b* transcript, leading to higher expression of *OsD2b* in PFG_3A-07287. Therefore, we selected two heterozygous lines for *OsD2*-knockdown (*OsD2*-KD; line numbers: 9000-2 and 9011-14) and *OsD2*-overexpression (*OsD2*-OE; line numbers: 9002-6 and 9012-15). To check whether *OsD2*-KD and *OsD2*-OE make *OsD2a* or *OsD2b* transcript, PCR was conducted using cDNA of Dongjin, *OsD2*-KD, and *OsD2*-OE (Figure 3B). *OsD2a* transcript was observed in Dongjin while amplicon was not detected in *OsD2*-KD and *OsD2*-OE. *OsD2b* transcript was observed in all materials and *OsD2*-KD displayed lower expression of *OsD2* transcript than Dongjin. *OsD2*-OE showed higher expression of *OsD2b* transcript than Dongjin and *OsD2*-KD. The two *OsD2*-KD lines showed erect plant architecture with smaller grain size than Dongjin plants, whereas the *OsD2*-OE lines displayed an open plant architecture with longer grains than Dongjin plants (Figures 3D,E, Supplementary Figure 4).

The LTG test was performed using two T-DNA lines and wild-type Dongjin (Figure 4). At the optimal germination temperature, Dongjin, *OsD2*-KD, and *OsD2*-OE seeds showed 100% germination rates at three DAI (Supplementary Figure 5). The LTG of *OsD2*-OE was significantly higher than that of Dongjin and *OsD2*-KD. *OsD2*-OE (9002-6) started germination at four DAI, while the other lines started germination at 5 or 6 DAI, indicating that *OsD2* plays a role in LTG regulation. However, *OsD2*-KD lines displayed a lower germination rate than Dongjin at 5 DAI, while the LTG differences between *OsD2*-KD lines and Dongjin were not significantly different except at 7 and 9 DAI. This is possibly due to the genetic background effects of seeds obtained from the heterozygous plants.

Haplotype analysis for *OsD2*

To validate the allelic effects of *OsD2*, haplotype analysis was conducted using 96 rice accessions from the KRICE_CORE set (Supplementary Table 2). Four haplotypes (Hap1–4) were found when all five SNPs were considered (Figure 5). Hap1 contained 64 accessions, including 50 temperate *japonica*, 12 tropical *japonica*, one admixture, and one aromatic rice. Hwaseong was classified as Hap1, and this haplotype showed 40.3% LTG at 6 DAI. Hap2 had 13 accessions, including one tropical *japonica*, nine *indica*, two Aus, and one admixture. The rice accessions in Hap2 displayed an average of 54.6% LTG. *Oryza rufipogon* belonged to Hap3, along with seven *indica*, four *indica* (Tongil), and one Aus accession. Under low-temperature conditions, Hap3 had 56.0% mean germination rate. Seven *indica* accessions were classified as Hap4 and showed 61.1% mean germination rate at 13°C. ANOVA was



Haplotype		No. of accessions	+76	+1188	+1706	+1726	+2970	Germination rate (%) [1]		Ecotypes classified by KRICE_CORE set [2]							Reference
		Nucleotide	A/G	G/T	C/T	G/A	A/G			*TEJ*	*TRJ*	*IND*	*IND (Tongil)*	*Aus*	*Adm*	*Aro*	
		Amino acid	Gly/Ser	Gly/Cys	Leu/Pro	Met/Val	Ile/Met										
HapA	Hap1	64	A	G	C	G	A	40.3 ± 3.4		50	12				1	1	Hwaseong
HapB	Hap2	13	G	G	T	A	G	56.6 ± 5.8	54.6 ± 10.7		1	9		2	1		
	Hap3	12	G	G	T	A	A		56.0 ± 9.5			7	4	1			*O. rufipogon*
	Hap4	7	G	T	T	A	A		61.1 ± 9.7			7					
ANOVA (*P*-value)			0.011*	0.170	0.011*	0.011*	0.249	0.011*	0.087	50	13	23	4	3	2	1	

FIGURE 5
Haplotype analysis of *OsD2* using 96 rice accessions from the KRICE_CORE set. 1) Germination rate was presented as mean±SE. 2) *TEJ*, temperate *japonica*; *TRJ*, tropical *japonica*; *IND*, *indica*; *IND* (Tongil), *indica* Tongil; *Aus*, Aus; *Adm*, admixture; and *Aro*, aromatic. * indicates significant difference at $p<0.05$ based on one-way ANOVA.

performed to determine whether LTG was significantly different among the four haplotypes at $p<0.05$ level. However, the difference was not significant, possibly because of a large amount of variation within the haplotypes ($p=0.087$). We then evaluated the effects of each of the five SNPs to identify informative SNPs associated with the LTG variation. At three SNPs (+76_A/G, +1706_C/T, and +1726_C/T), the four haplotypes were classified into two groups (Hap1 and Hap2–4). At the second SNP (+1188_G/T), two haplotype groups, Hap1–3 and Hap4, were formed. Hap2 formed one group, and the other three haplotypes, Hap1, 3, and 4, formed the other group based on the fifth SNP (+2970_A/G). ANOVA showed that the difference in LTG between the two groups was significant ($p=0.011$) at three SNPs (+76_A/G, +1706_C/T, and+1726_C/T), but not significant at the second or fifth SNP (Figure 5). Therefore, Hap1 and Hap2–4 were formed as new haplotype group as HapA and HapB based on the three significant SNPs, respectively (Figure 5). These results indicate that three SNPs (+76_A/G, +1706_C/T, and+1726_C/T) are informative in explaining the LTG variation, and these SNPs could be used in rice breeding programs for selecting parents with enhanced LTG.

BR is associated with low-temperature germination

OsD2-KD and *OsD2*-OE plants showed several differences in plant architecture compared to Dongjin plants (Figures 3D,E). *OsD2*-KD plants flowered 1 week later than the control Dongjin plants, had erect leaves with a narrow leaf angle, and did not show dwarfism. In contrast, *OsD2*-OE plants were shorter in plant height than Dongjin plants and had bent leaves, indirectly implying a higher amount of BR.

To determine whether BR content variation was responsible for the improved LTG in *OsD2*-OE, a feeding experiment was conducted. We investigated LTG following seed treatment with exogenous eBL and BRZ (a BL biosynthesis inhibitor; Figure 6). The low-temperature germination rates of Hwaseong, TR5, Dongjin, *OsD2*-KD (9000-2 and 9011-4), and *OsD2*-OE (9012-15) increased under eBL treatment, whereas the germination rates of *O. rufipogon* and *OsD2*-OE (9002-6) at low temperatures were not significantly different from the control condition. In contrast, BRZ treatment decreased the low-temperature germination rate of all lines. This experiment showed that the exogenous application of BR promoted LTG, whereas the application of a BR biosynthesis inhibitor decreased LTG. These results indicate that BR is associated with LTG.

Expression patterns of BR biosynthesis and signaling genes

Endogenous BR concentration is a major factor regulating the expression of BR biosynthesis and signaling pathway genes. The plant architecture of *OsD2*-KD and *OsD2*-OE plants indirectly indicated the endogenous BR content. The expression patterns of genes involved in BR signaling and biosynthetic pathways were analyzed using transgenic plants (Supplementary Figure 6). The transcript levels of the receptor-like kinase *BRASSINOSTEROID INSENSITIVE 1* (*OsBRI1*), co-receptor kinase *BRI1-ASSOCIATED KINASE 1* (*OsBAK1*), *BRASSINAZOLE-RESISTANT 1* (*OsBZR1*), and *INCREASE LEAF INCLINATION 4* (*OsILI4*) were considerably lower in both *OsD2*-KD and *OsD2*-OE than in Dongjin (Supplementary Figure 6A). Lower expression of two *GLYCOGEN SYNTHASE KINASE* (*OsGSK2* and *OsGSK3*) and *DWARF* and *LOW TILLERING* (*OsDLT*) genes was observed in *OsD2*-KD lines than in Dongjin, whereas the expression of the these three genes was higher in *OsD2*-OE lines than in Dongjin (Supplementary Figure 6A). Among the major BR biosynthesis genes, higher expression levels of *OsD11, OsCPD1,* and *OsBR6ox* were detected in the *OsD2*-OE lines than in Dongjin, whereas the BR-deficient gene, *OsBRD1*, was downregulated in the *OsD2*-OE

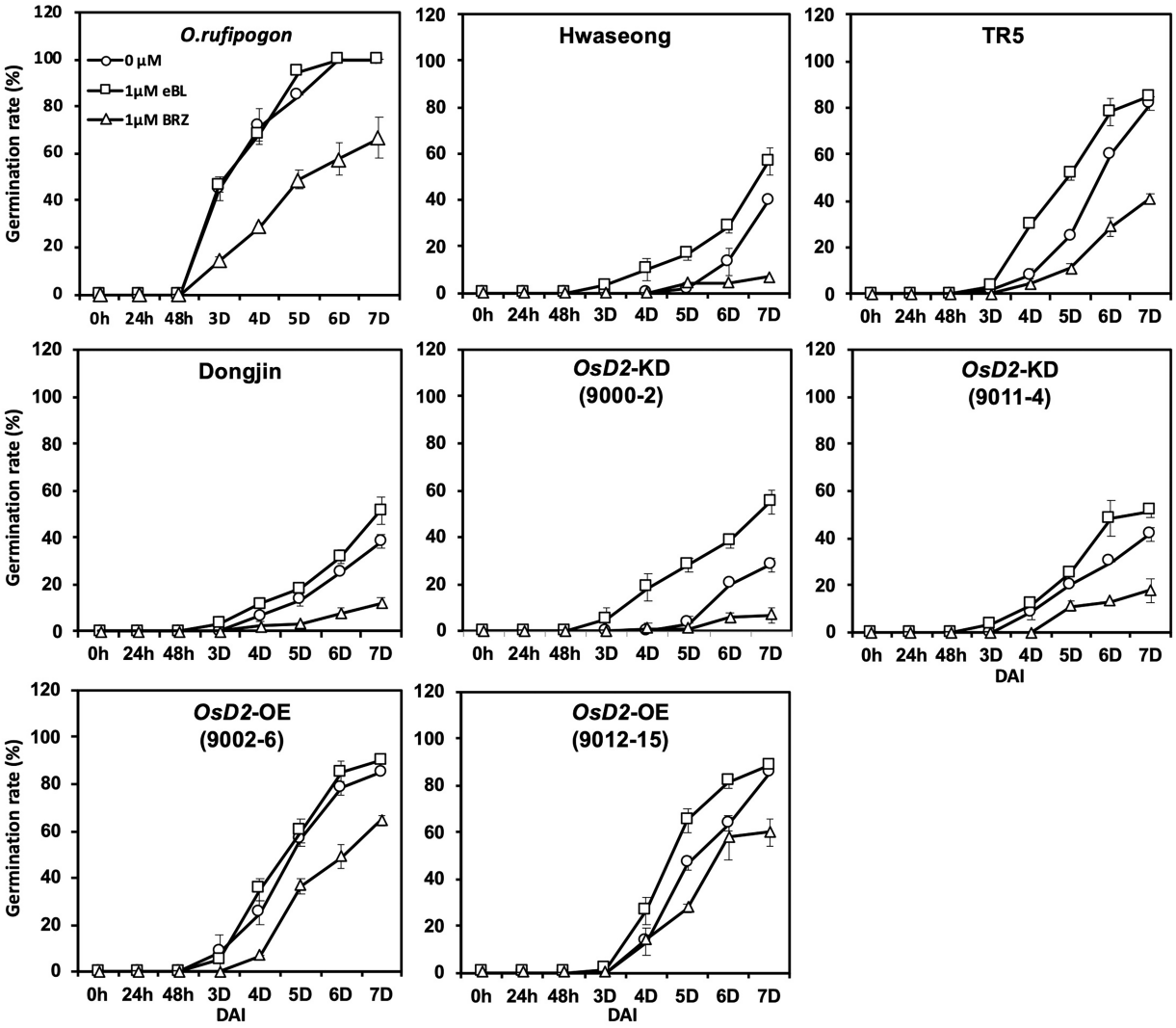


FIGURE 6
Comparison of the low-temperature germinability of eight plant materials under the exogenous brassinosteroid and brassinosteroid biosynthesis inhibitor treatment. Twenty-five seeds were incubated at 13°C with control (circle), 1 μM 24-epi-brassinolide (eBL; square), and 1 μM brassinazole (BRZ; triangle) for 7 days. Data are presented as mean±SD (n=3). DAI, days after incubation.

lines (Supplementary Figure 6B). Conversely, the *OsD2*-KD lines showed lower expression of *OsD11* and *OsBRD1* than Dongjin. The different expression levels of BR signaling and biosynthesis genes in T-DNA insertion lines might be associated with the difference in endogenous BR content, leading to the variation in LTG, plant architecture, and seed size in the T-DNA lines.

Discussion

Low-temperature germinability is an important trait for direct-seeded rice cultivation. Because of the significance and complex nature of LTG in rice, many genetic approaches have been used to identify genes involved in LTG (Satoh et al., 2016; Wang et al., 2018; Yang et al., 2020). To date, two genes, *qLTG3-1* and *OsSAP16* have been characterized using biparental and GWAS approaches, respectively (Fujino et al., 2008; Wang et al., 2018). In this study, we demonstrated that the BR biosynthesis gene *OsD2* is the causal gene for the LTG QTL, *qLTG1*. Two introgression lines, TR5 and TR20 (BC_3F_6), harboring the *O. rufipogon qLTG1* allele, showed significantly enhanced LTG compared to Hwaseong (Shim et al., 2020). Map-based cloning enabled us to delimit the QTL to a ~30-kb region flanked by two markers, CRM23 and CRM22. The region has four genes: prenylated rab acceptor (LOC_Os01g10010), ectonucleotide pyrophosphatase/phosphodiesterase family member 1 (LOC_Os01g10020), hypothetical protein (LOC_Os01g10030), and cytochrome P450 (LOC_Os01g10040, *OsD2*). To identify the causal gene for *qLTG1*, gene sequencing, expression analysis, and transgenic approaches were employed. Gene expression analysis indicated that the *O. rufipogon* and TR5 seeds showed higher *OsD2* expression than

the Hwaseong seeds during germination, whereas the three lines (Hwaseong, *O. rufipogon*, and TR5) did not show any consistent expression pattern in the other three genes (LOC_Os01g10010–LOC_Os01g10030). Based on the gene expression analysis, *OsD2* was selected as the candidate gene for *qLTG1*. Sequence comparison of *OsD2* between Hwaseong and *O. rufipogon* revealed five SNPs in the exon region and 14 sequence variants in the promoter region. To confirm the association of *OsD2* with LTG, we employed two T-DNA insertion lines: *OsD2*-OE and *OsD2*-KD plants in the Dongjin background. The LTG test using the two T-DNA lines and Dongjin showed that *OsD2*-OE had significantly higher LTG than Dongjin and *OsD2*-KD. To determine whether BR is responsible for the improved LTG in *OsD2*-OE, feeding experiments were conducted using exogenous eBL and BRZ (Figure 6). The low-temperature germination rates of Hwaseong, TR5, Dongjin, *OsD2*-KD (9000-2 and 9011-4), and *OsD2*-OE (9012-15) increased under eBL treatment, whereas the germination rates of *O. rufipogon* and *OsD2*-OE (9002-6) at low temperatures were not significantly different from the control condition. In contrast, BRZ treatment resulted in a decrease in the low-temperature germination rate of all lines. Studies have reported that LOC_Os01g10040 (*OsD2*) plays an important role in regulating plant height, grain size, and leaf angle (Hong et al., 2003; Li et al., 2013a; Fang et al., 2016). Our findings indicated that *OsD2* is also associated with LTG variation in rice.

The *ebisu dwarf* (*dwarf2* or *d2*), an *OsD2* gene mutant, has been characterized, and several *d2* alleles have been reported (Hong et al., 2003; Li et al., 2013a; Fang et al., 2016). Among these alleles, *d2-3*, *d2-4*, *d2-6*, and *ccdd1* showed severe dwarfism with undesirable traits, making their utilization impractical in breeding programs. Although the *d2-1*, *d2-2,* and *smg11* alleles of *OsD2* displayed non-dwarf phenotypes, plants with these alleles had reduced plant height, small grains, and erect plant architecture (Hong et al., 2003; Fang et al., 2016). To date, many studies have reported the function of BRs as a class of plant steroidal hormones and have focused on the effect of BR on agronomic and morphological traits (Fujioka and Yokota, 2003; Wu et al., 2008). In contrast, we analyzed the association of *OsD2* with germinability under stress conditions and found that the *O. rufipogon OsD2* allele improved LTG in the cultivated rice background without any deleterious phenotypes such as dwarfism. In addition, two introgression lines TR5 and TR20 harboring the *O. rufipogon qLTG1* allele showed improved agronomic traits, such as increased grain size, spikelets per panicle, 1,000-grain weight, and grain yield whereas their tall statue are unfavorable because of lodging (Yun et al., 2016, Supplementary Figures 7, 8). Although it is unclear whether these changes in agronomic traits are due to the pleiotropic effect of *OsD2* or other tightly linked genes, these results support the fact that wild rice is a valuable resource for improving yield stability and widening genetic diversity in rice breeding programs (Xiao et al., 1996; Yeo et al., 2014; Yun et al., 2016; Shim et al., 2019).

Haplotype analysis of 96 rice accessions with five non-synonymous SNPs in *OsD2* was performed to determine the association between haplotypes and LTG (Figure 5). When the five SNPs were combined, the 96 accessions were classified into four haplotypes (Hap1–4). However, the difference in LTG levels among the four haplotypes was not significant. Then, each of the five SNPs was considered individually, and the 96 accessions were classified into two haplotypes. At three SNP sites (+76, +1,706, and +1,726), Hap1 and Hap2–4 formed two different haplotypes (HapA and HapB), and two groups showed significant differences in LTG ($p < 0.05$, ANOVA). Interestingly, rice accessions in HapB shared the same genotype at +76, +1,706, and +1,726 SNPs, and three SNPs each showed significant differences in LTG based on ANOVA ($p = 0.011$; Supplementary Figure 3). It remains to be determined whether the amino acid sequence of *OsD2* between the two lines is directly associated with LTG. However, the possibility that the sequence difference in the *OsD2* promoter region is responsible for the variation in LTG cannot be ruled out. Although further studies on protein structure and enzymatic experiments are needed to answer this question, the molecular markers for +76, +1,706, and +1,726 SNPs would be useful in selecting high-LTG germplasms together with the previously designed *qLTG3-1* functional markers (Shim et al., 2020). In addition, pyramiding two *O. rufipogon* alleles at *qLTG1* and *qLTG3* will be an effective strategy for developing direct-seeding varieties with high LTG.

Two T-DNA insertion lines were used for *OsD2* characterization (Jeon et al., 2000; Jeong et al., 2006). Generally, homozygous T-DNA insertion lines are used to examine the effects of gene knockout or overexpression. However, in our study, we employed heterozygous T-DNA plants for phenotypic evaluation because the homozygous T-DNA insertion plants showed severe dwarfism and failed to set seeds. The absence of homozygous seeds made us to employ segregating seeds in the germination test and this could have affected the accuracy of the germination test. The seeds harvested from heterozygous plant will be segregating with 1:2:1, 3:1, or 1:3 ratio depending on the gene action (additive, dominant, or recessive). Therefore, 75% of seeds of heterozygous plants have at least one-copy T-DNA insertion and the number of seeds showing higher or lower *OsD2* expression than the WT Dongjin would be different depending on the gene action. This might have affected the accuracy of the LTG test especially in *OsD2*-KD lines.

The heterozygous plants (*OsD2*-KD) showed lower *OsD2* expression similar with other BR-deficient mutants phenotypes, such as *d2-1* and *d2-2,* and displayed erect leaves with decreased plant height and grain size (Figure 3; Hong et al., 2003). The *OsD2*-OE plants showed a higher expression of *OsD2,* although the insertion site of T-DNA was located near the site for *OsD2*-KD. It is possible that the CaMV 35S promoter enhancer in the vector border may have activated the expression of *OsD2b* transcript form (Jeong et al., 2006; Li et al., 2013a). Similar studies using T-DNA insertion lines for overexpression and decreased gene expression have been reported. Two T-DNA insertion mutants, *oswrky5-D* and *oswrky5* showed increased and decreased expression levels of *OsWRKY5*, respectively, and the expression of *OsWRKY5* was positively associated with the leaf senescence phenotype (Kim et al., 2019).

Studies have reported that BRs play an important role in promoting seed germination (Leubner-Metzger, 2001; Steber and McCourt, 2001). Leubner-Metzger (2001) observed that

brassinolide promotes tobacco seed germination under both light and dark incubation conditions. BR signaling is required to overcome the inhibition of germination by ABA, suggesting that BR stimulates seed germination (Steber and McCourt, 2001). It has been demonstrated that the exogenous application of BRs enhanced the seed germination rate under stress conditions in rice and cucumber (Anuradha and Rao, 2001; Wang et al., 2011). However, the effects of BR under low-temperature stress conditions have not yet been examined. In our feeding experiment, the application of 1 μM eBL improved the germination rate of Hwaseong, TR5, Dongjin, *OsD2*-KD (9000-2 and 9011-4), and *OsD2*-OE (9012-15) at low temperatures (Figure 6). In contrast, no significant difference was observed in in the germination rates of *O. rufipogon* and *OsD2*-OE (9002-6). This was possibly due to the high level of endogenous BR in the two lines, and the eBL treatment did not contribute to the increase in LTG compared to the control. On the other hand, BRZ treatment decreased LTG in all tested lines. These results clearly indicate that BR is associated with LTG.

In this study, the expression patterns of some BR signaling and biosynthesis genes were not consistent with those reported in previous studies. For example, the lower expression of *OsILI4* (*BRASSINOSTEROID UPREGULATED1, OsBU1*), encoding Helix–loop–Helix protein involved in BR signaling, in both *OsD2*-KD and *OsD2*-OE is not consistent with previous reports. Sakamoto et al. (2012) reported that the relative *OsBU1* expression levels in *d2-1*, *d2-3*, *d2-4,* and *d2-6* mutants were not significantly different from their corresponding wild types. Tanaka et al. (2009) reported that *OsBU1* overexpression plants showed enhanced bending of the lamina joint and enlarged grain size, which was not consistent with the characteristics of *OsD2*-OE lines in the present study. However, Li et al. (2013a) observed a decreased expression of *OsBU1* in the *OsD2* null mutant (*csdd1*), which is consistent with our observations. In our study, BR synthesis genes *OsD11* and *OsBRD1* which are regulated by the BR signaling gene *RELATED TO ABI3/VP1-LIKE1* (*RAVL1*), showed significantly decreased expression in *OsD2*-KD lines (Supplementary Figure 6; Je et al., 2010). However, *OsD2*-OE lines displayed increased expression levels of *OsD11* and decreased expression levels of *OsBRD1*. Li et al. (2013a) reported that the *OsD2* null mutant *csdd1* showed increased expression levels of *OsD11* and *OsDWARF* (*OsBRD1*) compared to wild-type plants. This discrepancy in the expression pattern of genes involved in BR signaling and biosynthesis may be due to the use of heterozygous *OsD2*-KD and *OsD2*-OE plants.

Low-temperature stress is a challenge in crop cultivation. In this study, a new QTL, *qLTG1*, controlling LTG was detected in an introgression line (TR5) derived from a cross between the *japonica* variety "Hwaseong" and the wild rice (*O. rufipogon*). Using map-based cloning, transgenic approaches, and gene expression analysis, we demonstrated that the rice BR biosynthesis gene *OsD2* is the causal gene for *qLTG1*. The *O. rufipogon OsD2* allele improved LTG in the cultivar background without any deleterious phenotypes. These results will benefit rice breeding programs by assisting in developing strategies to maximize the exploitation of invaluable genes from interspecific or intraspecific crosses.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SK, K-CS, and S-NA designed the experiments and wrote the manuscript. SK and K-CS carried out phenotype analysis (LTG test). H-SL and Y-AJ performed qRT-PCR analysis. CA and NL conducted the agronomic traits investigation. All authors contributed to the article and approved the submitted version.

Funding

This work was carried out with the support of "Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ015757)" Rural Development Administration, Republic of Korea.

Acknowledgments

We thank Young-Sook Kim for technical assistance and seed management.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.985559/full#supplementary-material

References

Anuradha, S., and Rao, S. S. S. (2001). Effect of brassinosteroids on salinity stress induced inhibition of seed germination and seedling growth of rice (Oryza sativa L.). *Plant Growth Regul.* 33, 151–153. doi: 10.1023/A:1017590108484

Causse, M. A., Fulton, T. M., Cho, Y. G., Ahn, S. N., Chunwongse, J., Wu, K. S., et al. (1994). Saturated molecular map of the Rice genome based on an interspecific backcross population. *Genetics* 138, 1251–1274. doi: 10.1093/genetics/138.4.1251

Chen, L., Lou, Q., Sun, Z., Xing, Y., Yu, X., and Luo, L. (2006). QTL mapping of low temperature germinability in rice. *Zhongguo Shuidao Kexue* 20, 159–164.

Fang, N., Xu, R., Huang, L. J., Zhang, B. L., Duan, P. G., Li, N., et al. (2016). SMALL GRAIN 11 controls grain size, grain number and grain yield in rice. *Rice* 9:64. doi: 10.1186/s12284-016-0136-z

Fujino, K., Sekiguchi, H., Matsuda, Y., Sugimoto, K., Ono, K., and Yano, M. (2008). Molecular identification of a major quantitative trait locus, qLTG3-1, controlling low-temperature germinability in rice. *Proc. Natl. Acad. Sci. U. S. A.* 105, 12623–12628. doi: 10.1073/pnas.0805303105

Fujino, K., Sekiguchi, H., Sato, T., Kiuchi, H., Nonoue, Y., Takeuchi, Y., et al. (2004). Mapping of quantitative trait loci controlling low-temperature germinability in rice (Oryza sativa L.). *Theor. Appl. Genet.* 108, 794–799. doi: 10.1007/s00122-003-1509-4

Fujioka, S., and Yokota, T. (2003). Biosynthesis and metabolism of brassinosteroids. *Annu. Rev. Plant Biol.* 54, 137–164. doi: 10.1146/annurev.arplant.54.031902.134921

Hong, Z., Ueguchi-Tanaka, M., Umemura, K., Uozu, S., Fujioka, S., Takatsuto, S., et al. (2003). A rice brassinosteroid-deficient mutant, ebisu dwarf (d2), is caused by a loss of function of a new member of cytochrome P450. *Plant Cell* 15, 2900–2910. doi: 10.1105/tpc.014712

Hyun, D. Y., Lee, G. A., Kang, M. J., Burkart-Waco, D., Kim, S. I., Kim, J. Y., et al. (2015). Development of low-temperature germinability markers for evaluation of rice (Oryza sativa L.) germplasm. *Mol. Breed.* 35:104. doi: 10.1007/s11032-015-0298-1

Iwata, N., Shinada, H., Kiuchi, H., Sato, T., and Fujino, K. (2010). Mapping of QTLs controlling seedling establishment using a direct seeding method in rice. *Breed. Sci.* 60, 353–360. doi: 10.1270/jsbbs.60.353

Je, B. I., Piao, H. L., Park, S. J., Park, S. H., Kim, C. M., Xuan, Y. H., et al. (2010). RAV-Like1 maintains Brassinosteroid homeostasis via the coordinated activation of BRI1 and biosynthetic genes in Rice. *Plant Cell* 22, 1777–1791. doi: 10.1105/tpc.109.069575

Jeon, Y.-A., Kang, Y.-J., Shim, K.-C., Lee, H.-S., Xin, L., Kang, J.-W., et al. (2018). Genetic analysis and mapping of genes for culm length and internode diameter in progeny from an interspecific cross in rice. *Plant Breed. Biotechnol.* 6, 140–146. doi: 10.9787/PBB.2018.6.2.140

Jeon, J. S., Lee, S., Jung, K. H., Jun, S. H., Jeong, D. H., Lee, J., et al. (2000). T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* 22, 561–570. doi: 10.1046/j.1365-313x.2000.00767.x

Jeon, Y. A., Lee, H. S., Kim, S. H., Shim, K. C., Kang, J. W., Kim, H. J., et al. (2021). Natural variation in rice ascorbate peroxidase gene APX9 is associated with a yield-enhancing QTL cluster. *J. Exp. Bot.* 72, 4254–4268. doi: 10.1093/jxb/erab155

Jeong, D. H., An, S., Park, S., Kang, H. G., Park, G. G., Kim, S. R., et al. (2006). Generation of a flanking sequence-tag database for activation-tagging lines in japonica rice. *Plant J.* 45, 123–132. doi: 10.1111/j.1365-313X.2005.02610.x

Jiang, N. F., Shi, S. L., Shi, H., Khanzada, H., Wassan, G. M., Zhu, C. L., et al. (2017). Mapping QTL for seed Germinability under low temperature using a new high-density genetic map of Rice. *Front. Plant Sci.* 8:1223. doi: 10.3389/fpls.2017.01223

Kim, T. S., He, Q., Kim, K. W., Yoon, M. Y., Ra, W. H., Li, F. P., et al. (2016). Genome-wide resequencing of KRICE_CORE reveals their potential for future breeding, as well as functional and evolutionary studies in the post-genomic era. *BMC Genomics* 17, 1–13. doi: 10.1186/s12864-016-2734-y

Kim, T., Kang, K., Kim, S. H., An, G., and Paek, N. C. (2019). OsWRKY5 promotes Rice leaf senescence via senescence-associated NAC and Abscisic acid biosynthesis pathway. *Int. J. Mol. Sci.* 20. doi: 10.3390/ijms20184437

Kumar, V., and Ladha, J. K. (2011). Direct seeding of Rice: recent developments and future research needs. *Adv. Agron.* 111, 297–413. doi: 10.1016/B978-0-12-387689-8.00001-1

Leubner-Metzger, G. (2001). Brassinosteroids and gibberellins promote tobacco seed germination by distinct pathways. *Planta* 213, 758–763. doi: 10.1007/s004250100542

Li, H., Jiang, L., Youn, J. H., Sun, W., Cheng, Z. J., Jin, T. Y., et al. (2013a). A comprehensive genetic study reveals a crucial role of CYP90D2/D2 in regulating plant architecture in rice (Oryza sativa). *New Phytol.* 200, 1076–1088. doi: 10.1111/nph.12427

Li, L. F., Liu, X., Xie, K., Wang, Y. H., Liu, F., Lin, Q. Y., et al. (2013b). qLTG-9, a stable quantitative trait locus for low-temperature germination in rice (Oryza sativa L.). *Theor. Appl. Genet.* 126, 2313–2322. doi: 10.1007/s00122-013-2137-2

Miura, K., Lin, S. Y., Yano, M., and Nagamine, T. (2001). Mapping quantitative trait loci controlling low temperature germinability in rice (Oryza sativa L.). *Breed. Sci.* 51, 293–299. doi: 10.1270/jsbbs.51.293

Nguyen, H. N., Park, I. K., Yeo, S. M., Yun, Y. T., and Ahn, S. N. (2012). Mapping quantitative trait loci controlling low-temperature germinability in rice. *Kor. J. Agric. Sci.* 39, 477–482. doi: 10.7744/cnujas.2012.39.4.477

Sakamoto, T., Ohnishi, T., Fujioka, S., Watanabe, B., and Mizutani, M. (2012). Rice CYP90D2 and CYP90D3 catalyze C-23 hydroxylation of brassinosteroids in vitro. *Plant Physiol. Biochem.* 58, 220–226. doi: 10.1016/j.plaphy.2012.07.011

Satoh, T., Tezuka, K., Kawamoto, T., Matsumoto, S., Satoh-Nagasawa, N., Ueda, K., et al. (2016). Identification of QTLs controlling low-temperature germination of the east European rice (*Oryza sativa* L.) variety Maratteli. *Euphytica* 207, 245–254. doi: 10.1007/s10681-015-1531-z

Shim, K. C., Kim, S., Le, A. Q., Lee, H. S., Adeva, C., Jeon, Y. A., et al. (2019). Fine mapping of a low-temperature Germinability QTL qLTG1 using introgression lines derived from Oryza rufipogon. *Plant Breed. Biotechnol.* 7, 141–150. doi: 10.9787/PBB.2019.7.2.141

Shim, K. C., Kim, S. H., Lee, H. S., Adeva, C., Jeon, Y. A., Luong, N. H., et al. (2020). Characterization of a new qLTG3-1 allele for low-temperature Germinability in Rice from the wild species Oryza rufipogon. *Rice* 13:10. doi: 10.1186/s12284-020-0370-2

Steber, C. M., and Mccourt, P. (2001). A role for brassinosteroids in germination in Arabidopsis. *Plant Physiol.* 125, 763–769. doi: 10.1104/pp.125.2.763

Tanaka, A., Nakagawa, H., Tomita, C., Shimatani, Z., Ohtake, M., Nomura, T., et al. (2009). BRASSINOSTEROID UPREGULATED1, encoding a helix-loop-helix protein, is a novel gene involved in Brassinosteroid signaling and controls bending of the lamina joint in Rice. *Plant Physiol.* 151, 669–680. doi: 10.1104/pp.109.140806

Vardhini, B. V., and Rao, S. S. R. (2003). Amelioration of osmotic stress by brassinosteroids on seed germination and seedling growth of three varieties of sorghum. *Plant Growth Regul.* 41, 25–31. doi: 10.1023/A:1027303518467

Wang, B. L., Zhang, J. L., Xia, X. Z., and Zhang, W. H. (2011). Ameliorative effect of brassinosteroid and ethylene on germination of cucumber seeds in the presence of sodium chloride. *Plant Growth Regul.* 65, 407–413. doi: 10.1007/s10725-011-9595-9

Wang, X., Zou, B., Shao, Q., Cui, Y., Lu, S., Zhang, Y., et al. (2018). Natural variation reveals that OsSAP16 controls low-temperature germination in rice. *J. Exp. Bot.* 69, 413–421. doi: 10.1093/jxb/erx413

Wu, C. Y., Trieu, A., Radhakrishnan, P., Kwok, S. F., Harris, S., Zhang, K., et al. (2008). Brassinosteroids regulate grain filling in rice. *Plant Cell* 20, 2130–2145. doi: 10.1105/tpc.107.055087

Xiao, J. H., Grandillo, S., Ahn, S. N., Mccouch, S. R., Tanksley, S. D., Li, J. M., et al. (1996). Genes from wild rice improve yield. *Nature* 384, 223–224. doi: 10.1038/384223a0

Yang, T. F., Zhou, L., Zhao, J. L., Dong, J. F., Liu, Q., Fu, H., et al. (2020). The candidate genes underlying a stably expressed QTL for low temperature Germinability in Rice (Oryza sativaL.). *Rice* 13:74. doi: 10.1186/s12284-020-00434-z

Yeo, S. M., Yun, Y. T., Kim, D. M., Chung, C. T., and Ahn, S. N. (2014). Validation of QTLs associated with spikelets per panicle and grain weight in rice. *Plant Genet. Res. Character. Utiliz.* 12, S151–S154. doi: 10.1017/S1479262114000562

Yun, Y. T., Chung, C. T., Lee, Y. J., Na, H. J., Lee, J. C., Lee, S. G., et al. (2016). QTL mapping of grain quality traits using introgression lines carrying O rufipogon chromosome segments in japonica Rice. *Rice* 9:62. doi: 10.1186/s12284-016-0135-0

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 October 2022
DOI 10.3389/fpls.2022.1005203

OPEN ACCESS

EDITED BY
Ryuji Ishikawa,
Hirosaki University, Japan

REVIEWED BY
Tian Qing Zheng,
Institute of Crop Sciences (CAAS) China
Wei Zong,
Indiana University, United States
Jianmin Bian,
Jiangxi Agricultural University, China
Xin-Yuan Huang,
Nanjing Agricultural University, China
Yoshiaki Ueda,
Japan International Research Center for Agricultural Sciences (JIRCAS), Japan
Jiang Shukun,
Heilongjiang Academy of Agricultural Sciences, China

*CORRESPONDENCE
Dali Zeng
dalizeng@126.com
Qian Qian
qianqian188@hotmail.com
Qing Li
liqing1986102@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Bioinformatics, a section of the journal Frontiers in Plant Science

RECEIVED 28 July 2022
ACCEPTED 13 October 2022
PUBLISHED 26 October 2022

CITATION
Dai L, Lu X, Shen L, Guo L, Zhang G, Gao Z, Zhu L, Hu J, Dong G, Ren D, Zhang Q, Zeng D, Qian Q and Li Q (2022) Genome-wide association study reveals novel QTLs and candidate genes for seed vigor in rice. *Front. Plant Sci.* 13:1005203. doi: 10.3389/fpls.2022.1005203

Genome-wide association study reveals novel QTLs and candidate genes for seed vigor in rice

Liping Dai[1†], Xueli Lu[1†], Lan Shen[1], Longbiao Guo[1], Guangheng Zhang[1], Zhenyu Gao[1], Li Zhu[1], Jiang Hu[1], Guojun Dong[1], Deyong Ren[1], Qiang Zhang[1], Dali Zeng[1,2*], Qian Qian[1*] and Qing Li[1*]

[1]State Key Laboratory for Rice Biology, China National Rice Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou, China, [2]The Key Laboratory for Quality Improvement of Agricultural Products of Zhejiang Province, College of Advanced Agricultural Sciences, Zhejiang A & F University, Hangzhou, China

Highly seed vigor (SV) is essential for rice direct seeding (DS). Understanding the genetic mechanism of SV-related traits could contribute to increasing the efficiency of DS. However, only a few genes responsible for SV have been determined in rice, and the regulatory network of SV remains obscure. In this study, the seed germination rate (GR), seedling shoot length (SL), and shoot fresh weight (FW) related to SV traits were measured, and a genome-wide association study (GWAS) was conducted to detect high-quality loci responsible for SV using a panel of 346 diverse accessions. A total of 51 significant SNPs were identified and arranged into six quantitative trait locus (QTL) regions, including one (*qGR1-1*), two (*qSL1-1*, *qSL1-2*), and three (*qFW1-1*, *qFW4-1*, and *qFW7-1*) QTLs associated with GR, SL, and FW respectively, which were further validated using chromosome segment substitution lines (CSSLs). Integrating gene expression, gene annotation, and haplotype analysis, we found 21 strong candidate genes significantly associated with SV. In addition, the SV-related functions of *LOC_Os01g11270* and *LOC_Os01g55240* were further verified by corresponding CRISPR/Cas9 gene-edited mutants. Thus, these results provide clues for elucidating the genetic basis of SV control. The candidate genes or QTLs would be helpful for improving DS by molecular marker-assisted selection (MAS) breeding in rice.

KEYWORDS

rice, GWAS, seed vigor, direct seeding, haplotype

Introduction

Rice direct seeding (DS) technology, in which seeds are directly sown in the field without seedling raising or transplanting, provides an advanced and simplified cultivation technology for mechanized rice production. With the decrease in the rural labor force, the rise in labor cost, and the development of mechanized rice planting technology, rice planting in Asia, the main rice-producing area, has a trend of developing from traditional or mechanical transplanting to mechanical direct planting due to its labor- and cost-saving. However, DS requires the strong seed vigor (SV) of cultivated rice varieties, including high and uniform seed germination rate, strong seedling establishment and growth capacity, and enhanced stress tolerances. Most transplanting rice varieties on the market, including conventional rice, hybrid rice, and super rice, cannot be used for DS due to their poor SV. Thus, it is urgent to develop high SV varieties suitable for rice mechanized direct seeding through molecular-assisted selective breeding and biotechnology breeding, which depends on understanding the regulatory mechanism of SV-related traits and identifying the essential genes underlying these traits.

SV is a comprehensive trait affected by genetic and environmental factors during seed development and maturation. At present, a large number of QTLs for SV-related characteristics, such as seed germination rate (Wang et al., 2010; Jiang et al., 2020; Najeeb et al., 2020), coleoptile/seedling length (Dang et al., 2014; Xie et al., 2014; Jin et al., 2018), and seedling weight (Zhang et al., 2017; Najeeb et al., 2020), have been found in rice. Most of these QTLs were identified by bi-parental mapping with traditional markers. However, due to the low density of traditional markers, only a few QTLs or genes for SV have been fine-mapped. For instance, Fujino et al. isolated *qLTG3-1*, a QTL associated with low-temperature germinability. It may accelerate vacuolization and weaken the tissues covering the embryo during seed germination, thereby reducing the mechanical resistance to the growth potential of coleoptile (Fujino et al., 2008). In addition, *qSE3*, a major QTL for rice seed germination and seedling establishment under salinity stress, was identified as a K^+ transporter OsHAK21 (He et al., 2019a). Physiological analysis revealed that *qSE3* improved seed germination and seedling establishment against salinity stress through mediating seed physiological states, including increasing K^+ and Na^+ uptake, promoting abscisic acid (ABA) biosynthesis, and reducing reactive oxygen species (ROS) accumulation. Very recently, *OsHIPL1*, which encodes a hedgehog-interacting protein-like 1 protein, was cloned as a causal gene for the major QTL *qSV3* associated with rice SV (He et al., 2022). And *OsHIPL1* may play its role by modulating ABA levels and *OsABIs* expression during rice seed germination.

Genome-wide association study (GWAS) has become an effective way to dissect complex traits and identify corresponding loci or candidate genes. Several QTLs or genes for SV have been detected *via* GWAS in rice. For example, four germination rate (GR)-related QTLs in rice were revealed by GWAS, and *OsOMT*, encoding the 2-oxoglutarate/malate translocator, was further validated as a causal gene for *qGR11*. *OsOMT* played an essential role in GR by modulating amino acid levels, and the processes of glycolysis and tricarboxylic acid cycle (Li et al., 2021a). Yang et al. identified 19 QTLs associated with rice seed germination by GWAS. They further found that *OsPK5* was a positive regulator of rice seed germination by encoding a pyruvate kinase and thereby modulating glycolytic metabolism and abscisic (ABA)/gibberellins (GA) balance (Yang et al., 2022). In addition, *OsCDP3.10*, encoding a cupin domain protein, was identified from *qSP3*, a significant QTL for rice seedling percentage by GWAS (Peng et al., 2022). Functional analysis revealed that *OsCDP3.10* affected rice SV mainly by modulating the levels of amino acids and H_2O_2 in the germinating seeds. These studies demonstrated that GWAS is an effective way to identify QTLs/genes for SV in rice.

However, so far, almost all studies on GWAS related to SV have been carried out in Petri dishes and under constant temperature conditions, which cannot fully simulate the natural environment of the field. Therefore, the present study intends to apply GWAS to identify QTLs/genes for rice SV in the field. We surveyed the GR, seedling shoot length (SL), and shoot fresh weight (FW) of 346 rice accessions in the field, and performed GWAS analysis to identify reliable loci responsible for rice SV. In total, six QTLs were detected, including one for GR, two for SL, and three for FW. Combining gene expression, gene annotation, and haplotype analysis, we screened 21 candidate genes significantly associated with SV from these QTLs, among which *LOC_Os01g11270* and *LOC_Os01g55240* were further validated by corresponding CRISPR/Cas9 gene-edited mutants. The detected QTLs and candidate genes will enhance our understanding of the genetic and molecular basis for SV and provide valuable resources for applying these potential elite loci in rice molecular breeding.

Materials and methods

Plant material and phenotyping

The 346 representative accessions distributed worldwide were selected for the GWAS panel, including 213 *indica* and 133 *japonica* rice accessions (Supplementary Table 1). All seeds of these accessions were harvested at their maturity stage in September and October 2020, dried, and stored in a refrigeration house. To break seed dormancy, these seeds were dried at 42 °C for seven days before sowing (He et al., 2019b; He et al., 2022). When seeds are sown, the waterlogged paddy field is silty and sticky without standing water. In May 2021, 50 well-filled seeds

of each accession were evenly sown in a circular box (6.4 cm with a diameter) on the surface of the paddy field at China National Rice Research Institute (Hangzhou, Zhejiang, China), following a randomized complete block experimental design with three replications. One week after sowing, we began to slowly saturate the paddy field, maintaining a thin layer of water with a depth of 1-2 cm on the soil's surface. After direct sowing for two weeks, the data of GR, SL, and FW were measured. The temperature during this period was 17 °C ~ 26 °C. The seed was regarded as germination when the coleoptile or radicle length was longer than 1 cm, and GR was defined as the ratio of the number of germinated seeds to the total number of seeds. Five plants for each accession were randomly selected to measure the SL and FW (shoot fresh weight of all five plants). The mean data of GR, SL, and FW were presented and obtained from the three replicates, including 344, 314, and 313 mean values for GR, SL, and FW, respectively (Supplementary Table 1). The lack of mean GR data for two accessions was due to the large SD of three replicates, which was considered inaccurate and consequently discarded. Similarly, the mean FW value of D188 accession, and SL and FW values of D299 accession were also discarded. In addition, to eliminate the influence of germination time on SL and FW, the SL and FW of 31 accessions with low GR (<38%) and shallow emergence rate one week after sowing were not recorded. These missing GR, SL, and FW data were marked as NA in Supplementary Table 1.

Chromosome segment substitution lines (CSSLs) used for *qGR1-1*, *qSL1-1*, *qSL1-2*, *qFW1-1*, *qFW4-1*, and *qFW7-1* validation were constructed by a combination of backcross and molecular maker-assisted selection (Dai et al., 2020). After backcrossing four times, each CSSL has a genetic background with recurrent receptor parent 9311 and only carries a target chromosomal segment from the donor parent Nipponbare (NPB). After drying at 42 °C for seven days, 50 well-filled seeds per replicate of 9311 and CSSLs were directly sown in ddH_2O at 30 °C for four days, then transplanted in a conventional Kimura B hydroponic medium at 30 °C for six days (maintain culture in ddH_2O for GR). The GR, SL, and FW data for each CSSL and its recurrent parent 9311 were collected ten days after cultivation. Three biological replicates were performed for GR and FW, and ten replicates were used for SL.

The CRISPR/Cas9-mediated gene editing mutants for *LOC_Os01g11270* and *LOC_Os01g55240* were obtained from Weimi Biotechnology Co., Ltd. (Lu et al., 2017). And the cultivation and measurement methods are the same as those of the CSSLs.

Genome resequencing and SNP genotyping

The genomic DNA of each accession for the GWAS population was extracted by the cetyltrimethylammonium bromide (CTAB) method (Murray and Thompson, 1980). The resequencing library was constructed according to the manufacturer's instructions (Illumina, San Diego, CA, USA). The paired-end reads in each library were generated by Illumina NovaSeq 6000 platform at Berry Genomics Company (http://www.berrygenomics.com/, Beijing, China). The clean reads were aligned to the NPB reference genome (IRGSP-1.0) with Burrows-Wheeler Aligner (BWA) (Li and Durbin, 2009). Genome Analysis Toolkit software (GATK) (Van der Auwera et al., 2013) was used for single nucleotide polymorphism (SNP) calling. The SNP annotations were analyzed by SnpEff (Cingolani et al., 2012) based on the GFF3 files of the NPB reference genome. A total of 2,748,212 high-quality SNPs with minor allele frequency (MAF) greater than 5% and a missing rate less than 20% were screened for further GWAS.

Population genetics analysis

We randomly selected 100,000 SNPs with a MAF > 5% and a missing rate < 20% to build a phylogenetic tree using SNPhylo (Lee et al., 2014; Li et al., 2021b; Feng et al., 2022), and plotted it with the R package of 'ggtree' (Yu et al., 2017). The Principal component analysis (PCA) and population structure analysis for all accessions were calculated by Plink (Purcell et al., 2007) and ADMIXTURE (Alexander et al., 2009).

Genome-wide association study

GWAS was conducted using vcf2gwas with a linear mixed model (Vogt et al., 2022). GR, SL, and FW values for 346 rice accessions were used as input data. Based on the Bonferroni-corrected method and previous reports (Yang et al., 2014; Li et al., 2021b; Feng et al., 2022), the *P-value* of the whole-genome significance cutoff was set as 1/n (n = the number of SNPs used in association analysis); that is, the threshold of $-\log_{10}(P)$ was about 6.4. Manhattan and quantile-quantile (Q-Q) plots were generated using the R package of Cmplot (https://github.com/YinLiLin/CMplot). Based on the previous report, the 200-kb upstream and downstream significant SNP site was regarded as a QTL region, and adjacent overlapped regions were merged into the same QTL (Li et al., 2022).

Haplotype analysis

The SNPs in the regions of 2-kb upstream, CDS, and 1-kb downstream sequence of each gene were selected for haplotype analysis. The rare haplotype with a frequency lower than ten were discarded. The top two or three haplotypes with the highest frequency were chosen for further phenotypic association analysis.

Result

Phenotypic variation of SV-related traits in 346 rice accessions

We observed significant variations of three SV-related traits (GR, SL, and FW) among 346 rice accessions. In detail, the GR ranged from 6.00 to 100.00%, with an average of 69.74% and a variation coefficient of 28.69% (Table 1). The SL changed from 8.40 to 26.20 cm, with a mean value of 15.42 cm and a variation coefficient of 22.65%. The FW varied from 0.1325 to 0.4800 g, with an average of 0.2592 g and a variation coefficient of 23.01%. These results indicate an extensive phenotypic variation (PV) in our experimental population. Furthermore, according to the kurtosis and skewness values, all traits showed skewed distribution, especially GR (Figures 1A, C, E; Table 1). We further analyzed the SV-related values in *indica* and *japonica* subpopulations. It revealed that the GR, SL, and FW of *indica* and *japonica* were 69.94% and 69.42%, 16.4 cm and 14.0 cm, 0.2722 g and 0.2391 g, respectively. Except for GR, a significant difference was observed for SL and FW between *indica* and *japonica* (Figures 1B, D, F). In sum, the GR, SL, and FW traits were suitable for subsequent GWAS mapping.

Population structure of 346 rice accessions

The 346 rice accessions were classified into two subgroups based on the phylogenetic tree, roughly corresponding to the classification of *indica* and *japonica* (Figure 2A). Similarly, PCA and the population structure base on the possible clusters (k) method also showed that these rice accessions could be divided into two subgroups as the phylogenetic tree classification (Figures 2B, C). Taken together, the category of population structure has a similar result compared to a previous study (Shi et al., 2021; Li et al., 2022).

Mining loci associated with SV by GWAS

To explore the novel loci related to SV, we performed GWAS on GR, SL, and FW in 346 rice accessions, respectively. The Manhattan plots of GWAS showed 18, 30, and 3 significant SNPs were associated with GR, SL, and FW, respectively, and assigned as six QTLs (*qGR1-1*, *qSL1-1*, *qSL1-2*, *qFW1-1*, *qFW4-1*, and *qFW7-1*) according to their physical position in the genome (Figure 3; Table 2 and Supplementary Table 2). For GR and SL, 18, 2, and 28 significant SNPs were detected in *qGR1-1*, *qSL1-1*, and *qSL1-2* on chromosome 1, contributing 7.76 ~ 10.53%, 7.92%, and 7.78 ~ 9.33% of the phenotypic variation explanation (PVE), respectively (Table 2, Supplementary Table 2). Whereas FW, *qFW1-1*, *qFW4-1*, and *qFW7-1* located on chromosomes 1, 4, and 7 respectively, all contain only one associated SNP that explains 7.98 ~ 9.28% of the PVE (Table 2, Supplementary Table 2).

Validation of GWAS-associated loci using CSSLs

To verify the GWAS-associated QTLs, we screened five CSSLs from the CSSL library with 9311 as recurrent parent and NPB as donor parent. We named them CSSL14 (containing *qGR1-1*), CSSL16 (containing *qSL1-1* and *qSL1-2*), CSSL3 (containing *qFW1-1*), CSSL18 (containing *qFW4-1*), and CSSL15 (containing *qFW7-1*), respectively (Figure 4A). A comparison of SV-related traits between 9311 and corresponding CSSLs revealed significant differences in GR, SL, and FW, respectively. In detail, the GR of CSSL14 increased by 7.34% compared with that of 9311 (Figure 4B), while the SL of CSSL16 and FW of CSSL3, CSSL18, and CSSL15 decreased by 12.86%, 13.27%, 29.50%, and 22.08% compared with that of 9311, respectively (Figures 4C, D). These data suggest that the above GWAS results are reliable and imply that the NPB allele is the dominant allele of *qGR1-1*, whereas the allele from 9311 is the superior allele for *qSL1-1*, *qSL1-2*, *qFW1-1*, *qFW4-1*, and *qFW7-1*.

Identification of SV-related candidate genes in GWAS-associated loci

To search for the candidate genes in six QTLs, we excavated all 540 annotated genes, including 135, 60, 169, 59, 58, and 59 genes in *qGR1-1*, *qSL1-1*, *qSL1-2*, *qFW1-1*, *qFW4-1*, and *qFW7-1*, respectively (Supplementary Table 3). We used three strategies to narrow down the candidate genes: Firstly, 76 genes not expressed in seeds (for GR) or seeds and leaves (for SL and FW) were deleted based on the expression profile in Rice Expression Database

TABLE 1 The SV-related traits in 346 rice accessions.

Phenotype	Mean	Range	Coefficient of variation (%)	Kurtosis	Skewness
GR	69.74%	6.00-100.00%	28.69	0.22	-0.91
SL	15.42 cm	8.40-26.20 cm	22.65	0.34	0.68
FW	0.2592 g	0.1325-0.4800 g	23.01	0.30	0.51

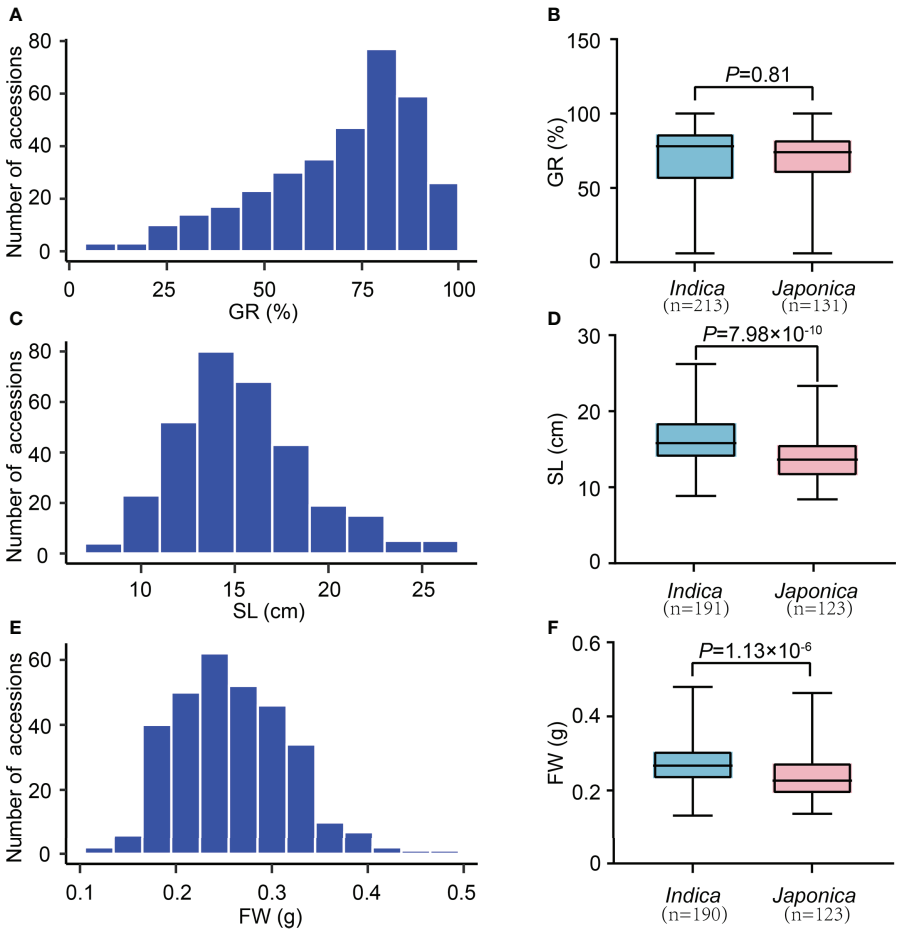


FIGURE 1
The variation of SV-related traits in 346 rice accessions. **(A)** Distribution of GR values. **(C)** Distribution of SL values. **(E)** Distribution of FW values. **(B)** Comparison of GR values between *indica* and *japonica*. **(D)** Comparison of SL values between *indica* and *japonica*. **(F)** Comparison of FW values between *indica* and *japonica*. n indicates the number of accessions. The p-value is obtained from the T-test.

(http://expression.ic4r.org/) (Supplementary Table 3). Secondly, 65 genes encoding retrotransposon or transposon protein were excluded based on their function annotations (Supplementary Table 3). Meanwhile, 28 genes considered highly related to the SV phenotype were screened from the remaining genes (Supplementary Table 3). The common feature of these 28 genes is that they and their homologs are at least associated with SV-related traits or involved in hormone metabolism or signaling pathways such as ABA and GA. Lastly, haplotype analysis was carried out using the SNPs in 346 accessions, and 21 candidate genes associated with significant SV-related phenotypic differences between their top two or three haplotypes were screened (Supplementary Tables 1, 3). Their distances from the nearest significant SNPs are shown in Supplementary Table 4. However, it is important to note that the phenotypic differences between haplotypes of some candidate genes may result from the phenotypic differences between *indica* and *japonica* subpopulations, since SL and FW differences exist in the two subspecies.

Among them, *LOC_Os01g11270* located in *qGR1-1* was annotated as a cytochrome P450 (CYP450) (Supplementary Table 3), and its homologs played a crucial role in hormone metabolisms, such as gibberellin (GA) (Zhu et al., 2005; Wu et al., 2014) and abscisic acid (ABA) (Zhu et al., 2009; Zhang et al., 2020). GA and ABA are two antagonistic phytohormones with positive and negative regulatory effects on seed germination and seedling establishment, respectively. Therefore, *LOC_Os01g11270* is a strong candidate associated with GR. Haplotype analysis revealed that *LOC_Os01g11270* has two major haplotypes according to an SNP located in its promoter (Figure 5A). Haplotype 1 (Hap1, reference sequence) was mainly found in *japonica* accessions, while Hap2 only occurred in *indica* accessions (Figure 5A; Supplementary Table 1). Comparatively, Hap1 had a lower GR than Hap2 (Figure 5B). To further validate

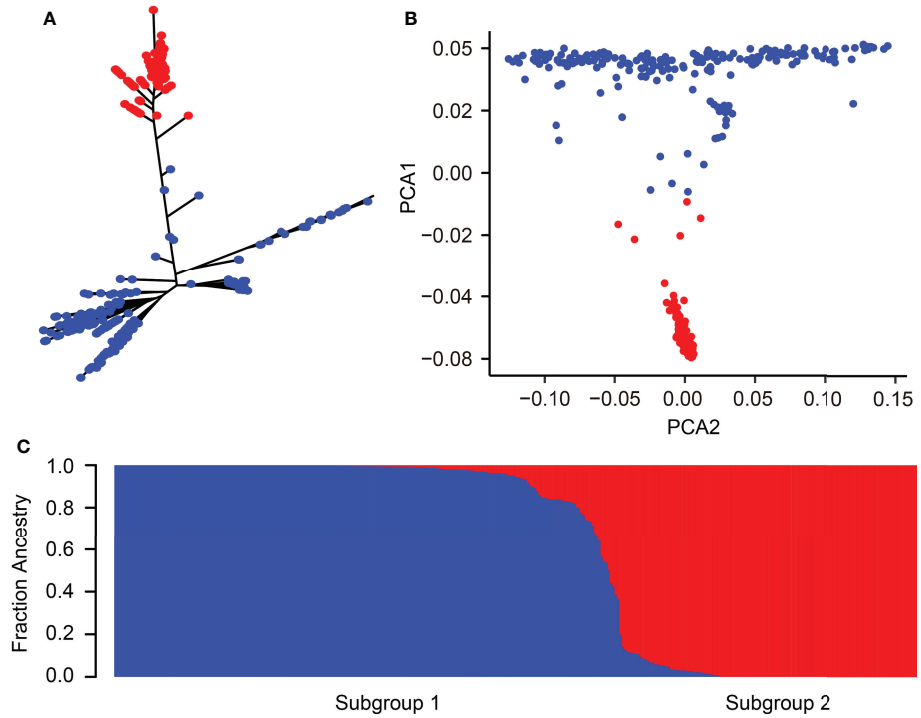


FIGURE 2
Population structure of 346 rice accessions. **(A)** Phylogenetic tree of 346 rice accessions. **(B)** PCA of 346 rice accessions. The first and second principal components are represented by PCA1 and PCA2, respectively. **(C)** Cluster analysis of 346 rice accessions. Blue and red colors correspond to *indica* and *japonica* accessions, respectively.

the role of *LOC_Os01g11270* in rice seed germination, we compared the GR of the *LOC_Os01g11270* gene-edited mutant (KO#270) and its wild-type Zhonghua 11 (ZH11). In KO#270, an 8-bp deletion at base 799 to 806 in the *LOC_Os01g11270* CDS caused a frameshift mutation (Figure 5C). The statistical result showed that the GR of KO#270 decreased by 15% compared to that of ZH11 (Figures 5C, D), suggesting a positive regulation effect of *LOC_Os01g11270*.

LOC_Os01g55240 in *qSL1-2* was identified as a GA inactivation gene (Supplementary Table 3), encoding a gibberellin 2-beta-dioxygenase (OsGA2ox3) (Tong et al., 2014), and thus may control seedling growth. Haplotype analysis of *OsGA2ox3* revealed two major haplotypes, Hap1 (reference sequence) and Hap2, that existed in 346 rice accessions (Figure 6A). The Hap1 associated with short SL was mainly found in *japonica* accessions, whereas Hap2 related to long SL only occurred in *indica* accessions (Figures 6A, B). To know whether the differential gene expression levels cause the SL variation in different haplotypes, we compared the transcripts of *OsGA2ox3* between Hap1 and Hap2 accessions using the leaf tissues (the uppermost complete leaves from four 3-week-old plants) RNA-seq dataset from our laboratory. The result showed that there was no significant difference between the average gene expression levels of Hap1 and Hap2 (Figure 6C; Supplementary Table 1), suggesting that the SL difference between Hap1 and Hap2 was not caused by their differential gene expression levels. To further validate the effect of *OsGA2ox3* in rice seedling growth, we compared the SL between the *OsGA2ox3* gene-edited mutant (KO#240) and its wild-type ZH11. In KO#240, one bp insertion between base 104 to 105 in the *OsGA2ox3* CDS resulted in a frameshift mutation (Figure 6D). It showed that the SL of KO#240 was 15% longer than that of the wild type (Figures 6D, E), suggesting a negative regulation effect of *OsGA2ox3*.

Discussion

Seed vigor is crucial in determining the yield of direct seeding rice. Good SV contains strong and uniform seed germination, rapid seedling growth, and increased stress tolerance. In contrast, apart from without these advantages, the poor SV also leads to weed bloom due to the weak seedling competitiveness, which further reduces rice production (Rajjou et al., 2012). Therefore, deciphering the molecular regulatory mechanisms of SV, and identifying the causative genes underlying SV have been the focus of rice biology. In the past decade, GWAS has been widely used to detect the potential loci or alleles underlying complex agronomic traits in rice. This study used three indicators, GR, SL, and FW,

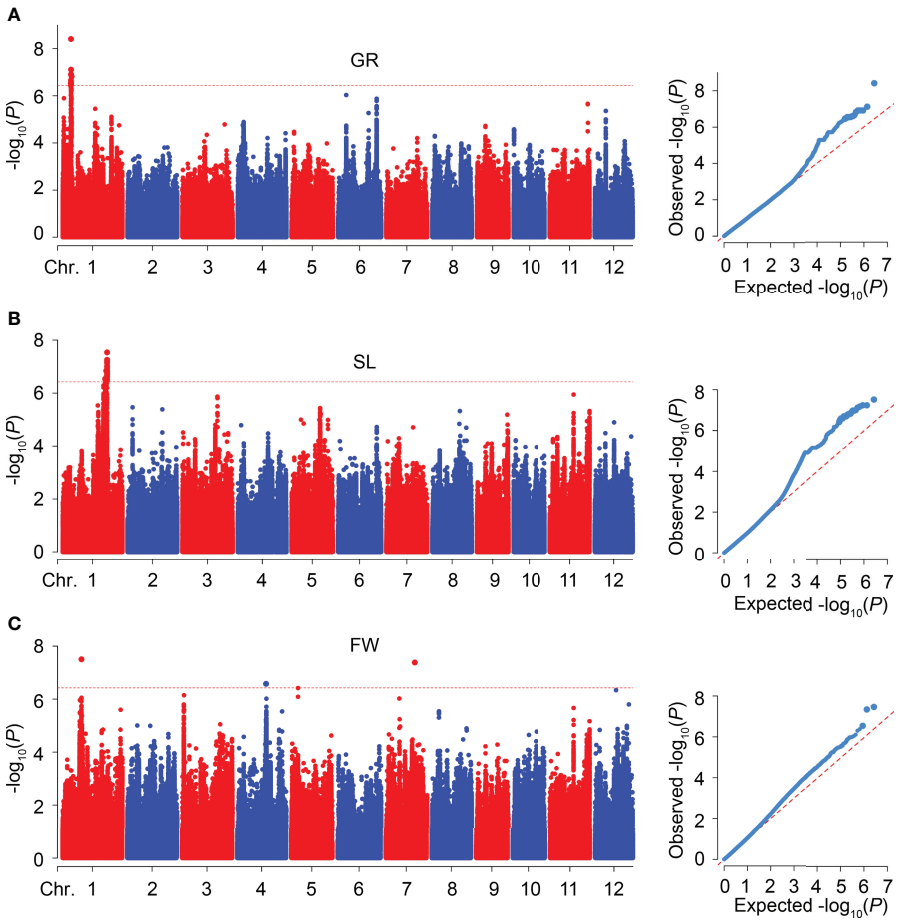


FIGURE 3
GWAS for SV-related traits in 346 rice accessions. **(A-C)** Manhattan and quantile-quantile (Q-Q) plots of GWAS for GR **(A)**, SL **(B)**, and FW **(C)**, respectively.

closely related to seed germination and seedling growth, to evaluate SV and perform GWAS. The results revealed a considerable variation and nearly normal distribution of GR, SL, and FW in 346 accessions, respectively (Figures 1A, C, E; Table 1). In addition, the *indica* accessions may be more suitable for DS because they had larger SL and FW than *japonica* accessions (Figures 1D, F). A total of six QTLs (*qGR1-1* for GR, *qSL1-1* and *qSL1-2* for SL, *qFW1-1*, *qFW4-1*, and *qFW7-1* for FW) were identified through GWAS using the whole population (Figure 3; Table 2). Among them, we found two QTLs co-located with previously reported SV-related QTLs using the traditional mapping method, suggesting that our GWAS results are pretty

TABLE 2 Six GWAS regions associated with SV.

QTL	Chr.	Region (nt)	Significant SNPs	Lead SNP		Co-location loci	References
				Position (nt)	-log10 (*P*)		
qGR1-1	1	5,689,847-6,629,529	18	5,990,942	8.41	*QAlGC1.2*	Janhan et al., 2021
qSL1-1	1	30,535,752-30,935,754	2	30,735,752	6.53		
qSL1-2	1	31,306,390-32,339,392	28	32,114,784	7.54	*qASRI1c*	Islam et al., 2022
qFW1-1	1	13,379,050-13,779,050	1	13,579,050	7.50	*qSWW1b*	Zhang et al., 2017
qFW4-1	4	20,256,553-20,656,553	1	20,456,553	6.58		
qFW7-1	7	20,402,053-20,802,053	1	20,602,053	7.38		

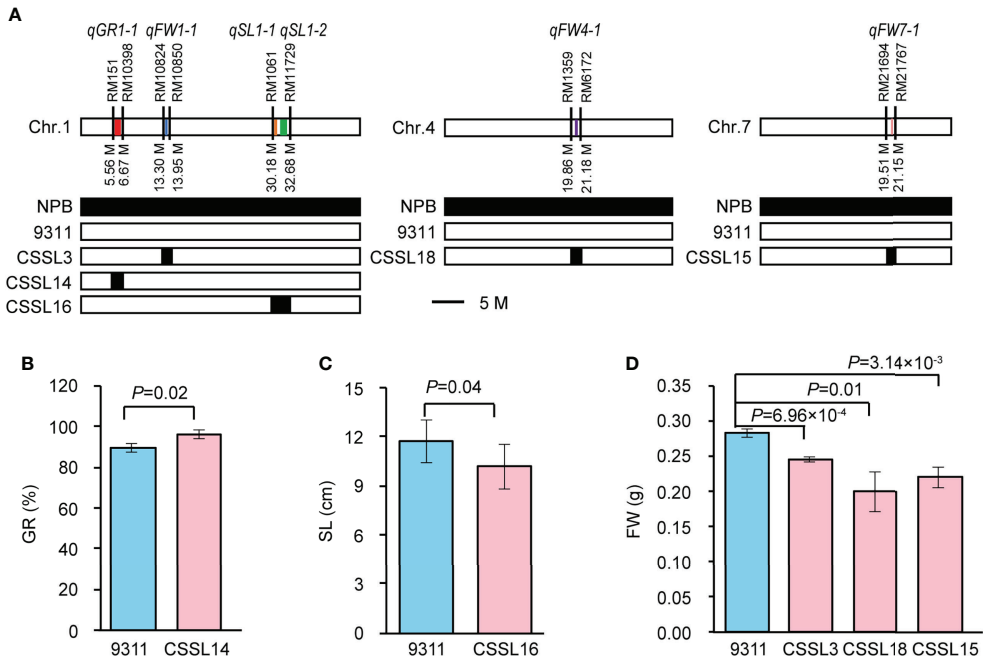


FIGURE 4
Validation of six GWAS-associated QTLs using CSSLs. **(A)** The genotype diagrams of CSSLs. *qGR1-1*, *qSL1-1*, *qSL1-2*, *qFW1-1*, *qFW4-1*, *qFW7-1* are represented by red, orange, green, blue, purple, and yellow lines, respectively. Line width indicates the size of QTL interval. **(B)** Comparison of GR between CSSL14 and 9311. Data represent mean ± SD (n=3). **(C)** Comparison of SL between CSSL16 and 9311. Data represent mean ± SD (n=10). **(D)** Comparison of FW among CSSL3, CSSL18, CSSL15, and 9311. Data represent mean ± SD (n=3). n indicates the number of replicates. The p-value is obtained from the T-test.

reliable (Table 2). For instance, *qGR1-1* corresponds to *QAlGC1.2*, a QTL for germination capacity under aluminum stress and mapped between markers RM14 and RM243 (Jahan et al., 2021). *qFW1-1* shares the same location with *qSWW1b*, a major QTL for the wet seedling weight (Zhang et al., 2017). In addition, we also found *qSL1-2* co-located with previously reported *qASRI1c*, a QTL for anaerobic salt response index (ASRI) of coleoptile identified by GWAS (Islam et al., 2022). However, other GWAS signals do not overlap with the previous GWAS results for SV-related traits (Li et al., 2021a; Peng et al., 2022; Yang et al., 2022). Besides the different rice accessions we used, one possible reason for the different results of GWAS is that phenotypic statistics of previous GWAS were mainly performed in Petri dishes and under constant temperature conditions. In contrast, we conducted phenotypic statistics in field conditions. In this study, we further used CSSLs to validate the GWAS-associated loci, and this strategy is relatively rare in GWAS studies.

By scanning the 540 annotated genes in six QTLs, we screened 21 strong candidate genes associated with SV (Supplementary Table 3). Based on functional annotations and literature review, these genes can be broadly classified into hormone homeostasis and signal-related, CYP450s, calcium signal-related, transcription factors, and so on. Here, we chose some representative ones for discussion. For example, *LOC_Os01g11150* (*OsGA2ox7*), *LOC_Os01g55240* (*OsGA2ox3*), *LOC_Os07g34240*, and *LOC_Os07g34370* were annotated as gibberellin 2-beta-dioxygenase or putative gibberellin receptor (Supplementary Table 3). Therefore, they may influence SV by regulating GA activity or signaling pathway. In fact, *OsGA2ox3* is mainly expressed in seedlings, and the *OsGA2ox3*-activated mutant or overexpression transgenic plant was severely dwarfed and maintained a vegetative growth (Lo et al., 2008; Takehara et al., 2020; Hsieh et al., 2021). In contrast, *OsGA2ox3* CRISPR/Cas9 knockout mutant showed a taller plant height due to elongated internodes and leaves (Takehara et al., 2020; Hsieh et al., 2021). These results suggest that *OsGA2ox3* is a negative regulator of plant height. Similarly, we also confirmed that *OsGA2ox3* down-regulated SL using *OsGA2ox3* CRISPR/Cas9 knockout mutant (Figures 6D, E). Furthermore, we found that the Hap2 in *indica* accessions was the dominant haplotype compared to Hap1 in *japonica* accessions (Figures 6A, B), implicating a great potential for improving SV-related traits in *japonica* accessions by *OsGA2ox3*. Compared with *OsGA2ox3*, *OsGA2ox7* in *qGR1-1* is an attenuated functional gene since its overexpression transgenic plants exhibited a semi-dwarf phenotype, whereas no phenotypic change on plant height was observed in *osga2ox7* knockout mutant (Hsieh et al., 2021). Here, we found apparent differences in GR and SL between the

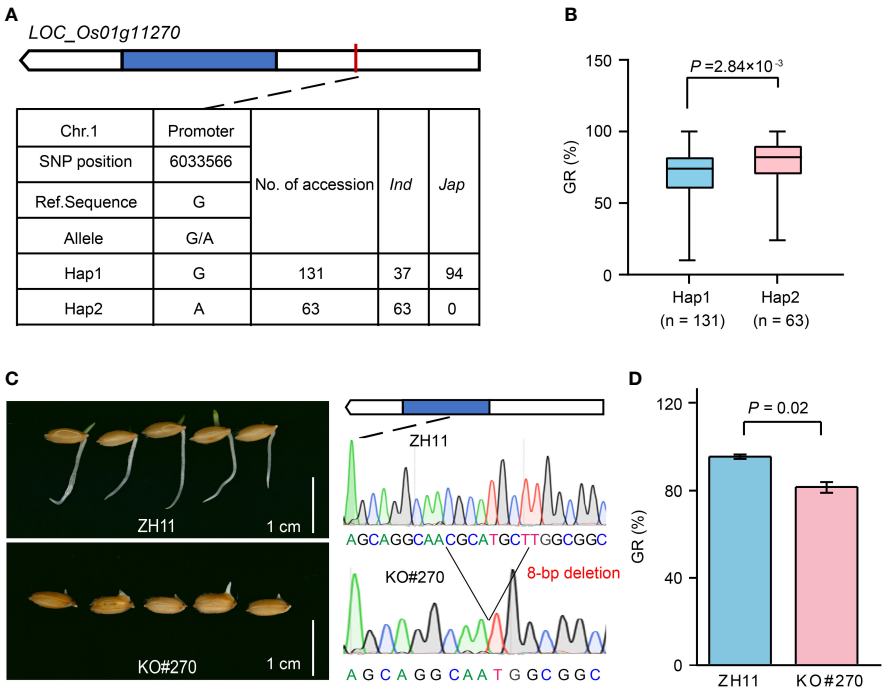


FIGURE 5
Effect of *LOC_Os01g11270* on rice seed germination. **(A)** The top two haplotypes of *LOC_Os01g11270* in 346 rice accessions. The blue box represents the exon, the white box indicates 2-kb upstream region of the start codon or 1-kb downstream region of the stop codon, and the black line represents the intron. **(B)** Comparison of GR between accessions containing Hap1 and Hap2. **(C)** Comparison of seed germination status after imbibition for five days between *LOC_Os01g11270* gene-edited mutant (KO#270) and its wild-type ZH11. **(D)** Comparison of GR between KO#270 and ZH11. Data represent mean ± SD from three replicates. The p-value is obtained from the T-test.

two main haplotypes of *OsGA2ox7* (Supplementary Figures 1A, B). Similarly, the SL and FW of the two main haplotypes of *LOC_Os07g34240* and *LOC_Os07g34370* in *qFW7-1* also differ significantly (Supplementary Figures 1C–F). These results suggest that GA is a vital hormone for regulating SV, and there is a great potential to improve SV by targeting genes related to the GA pathway.

Auxin is a universal coordinator for plant growth and development and stress responses. It is well known to promote coleoptile elongation and seedling growth (Gallei et al., 2020). Of 21 strong candidates, *LOC_Os01g54990* (*OsARF3*), *LOC_Os01g10970*, and *LOC_Os01g55940* (*OsGH3.2*) were regarded as auxin response factor (ARF) or indole-3-acetic acid (IAA)-amido synthetase (conversion of active IAA to an inactive form) (Supplementary Table 3), and thus may control SV through modulating auxin homeostasis or signaling pathway. Actually, *OsGH3.2* overexpression not only resulted in remarkable morphological changes in transgenic lines, such as dwarfism, shortened leaf length, and small panicles and internodes, but also showed increased cold stress tolerance and drought hypersensitivity through the modulation of IAA and ABA homeostasis. In contrast, no significant change in these phenotypes was detected in *OsGH3.2*-suppressed plants (Du et al., 2012). Recently, OsGH3.2 was demonstrated as a negative regulator of seed longevity by regulating the accumulation of ABA (Yuan et al., 2021). In this study, we observed significant differences in SV-related traits between the two main haplotypes of *OsGH3.2* and *OsARF3* in *qSL1-2* and *LOC_Os01g10970* in *qGR1-1* (Supplementary Figures 2A–C). These data suggest that auxin acts as an essential factor in modulating SV.

LOC_Os01g11270, *LOC_Os01g11280*, and *LOC_Os01g11340* in *qGR1-1* were predicted as CYP450s (Supplementary Table 3), which are the biggest enzymatic protein family involved in NADPH- and oxygen-dependent hydroxylation reactions in plants (Pandian et al., 2020). As versatile catalysts, CYP450s play vital roles in the biosynthesis of primary and secondary metabolites, antioxidants, and phytohormones, thereby regulating plants' growth and development, and protecting them from stresses (Xu et al., 2015; Pandian et al., 2020). *CYP714D1*, *CYP714B1*, and *CYP714B2* act as the GA deactivating enzymes in rice, and their mutants and overexpressing transgenic plants showed tall and dwarfed phenotypes, respectively (Zhu et al., 2006; Magome et al., 2013). *CYP90D2* encodes a brassinosteroid (BR) biosynthetic enzyme whose mutants were dwarfed (Hong et al., 2003; Fang et al., 2016). *CYP96B4* modulates rice growth

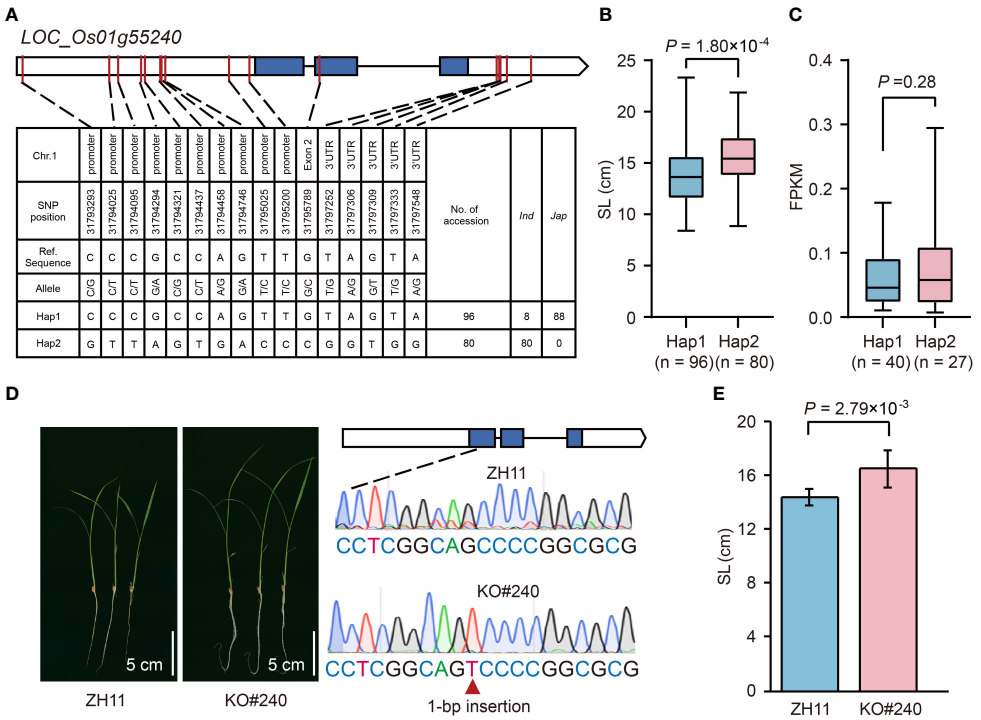


FIGURE 6
Effects of *LOC_Os01g55240* on rice seedling growth. **(A)** The top two haplotypes of *LOC_Os01g55240* in 346 rice accessions. The blue box represents the exon, the white box indicates 2-kb upstream region of the start codon or 1-kb downstream region of the stop codon, and the black line represents the intron. **(B)** Comparison of SL between accessions with Hap1 and Hap2. **(C)** Comparison of gene expression in leaves between accessions with Hap1 and Hap2. The mean FPKM value of three replicates is represented as the gene expression value for each accession. **(D)** Comparison of seedling growth status after imbibition for ten days between *LOC_Os01g55240* gene-edited mutant (KO#240) and its wild-type ZH11. **(E)** Comparison of SL between KO#240 and ZH11. Data represent mean ± SD from ten replicate. The p-value is obtained from the T-test.

and stress tolerance by finetuning the balance of GA and ABA. *cyp96b4* mutant showed pleiotropic abnormal phenotypes, including dwarfism, delayed germination and early growth, and enhanced tolerance to drought (Tamiru et al., 2015). In addition, *CYP71D8L* is a crucial regulator for GA and cytokinin (CK) homeostasis, thereby controlling multiple agronomic traits and stress responses (Zhou et al., 2020). These results imply that *LOC_Os01g11270*, *LOC_Os01g11280*, and *LOC_Os01g11340* could influence SV by mediating phytohormones' homeostasis. Here, significant variations in GR between the two main haplotypes of these three genes were detected (Figures 5A, B, Supplementary Figures 3A, B), which support the above speculation. And the effect of *LOC_Os01g11270* on GR was further confirmed using its CRISPR/Cas9 knockout mutant (Figures 5C, D). In addition, the Hap2 with a higher GR may be an elite allele available for SV improvement.

Ca^{2+} acts as an essential nutrient and a universal second messenger for plants' growth and development in normal and stressful conditions. The Ca^{2+} signals are decoded through the interactions between a unique set of Ca^{2+} sensors named calcineurin B-like proteins (CBLs) and the major downstream signaling components called CBL-interacting protein kinases (CIPKs). These CBLs-CIPKs interactions define the complexity of the Ca^{2+} signaling networks for the perception and transduction of stress signals under various environmental stresses (Tang et al., 2020). In rice, *OsCIPK3*, *OsCIPK9*, and *OsCIPK23* are involved in cold, salinity, and drought stress responses, respectively (Xiang et al., 2007; Yang et al., 2008; Shabala et al., 2021). *OsCIPK31* functions in seed germination and seedling growth under abiotic stress conditions (Piao et al., 2010). *OsCIPK15* is related to rice germination and subflooding tolerance under oxygen deprivation (Lee et al., 2009; Kudahettige et al., 2011). Of 21 strong candidates, two *CIPK* genes, *LOC_Os01g55450* (*OsCIPK12*) in *qSL1-2* and *LOC Os01g10870* (*OsCIPK13*) in *qGR1-1* were observed (Supplementary Table 3). We found apparent variations in SL and GR between the two main haplotypes of *OsCIPK12* and *OsCIPK13*, respectively (Supplementary Figures 4A, C). In addition, the longer SL in Hap2 of *OsCIPK12* corresponds to the higher gene expression level, indicating that *OsCIPK12* is a positive regulator of SL (Supplementary Figure 4B). Our results, combined with the overexpression of *OsCIPK12* resulted in

increased drought tolerance in seedlings (Xiang et al., 2007), suggesting that *OsCIPK12* is a promising SV improvement gene.

As critical regulatory elements of gene expression, we identified three transcription factors (TFs) in *qSL1-1* and *qSL1-2*, including one bZIP family TF (*LOC_Os01g55150*) and two OVATE family TFs (*LOC_Os01g53160* and *LOC_Os01g54570*) (Supplementary Table 3). The bZIP family proteins have been shown to regulate a set of plant growth and development processes, such as photomorphogenesis, floral induction and development, and seed maturation and germination. They also involve stress responses (Nijhawan et al., 2008). Here, we observed that the Hap2 of *LOC_Os01g55150* had longer SL, heavier FW, and higher gene expression than those of Hap1 (reference sequence) (Supplementary Figures 5A-C), implicating that *LOC_Os01g55150* acts as a positive regulator. Compared with the bZIP family, the OVATE family with a conserved DUF623 domain is relatively small and poorly understood. In rice, several OVATE family proteins (OsOFPs) have been cloned and shown to regulate multiple aspects of plant growth and development, including seedling growth, plant architecture, leaf morphology, grain shape, and abiotic stresses (Schmitz et al., 2015; Yang et al., 2016; Ma et al., 2017; Xiao et al., 2017; Yang et al., 2018; Xiao et al., 2020). Among them, *LOC_Os01g53160*, also named *OsOFP3*, acts as a repressor of both BR biosynthesis and signal and incorporates into a TF complex to regulate BR signaling, thereby controlling the proper development of plants. In this study, we found apparent variations in SL and FW between the top two haplotypes of *LOC_Os01g53160* and *LOC_Os01g54570* (Supplementary Figures 5D–G). Furthermore, the larger values of SL and FW in Hap2 of *LOC_Os01g54570* were associated with its higher gene expression level (Supplementary Figure 5H), implicating a positive role of *LOC_Os01g54570* on SV.

In sum, we detected six QTLs for GR, SL, and FW traits by GWAS within 346 rice accessions, and screened 21 strong candidate genes with differential SV-related features between their top two haplotypes. Two genes, *LOC_Os01g11270* and *LOC_Os01g55240*, were further validated. The identification of those candidate genes and their elite haplotypes provides a promising source for molecular breeding for high vigor seeds in rice.

Data availability statement

The original SNP data presented in the study are publicly available. This data can be found here: https://github.com/LipingDai/GWAS-for-rice-vigor/. The raw sequencing data presented in this article are not readily available for ownership reasons. Requests to access the datasets should be directed to the corresponding authors.

Author contributions

DZ, QQ, and QL designed and supervised the study. LD and XL performed the experiments. LD conducted bioinformatics analysis and interpretation of the data. LS, LG, GZ, ZG, LZ, JH, DR, QZ, DZ, QQ, and QL participated in rice material collection and provided resequencing resources. LD and QL drafted the manuscript. DZ revised the manuscript. LD and QL responded to review comments. All authors contributed to the article and approved the submitted version.

Funding

This work was financially supported by the Hainan Yazhou Bay Seed Laboratory Project (B21HJ0220), the National Natural Science Foundation of China (32101755), the China Postdoctoral Science Foundation (2020M680778), and the Key Research and Development Program of Zhejiang Province (2021C02056).

Acknowledgments

We thank Biogle Genome Editing Centre (BGEC) for providing the rice gene editing seeds.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1005203/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 (9), 1655–1664. doi: 10.1101/gr.094052.109

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w^{1118}; *iso-2*; *iso-3*. *Fly* 6 (2), 80–92. doi: 10.4161/fly.19695

Dai, L. P., Lu, X. L., Zou, W. W., Wang, C. J., Shen, L., Hu, J., et al. (2020). Mapping of QTLs for source and sink associated traits under elevated CO_2 in rice (*Oryza sativa* L.). *Plant Growth Regul.* 90 (2), 359–367. doi: 10.1007/s10725-019-00564-5

Dang, X., Thi, T. G. T., Dong, G., Wang, H., Edzesi, W. M., and Hong, D. (2014). Genetic diversity and association mapping of seed vigor in rice (*Oryza sativa* L.). *Planta* 239 (6), 1309–1319. doi: 10.1007/s00425-014-2060-z

Du, H., Wu, N., Fu, J., Wang, S., Li, X., Xiao, J., et al. (2012). A GH3 family member, OsGH3-2, modulates auxin and abscisic acid levels and differentially affects drought and cold tolerance in rice. *J. Exp. Bot.* 63 (18), 6467–6480. doi: 10.1093/jxb/ers300

Fang, N., Xu, R., Huang, L., Zhang, B., Duan, P., Li, N., et al. (2016). SMALL GRAIN 11 controls grain size, grain number and grain yield in rice. *Rice* 9 (1), 64. doi: 10.1186/s12284-016-0136-z

Feng, Z., Li, L. B., Tang, M. Q., Liu, Q. B., Ji, Z. H., Sun, D. L., et al. (2022). Detection of stable elite haplotypes and potential candidate genes of boll weight across multiple encironments via GWAS in upland cotton. *Front. Plant Sci.* 13, 929168. doi: 10.3389/fpls.2022.929168

Fujino, K., Sekiguchi, H., Matsuda, Y., Sugimoto, K., Ono, K., and Yano, M. (2008). Molecular identification of a major quantitative trait locus, *qLTG3–1*, controlling low-temperature germinability in rice. *Proc. Natl. Acad. Sci.* 105 (34), 12623–12628. doi: 10.1073/pnas.0805303105

Gallei, M., Luschnig, C., and Friml, J. (2020). Auxin signalling in growth: Schrödinger's cat out of the bag. *Curr. Opin. Plant Bio.* 53, 43–49. doi: 10.1016/j.pbi.2019.10.003

He, Y. Q., Cheng, J. P., He, Y., Yang, B., Cheng, Y. H., Yang, C., et al. (2019b). Influence of isopropylmalate synthase OsIPMS1 on seed vigour associated with amino acid and energy metabolism in rice. *Plant Biotechnol. J.* 17 (2), 322–337. doi: 10.1111/pbi.12979

He, Y., Chen, S., Liu, K., Chen, Y., Cheng, Y., Zeng, P., et al. (2022). OsHIPL1, a hedgehog-interacting protein-like 1 protein, increases seed vigour in rice. *Plant Biotechnol. J.* 20 (7), 1346–1362. doi: 10.1111/pbi.13812

He, Y., Yang, B., He, Y., Zhan, C., Cheng, Y., Zhang, J., et al. (2019a). A quantitative trait locus, *qSE 3*, promotes seed germination and seedling establishment under salinity stress in rice. *Plant J.* 97 (6), 1089–1104. doi: 10.1111/tpj.14181

Hong, Z., Ueguchi-Tanaka, M., Umemura, K., Uozu, S., Fujioka, S., Takatsuto, S., et al. (2003). A rice brassinosteroid-deficient mutant, *ebisu dwarf* (*d2*), is caused by a loss of function of a new member of cytochrome P450. *Plant Cell* 15 (12), 2900–2910. doi: 10.1105/tpc.014712

Hsieh, K. T., Chen, Y. T., Hu, T. J., Lin, S. M., Hsieh, C. H., Liu, S. H., et al. (2021). Comparisons within the rice GA 2-oxidase gene family revealed three dominant paralogs and a functional attenuated gene that led to the identification of four amino acid variants associated with GA deactivation capability. *Rice* 14 (1), 70. doi: 10.1186/s12284-021-00499-4

Islam, M. R., Naveed, S. A., Zhang, Y., Li, Z., Zhao, X., Fiaz, S., et al. (2022). Identification of candidate genes for salinity and anaerobic tolerance at the germination stage in rice by genome-wide association analyses. *Front. Genet.* 13, 822516. doi: 10.3389/fgene.2022.822516

Jahan, N., Javed, M. A., Khan, A., Manan, F. A., and Tabassum, B. (2021). Genetic architecture of Al^{3+} toxicity tolerance in rice $F_{2:3}$ populations determined through QTL mapping. *Ecotoxicology* 30 (5), 794–805. doi: 10.1007/s10646-021-02413-6

Jiang, S., Yang, C., Xu, Q., Wang, L., Yang, X., Song, X., et al. (2020). Genetic dissection of germinability under low temperature by building a resequencing linkage map in japonica rice. *Int. J. Mol. Sci.* 21 (4), 1284. doi: 10.3390/ijms21041284

Jin, J., Long, W., Wang, L., Liu, X., Pan, G., Xiang, W., et al. (2018). QTL mapping of seed vigor of backcross inbred lines derived from *Oryza longistaminata* under artificial aging. *Front. Plant Sci.* 9, 1909. doi: 10.3389/fpls.2018.01909

Kudahettige, N., Pucciariello, C., Parlanti, S., Alpi, A., and Perata, P. (2011). Regulatory interplay of the Sub1A and CIPK15 pathways in the regulation of α-amylase production in flooded rice plants. *Plant Biol.* 13 (4), 611–619. doi: 10.1111/j.1438-8677.2010.00415.x

Lee, K. W., Chen, P. W., Lu, C. A., Chen, S., Ho, T. H. D., and Yu, S. M. (2009). Coordinated responses to oxygen and sugar deficiency allow rice seedlings to tolerate flooding. *Sci. Signal.* 2 (91), ra61. doi: 10.1126/scisignal.2000333

Lee, T. H., Guo, H., Wang, X., Kim, C., and Paterson, A. H. (2014). SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genom.* 15, 162. doi: 10.1186/1471-2164-15-162

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25 (14), 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, Q., Lu, X., Wang, C., Shen, L., Dai, L., He, J., et al. (2022). Genome-wide association study and transcriptome analysis reveal new QTL and candidate genes for nitrogen-deficiency tolerance in rice. *Crop J.* 10 (4), 942–951. doi: 10.1016/j.cj.2021.12.006

Li, W., Yang, B., Xu, J., Peng, L., Sun, S. , Huang, Z., et al. (2021a). A genome-wide association study reveals that the 2-oxoglutarate/malate translocator mediates seed vigor in rice. *Plant J.* 108 (2), 478–491. doi: 10.1111/tpj.15455

Li, L., Zhang, C., Huang, J., Liu, Q., Wei, H. , Wang, H., et al. (2021b). Genomic analyses reveal the genetic basis of early maturity and identification of loci and candidate genes in upland cotton (*Gossypium hirsutum* l.). *Plant Biotechnol. J.* 19 (1), 109–123. doi: 10.1111/pbi.13446

Lo, S. F., Yang, S. Y., Chen, K. T., Hsing, Y. I., Zeevaart, J. A., Chen, L. J., et al. (2008). A novel class of gibberellin 2-oxidases control semidwarfism, tillering, and root development in rice. *Plant Cell* 20 (10), 2603–2618. doi: 10.1105/tpc.108.060913

Lu, Y., Ye, X., Guo, R., Huang, J., Wang, W., Tang, J., et al. (2017). Genome-wide targeted mutagenesis in rice using the CRISPR/Cas9 system. *Mol. Plant* 10 (9), 1242–1245. doi: 10.1016/j.molp.2017.06.007

Magome, H., Nomura, T., Hanada, A., Takeda-Kamiya, N., Ohnishi, T., Shinma, Y., et al. (2013). CYP714B1 and CYP714B2 encode gibberellin 13-oxidases that reduce gibberellin activity in rice. *Proc. Natl. Acad. Sci.* 110 (5), 1947–1952. doi: 10.1073/pnas.1215788110

Ma, Y., Yang, C., He, Y., Tian, Z., Li, J., and Sunkar, R. (2017). Rice OVATE family protein 6 regulates plant development and confers resistance to drought and cold stresses. *J. Exp. Bot.* 68 (17), 4885–4898. doi: 10.1093/jxb/erx309

Murray, M., and Thompson, W. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8 (19), 4321–4326. doi: 10.1093/nar/8.19.4321

Najeeb, S., Ali, J., Mahender, A., Pang, Y., Zilhas, J., Murugaiyan, V., et al. (2020). Identification of main-effect quantitative trait loci (QTLs) for low-temperature stress tolerance germination-and early seedling vigor-related traits in rice (*Oryza sativa* L.). *Mol. Breed.* 40 (1), 1–25. doi: 10.1007/s11032-019-1090-4

Nijhawan, A., Jain, M., Tyagi, A. K., and Khurana, J. P. (2008). Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. *Plant Physiol.* 146 (2), 333–350. doi: 10.1104/pp.107.112821

Pandian, B. A., Sathishraj, R., Djanaguiraman, M., Prasad, P. V., and Jugulam, M. (2020). Role of cytochrome P450 enzymes in plant stress response. *Antioxidants* 9 (5), 454. doi: 10.3390/antiox9050454

Peng, L., Sun, S., Yang, B., Zhao, J., Li, W., Huang, Z., et al. (2022). Genome-wide association study reveals that the cupin domain protein OsCDP3.10 regulates seed vigour in rice. *Plant Biotechnol. J.* 20 (3), 485–498. doi: 10.1111/pbi.13731

Piao, H. L., Xuan, Y. H., Park, S. H., Je, B. I., Park, S. J., Park, S. H., et al. (2010). OsCIPK31, a CBL-interacting protein kinase is involved in germination and seedling growth under abiotic stress conditions in rice plants. *Mol. Cells* 30 (1), 19–27. doi: 10.1007/s10059-010-0084-1

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81 (3), 559–575. doi: 10.1086/519795

Rajjou, L., Duval, M., Gallardo, K., Catusse, J., Bally, J., Job, C., et al. (2012). Seed germination and vigor. *Annu. Rev. Plant Biol.* 63, 507–533. doi: 10.1146/annurev-arplant-042811-105550

Schmitz, A. J., Begcy, K., Sarath, G., and Walia, H. (2015). Rice ovate family protein 2 (OFP2) alters hormonal homeostasis and vasculature development. *Plant Sci.* 241, 177–188. doi: 10.1016/j.plantsci.2015.10.011

Shabala, S., Alnayef, M., Bose, J., Chen, Z. H., Venkataraman, G., Zhou, M., et al. (2021). Revealing the role of the calcineurin b-like protein-interacting protein kinase 9 (CIPK9) in rice adaptive responses to salinity, osmotic stress, and k^+ deficiency. *Plants* 10 (8), 1513. doi: 10.3390/plants10081513

Shi, J., Shi, J., Liang, W., and Zhang, D. (2021). Integrating GWAS and transcriptomics to identify genes involved in seed dormancy in rice. *Theor. Appl. Genet.* 134 (11), 3553–3562. doi: 10.1007/s00122-021-03911-1

Takehara, S., Sakuraba, S., Mikami, B., Yoshida, H., Yoshimura, H., Itoh, A., et al. (2020). A common allosteric mechanism regulates homeostatic inactivation of auxin and gibberellin. *Nat. Commun.* 11, 2143. doi: 10.1038/s41467-020-16068-0

Tamiru, M., Undan, J. R., Takagi, H., Abe, A., Yoshida, K., Undan, J. Q., et al. (2015). A cytochrome P450, OsDSS1, is involved in growth and drought stress responses in rice (*Oryza sativa* L.). *Plant Mol. Biol.* 88 (1), 85–99. doi: 10.1007/s11103-015-0310-5

Tang, R. J., Wang, C., Li, K., and Luan, S. (2020). The CBL-CIPK calcium signaling network: unified paradigm from 20 years of discoveries. *Trends Plant Sci.* 25 (6), 604–617. doi: 10.1016/j.tplants.2020.01.009

Tong, H., Xiao, Y., Liu, D., Gao, S., Liu, L., Yin, Y., et al. (2014). Brassinosteroid regulates cell elongation by modulating gibberellin metabolism in rice. *Plant Cell* 26 (11), 4376–4393. doi: 10.1105/tpc.114.132092

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Proto. Bioinf.* 43 (1), 11.10. 1–11.10. 33. doi: 10.1002/0471250953.bi1110s43

Vogt, F., Shirsekar, G., and Weigel, D. (2022). vcf2gwas: Python API for comprehensive GWAS analysis using GEMMA. *Bioinformatics* 38 (3), 839–840. doi: 10.1093/bioinformatics/btab710

Wang, Z. F., Wang, J. F., Bao, Y. M., Wang, F. H., and Zhang, H. S. (2010). Quantitative trait loci analysis for rice seed vigor during the germination stage. *J. Zhejiang Univ. Sci. B.* 11 (12), 958–964. doi: 10.1631/jzus.B1000238

Wu, J., Zhu, C., Pang, J., Zhang, X., Yang, C., Xia, G., et al. (2014). $OsLOL_1$, a C_2C_2-type zinc finger protein, interacts with OsbZIP58 to promote seed germination through the modulation of gibberellin biosynthesis in *Oryza sativa*. *Plant J.* 80 (6), 1118–1130. doi: 10.1111/tpj.12714

Xiang, Y., Huang, Y., and Xiong, L. (2007). Characterization of stress-responsive CIPK genes in rice for stress tolerance improvement. *Plant Physiol.* 144 (3), 1416–1428. doi: 10.1104/pp.107.101295

Xiao, Y., Liu, D., Zhang, G., Tong, H., and Chu, C. (2017). Brassinosteroids regulate OFP1, a DLT interacting protein, to modulate plant architecture and grain morphology in rice. *Front. Plant Sci.* 8, 1698. doi: 10.3389/fpls.2017.01698

Xiao, Y., Zhang, G., Liu, D., Niu, M., Tong, H., and Chu, C. (2020). GSK2 stabilizes OFP3 to suppress brassinosteroid responses in rice. *Plant J.* 102 (6), 1187–1201. doi: 10.1111/tpj.14692

Xie, L., Tan, Z., Zhou, Y., Xu, R., Feng, L., Xing, Y., et al. (2014). Identification and fine mapping of quantitative trait loci for seed vigor in germination and seedling establishment in rice. *J. Integr. Plant Biol.* 56 (8), 749–759. doi: 10.1111/jipb.12190

Xu, J., Wang, X. Y., and Guo, W. Z. (2015). The cytochrome P450 superfamily: Key players in plant development and defense. *J. Integr. Agr.* 14 (9), 1673–1686. doi: 10.1016/S2095-3119(14)60980-1

Yang, B., Chen, M., Zhan, C., Liu, K., Cheng, Y., Xie, T., et al. (2022). Identification of *OsPK5* involved in rice glycolytic metabolism and GA/ABA balance for improving seed germination *via* GWAS. *J. Exp. Bot.* 11 (2), 3446–3461. doi: 10.1093/jxb/erac071

Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., et al. (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* 5, 5087. doi: 10.1038/ncomms6087

Yang, W., Kong, Z., Omo-Ikerodah, E., Xu, W., Li, Q., and Xue, Y. (2008). Calcineurin b-like interacting protein kinase OsCIPK23 functions in pollination and drought stress responses in rice (*Oryza sativa* L.). *J. Genet. Genomics* 35 (9), 531–543. doi: 10.1016/S1673-8527(08)60073-9

Yang, C., Ma, Y., He, Y., Tian, Z., and Li, J. (2018). OsOFP19 modulates plant architecture by integrating the cell division pattern and brassinosteroid signaling. *Plant J.* 93 (3), 489–501. doi: 10.1111/tpj.13793

Yang, C., Shen, W., He, Y., Tian, Z., and Li, J. (2016). OVATE family protein 8 positively mediates brassinosteroid signaling through interacting with the GSK3-like kinase in rice. *PloS Genet.* 12 (6), e1006118. doi: 10.1371/journal.pgen.1006118

Yuan, Z., Fan, K., Wang, Y., Tian, L., Zhang, C., Sun, W., et al. (2021). *OsGRETCHENHAGEN3-2* modulates rice seed storability *via* accumulation of abscisic acid and protective substances. *Plant Physiol.* 186 (1), 469–482. doi: 10.1093/plphys/kiab059

Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T. Y. (2017). Ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8 (1), 28–36. doi: 10.1111/2041-210X.12628

Zhang, A., Liu, C., Chen, G., Hong, K., Gao, Y., Tian, P., et al. (2017). Genetic analysis for rice seedling vigor and fine mapping of a major QTL *qSSL1b* for seedling shoot length. *Breed. Sci.* 67 (3), 307–315. doi: 10.1270/jsbbs.16195

Zhang, Y., Wang, X., Luo, Y., Zhang, L., Yao, Y., Han, L., et al. (2020). *OsABA8ox2*, an ABA catabolic gene, suppresses root elongation of rice seedlings and contributes to drought response. *Crop J.* 8 (3), 480–491. doi: 10.1016/j.cj.2019.08.006

Zhou, J., Li, Z., Xiao, G., Zhai, M., Pan, X., Huang, R., et al. (2020). CYP71D8L is a key regulator involved in growth and stress responses by mediating gibberellin homeostasis in rice. *J. Exp. Bot.* 71 (3), 1160–1170. doi: 10.1093/jxb/erz491

Zhu, S., Gao, F., Cao, X., Chen, M., Ye, G., Wei, C., et al. (2005). The rice dwarf virus P2 protein interacts with ent-kaurene oxidases *in vivo*, leading to reduced biosynthesis of gibberellins and rice dwarf symptoms. *Plant Physiol.* 139 (4), 1935–1945. doi: 10.1104/pp.105.072306

Zhu, Y., Nomura, T., Xu, Y., Zhang, Y., Peng, Y., Mao, B., et al. (2006). *ELONGATED UPPERMOST INTERNODE* encodes a cytochrome P450 monooxygenase that epoxidizes gibberellins in a novel deactivation reaction in rice. *Plant Cell* 18 (2), 442–456. doi: 10.1105/tpc.105.038455

Zhu, G., Ye, N., and Zhang, J. (2009). Glucose-induced delay of seed germination in rice is mediated by the suppression of ABA catabolism rather than an enhancement of ABA biosynthesis. *Plant Cell Physiol.* 50 (3), 644–651. doi: 10.1093/pcp/pcp022

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 29 November 2022
DOI 10.3389/fpls.2022.1068769

OPEN ACCESS

EDITED BY
Xiaojin Luo,
Fudan University, China

REVIEWED BY
Shirong Zhou,
Nanjing Agricultural University, China
Yulong Ren,
Institute of Crop Sciences (CAAS), China

*CORRESPONDENCE
Bing-Ran Zhao
brzhao123@163.com
Bi-Gang Mao
mbg@hhrrc.ac.cn

SPECIALTY SECTION
This article was submitted to Plant Bioinformatics, a section of the journal Frontiers in Plant Science

RECEIVED 13 October 2022
ACCEPTED 14 November 2022
PUBLISHED 29 November 2022

CITATION
Zheng W-J, Li W-Q, Peng Y, Shao Y, Tang L, Liu C-T, Zhang D, Zhang L-J, Li J-H, Luo W-Z, Yuan Z-C, Zhao B-R and Mao B-G (2022) E2Fs co-participate in cadmium stress response through activation of MSHs during the cell cycle.
Front. Plant Sci. 13:1068769.
doi: 10.3389/fpls.2022.1068769

E2Fs co-participate in cadmium stress response through activation of MSHs during the cell cycle

Wen-Jie Zheng[1], Wang-Qing Li[1], Yan Peng[2], Ye Shao[2], Li Tang[2], Ci-Tao Liu[3], Dan Zhang[2,3], Lan-Jing Zhang[3], Ji-Huan Li[3], Wu-Zhong Luo[2], Zhi-Cheng Yuan[2], Bing-Ran Zhao[1,2,3]* and Bi-Gang Mao[1,2]*

[1]Longping Branch, College of Biology, Hunan University, Changsha, China, [2]State Key Laboratory of Hybrid Rice, Hunan Hybrid Rice Research Center, Changsha, China, [3]College of Agricultural, Hunan Agricultural University, Changsha, China

Cadmium is one of the most common heavy metal contaminants found in agricultural fields. MutSα, MutSβ, and MutSγ are three different MutS-associated protein heterodimer complexes consisting of MSH2/MSH6, MSH2/MSH3, and MSH2/MSH7, respectively. These complexes have different mismatch recognition properties and abilities to support MMR. However, changes in mismatch repair genes (*OsMSH2*, *OsMSH3*, *OsMSH6*, and *OsMSH7*) of the MutS system in rice, one of the most important food crops, under cadmium stress and their association with E2Fs, the key transcription factors affecting cell cycles, are poorly evaluated. In this study, we systematically categorized six rice E2Fs and confirmed that *OsMSHs* were the downstream target genes of E2F using dual-luciferase reporter assays. In addition, we constructed four *msh* mutant rice varieties (*msh2*, *msh3*, *msh6*, and *msh7*) using the CRISPR-Cas9 technology, exposed these mutant rice seedlings to different concentrations of cadmium (0, 2, and 4 mg/L) and observed changes in their phenotype and transcriptomic profiles using RNA-Seq and qRT-PCR. We found that the difference in plant height before and after cadmium stress was more significant in mutant rice seedlings than in wild-type rice seedlings. Transcriptomic profiling and qRT-PCR quantification showed that cadmium stress specifically mobilized cell cycle-related genes *ATR*, *CDKB2;1*, *MAD2*, *CycD5;2*, *CDKA;1*, and *OsRBR1*. Furthermore, we expressed OsE2Fs in yeasts and found that heterologous E2F expression in yeast strains regulated cadmium tolerance by regulating *MSHs* expression. Further exploration of the underlying mechanisms revealed that cadmium stress may activate the CDKA/CYCD complex, which phosphorylates RBR proteins to release E2F, to regulate downstream *MSHs* expression and subsequent DNA damage repairment, thereby enhancing the response to cadmium stress.

KEYWORDS

cadmium, cell cycle, E2Fs, MSHs, DNA damage, RNA-seq

Introduction

Cadmium (Cd) is toxic to animals and plants. Extensive Cd contamination in soils can lead to cadmium accumulation in edible parts of crops, especially rice, threatening the food safety of rice consumption (Uraguchi and Fujiwara, 2012). To minimize the risk of soil Cd entering the food chain, understanding plants' molecular response network to Cd and in-depth exploring the molecular mechanism of Cd stress have become an important direction in agriculture and environment studies (Pena et al., 2012; Wang et al., 2016; Chang et al., 2019).

The development of multicellular organisms relies on the orchestrated spatiotemporal regulation of cell division, including mitotic cell cycle and cell expansion. Therefore, cell cycles must be integrated into a complex histogenesis and organogenesis system (Mironov et al., 1997; Meijer and Murray, 2001; Stals and Inzé, 2001; Veylder et al., 2003). The plant cell cycle consists of four distinct phases: postmitotic interphase G1, DNA synthesis phase S, pre-mitotic interphase G2, and mitosis/cytokinesis phase M (Qi and Zhang, 2019). During the G1/S transition, the A-type cyclin-dependent kinase/Cyclin D (CDKA/CYCD) complex, *via* phosphorylating retinoblastoma-related (RBR) proteins, activates the S-phase E2 promoter binding factor (E2F). Activated E2F then regulates the expression of genes involved in DNA replication, cell cycle progression, and chromatin dynamics transition, thereby promoting the G1/S transition (del Pozo et al., 2006).

Both typical and atypical E2F transcription factors are key components of the cyclin D/RB/E2F pathway. Typical E2F proteins possess a homologous DNA-binding domain and form heterodimers with Dimerization part (DP) proteins to bind DNA *via* the leucine zipper in their dimerization region. In contrast, atypical E2F/DEL factors do not interact with DPs due to lacking a dimerization region. However, they have a homologous DNA-binding domain for replication, enabling them to bind DNA autonomously (Perrotta et al., 2021). In addition, typical E2F proteins also possess a conserved C-terminal transactivation domain that is absent in DPs and in atypical E2F/DEL proteins, allowing transcriptional activation of their target genes (Dimova and Dyson, 2005). *Arabidopsis* E2F/DP family includes three typical E2Fs (*AtE2Fa*, *AtE2Fb*, and *AtE2Fc*), three atypical E2Fs (*AtE2Fd/DEL2*, *AtE2Fe/DEL1*, and *AtE2Ff/DEL3*) and two DPs (Mariconti et al., 2002). Carrot E2F/DP family consists of four typical E2Fs, three atypical E2Fs, and three DPs (Perrotta et al., 2021). Yeast, which is widely used as an *in vivo* model to explore cell cycle response mechanisms, also utilizes E2F homologs MCB-binding factor (MBF) and SCB-binding factor (SBF) proteins to regulate G1/S transition and cell proliferation (Morshed et al., 2020). Previous studies have identified four E2Fs, three DPs, and two DP-E2F-like (DELs) in the rice genome, but little is known about their taxonomy and functions (Guo et al., 2007).

Proper cell division and cycling requires tight regulation of key cell cycle-related genes. RBR proteins act as negative regulators of E2F transcription factors and are crucial for E2F to function properly (Weintraub et al., 1992). The CDKA/CYCD complex regulates E2F/DP family activity *via* RBR proteins, which mediate G1/S transition. CDKA forms complexes with A, B, or D-type cyclins to drive G2/M transition (Gutierrez, 2009). The dramatic repression of Mitotic arrest deficient protein 2 (*MAD2*) expression may mediate G2/M arrest through a dual mechanism to regulate chromosome segregation (Cao et al., 2018). ATM and Rad3-related (*ATR*)is a key component of the G2/M checkpoint in plant cells and is activated by the mitogen-activated protein kinase (MAPK) signaling pathway *via* phosphorylation (Yoshioka et al., 2006).

Many replication- and mismatch repair-related genes in *Arabidopsis* contain conserved E2F binding sequences in their predicted promoter regions (Kosugi and Ohashi (2002); Vandepoele et al., 2005). Previous work on the *Arabidopsis* genome at E2F binding sites has identified over 180 potential E2F target genes associated with transcription, stress, DNA damage, plant defense, and signaling transduction, in addition to the cell cycle (Ramirez-Parra, 2003). Among them are genes related to mismatch repair (MMR), such as Mutated S homologue 3 (*MSH3*), Mutated S homologue 6 (*MSH6*), and Postmeiotic segregation increased 2 (*PMS2*). In addition, the putative rice homologs are listed.

The most important role of the MMR system is to identify and correct mispaired or unpaired bases (Schofield and Hsieh, 2003; Iyer et al., 2006). Eukaryotic cells rely on high-fidelity DNA replication to maintain genomic integrity and depend on DNA MMR to ensure proofreading of incorrect pairings. Thus, any deletion in MMR genes can lead to spontaneous mutations in organisms (Zhang et al., 2005). In addition to correcting base-base mismatches, MMR genes are involved in suppressing mutations and inducing protective responses to various types of DNA damage. MMR plays multiple roles in response to various DNA damage inducers, such as nucleotide methylation, oxidative DNA damage, and UV-induced DNA damage, among other degenerative damages (Ijsselsteijn et al., 2020). Interestingly, cadmium interacts with DNA repair systems, cell cycle checkpoints, apoptosis-related epigenetic factors, and factors controlling gene expression. Cadmium can bind directly to DNA at very low concentrations, inducing various DNA damages such as base-base mismatches, insertion/deletion loops, strand cross-links, and breaks (Filipic, 2012). DNAs under stress can induce both complex and precise repair mechanisms and signal transduction pathways in eukaryotic cells and a stage-specific arrest of the cell cycle process (Wang et al., 2013; Hu et al., 2016; Xiang et al., 2017). Mutated S α (MutSα), Mutated S β (MutSβ), and Mutated S γ (MutSγ) are three different MutS-associated protein heterodimer complexes consisting of MSH2/MSH6, MSH2/MSH3, and MSH2/MSH7, respectively. These complexes have different

mismatch recognition properties and abilities to support MMR (Wang et al., 2020). In plants, Knockdown of OsMSH2 does not cause gene variants at other locations in the genome (Karthika et al., 2021). Mutated S homologue 2 (MSH2) preferentially activates ATR to trigger DNA damage responses (DDR), including regulation of cell cycle, endoreplication, cell death, and recruitment of other DNA repair, thereby enhancing plant tolerance to cadmium (Wang et al., 2020). OsMSH6 affects rice microsatellite stability and homologous recombination and plays an important role in ensuring genome stability and genetic (Jiang et al., 2020). MSH7 is a plant-specific protein similar to MSH6, and in rice OsMSH7 is able to interact with Meiotic chromosome association 1 (MEICA1) and play a role in meiotic recombination (Hu et al., 2017). *AtMSH2* and *AtMSH6* are involved in G2/M or G1/S transitions in *Arabidopsis* and soybeans. *MSH2* and *MSH6* may be the direct sensors of cadmium-mediated DNA damage. Expression of DNA mismatch repair-related genes AtMSH2, AtMSH3, AtMSH6, and AtMSH7 can be used as potential biomarkers for evaluating cadmium exposure in Arabidopsis seedlings (Liu et al., 2009).

In this study, we systematically classified E2Fs in rice, determined its relationship with downstream MSH target genes, and then explored the phenotype of rice E2F and MSH under cadmium stress. These conclusions combined with transcriptome and quantitative results, established a model hypothesis of E2Fs-MSHs.The study revealed that E2Fs co-participate in responses to cadmium stress by binding to MSHs during the cell cycle.

Materials and methods

Plant materials and growth conditions

Indica rice variety Huazhan (HZ) grains were obtained from the State Key Laboratory of Hybrid Rice. *Osmsh2*, *Osmsh3*, *Osmsh6*, and *Osmsh7* knockout rice were established using the CRISPR-Cas9 technology *via* transformation, screened by hygromycin, and sequencing verified in each generation. The pure T3 generation seeds were obtained and used for subsequent experiments. All plant materials were grown in the transgenic test base, and seedlings were cultured in an artificial climate chamber (Hunan Changsha).

The *Osmsh2* exon 4, *Osmsh3* exon 2, *Osmsh6* exon 1, and *Osmsh7* exon 2 were selected as the target sites to construct the gene knockout vectors pYLCRISPR/Cas9-MT(I)-*OsMSH2*, *OsMSH3*, *OsMSH6*, and *OsMSH7*, respectively, using target connector primers Cas9-*OsMSH2*-F and Cas9-*OsMSH2*-R and the target-linked primers Cas9-F and Cas9-R. All primer pairs are listed in Supplemental Table 1. Accession numbers for all genes are listed in Supplementary Table 2.

Rice seeds were germinated at 37°C and 60% relative humidity after sterilization first in H_2O_2 (10% v/v) for 0.5 h and then in NaClO (0.1% v/v) for 1 day and transferred to 96-well nursery plates. Seedlings were cultivated in 1/4 nutrient solution as recommended by the International Rice Research Institute (Ponnamperuma, 1977). The 10-day-old seedlings were subjected to $CdCl_2$ treatment at 0, 2, and 4 mg/L concentrations for 10 days. The seedling leaves were collected at 6 hours of treatment, snap-frozen in liquid nitrogen, and stored at -80°C for RNA extraction. Total RNAs were extracted from seedling leaves using an RNAprep pure Plant Kit (Magen, China) and reverse transcribed using a SuperScript II kit (TransGen, China). RNA concentration and quality were determined with a NanoDrop 2000 (Thermo Scientific, Waltham, MA, USA). Each RNA sample was divided into two aliquots for RNA-seq and qRT-PCR, respectively.

Phylogenetic tree construction, gene structure, protein structure and expression pattern analysis

Clustal_W was used to compare the E2F sequences of Arabidopsis Thaliana and rice, and GeneDoc was used to output the amino acid alignment map. The phylogenetic tree was constructed using the adjacency method (NJ) of MEGA6.0 software, and the Bootstrap value was set to 1000. From the Ensembl the Plants get the length of the candidate genes or cDNA sequence, and CDS use GSDS 2.0 (http://gsds.cbi.pku.edu.cn/) to analyze E2F gene exon/embedded substructure, The conserved domain was analyzed using MEME tool (http://meme-suite.org/index.html). From the Rice Expression database IC4R (Information Commons for Rice; The expression levels of E2F genes in different tissues were obtained in http://ic4r.org, and then plotted with TBtools (Chen et al., 2020).

Dual-luciferase reporter assays

The full-length coding regions of *OsE2Fa-1*, *OsE2Fa-2*, *OsE2Fa-3*, *OsE2Fc*, *OsE2Fe-1*, and *OsE2Fe-2* were amplified using specific primers listed in Supplemental Table 1 and cloned into a pGreen II 62-SK vector. After that, pGreenII 0800-LUC double-reporter vector was fused to the promoter fragments of *OsMSHs*. The above constructs were transferred into *A. tumefaciens* strain GV3101 (pSoup-p19) to generate pro35S:*OsE2Fa-1*, pro35S:*OsE2Fa-2*, pro35S:*OsE2Fa-3*, pro35S:*OsE2Fc*, pro35S:*OsE2Fe-1*, pro35S:*OsE2Fe-2*, proOsMSH2:LUC, proOsMSH3:LUC, proOsMSH6:LUC, and proOsMSH7:LUC recombinant strains. *N. benthamiana* leaves were co-infiltrated with pro35S:E2F and proOsMSH2/3/6/7:LUC and cultivated for 3 days in a growth chamber. *N. benthamiana* leaves infiltrated

with proOsMSH2/3/6/7:LUC were used as the internal control. After inoculation and a transient incubation of 72 h, the relative LUC activity was measured using a dual-LUC reporter assay system (Promega), which included firefly LUC and Renilla (REN) LUC. Leaf discs with an area of 2 cm^2 leaf were sampled and finely ground in 500 mL of Passive Lysis Buffer. Crude extracts (8 µL) were mixed with 40 µL of Luciferase Assay Buffer (Hellens et al., 2005), and the promoter activity was determined as the LUC/REN value using a luminometer (Modulus™, Promega). Each measurement included three independent biological replicates.

Heterologous OsE2Fs expression in yeasts

The full-length ORFs of OsE2Fs were amplified from *Oryza sativa* Nipponbare cDNA using primer pairs listed in Supplemental Table 1, digested with *Kpn*I and *Bam*HI and cloned into the corresponding sites of yeast expression vector pYES2 (Invitrogen). The resulting pYES2-OsMSHs constructs and the empty vector were transformed into yeast strain Δycf1 (BY4741; *MATa*; *his3Δ1*; *leu2Δ0*; *met15Δ0*; *ura3Δ0*; *YDR135c::kanMX4*) (Tan et al., 2019). Positive OsE2F transformants were selected on synthetic media lacking uracil. To analyze cadmium tolerance, individual transformant cultures were diluted, spotted on solid media containing 2% galactose and 0, 5, 10, or 20 µM $CdCl_2$, incubated at 30°C for 4 days, and photographed.

Subcellular localization

Subcellular localization of OsMSHs was investigated by transiently overexpressing 35S:OsMSH2-GFP, 35S:OsMSH3-GFP, 35S:OsMSH6-GFP, and 35S:OsMSH7-GFP in tobacco (Nicotiana tabacum) leaves *via* Agrobacterium-mediated transformation (Sparkes et al., 2006). Nuclear marker (m-Cherry) and GFP signals were observed under a confocal scanning microscope (Model LSM 880, Zeiss, Jena, Germany).

Transcriptome sequencing analysis

RNA quality control, illumina library construction, and sequencing

mRNA samples were isolated using VAHTSTM mRNA Capture Beads (Vazyme Biotech) and analyzed using Qubit (Invitrogen) and bioanalyzer (Shanghai Furi Science and Technology) to determine their concentration and contamination. After that, RNA-sequencing libraries were established. Briefly, the 1st strand cDNAs were synthesized using the 1st strand buffer and enzyme mix, and the 2nd strand DNAs were subsequently obtained by adding the 2nd strand buffer and enzyme mix. Double-stranded DNAs were purified using 1.8×VAHTS™ DNA Clean Beads (Vazyme Biotech) and subjected to terminal repair and elongation with dA-tailing, as well as ligation with adaptors. The target fragments were size-selected by 0.7× and 0.1×VAHTS™ DNA Clean Beads (Vazyme Biotech), amplified as sequencing templates and sequenced on an Illumina HiSeq™ system (LC Sciences) following the manufacturer's protocol. Clean reads were mapped to a reference genome (*Oryza sativa* Group 4.0, https://www.ncbi.nlm.nih.gov/nuccore/255672756?report=fasta), and gene expression levels were calculated by quantifying the cDNA fragments per kilobase of transcript per million fragments mapped (FPKM).

Quantitative real-time reverse transcription-PCR

Quantitative real-time reverse transcription-PCR (qRT-PCR) was performed in a 384-well plate using the SYBR premix Ex TaqTM kit (Vazyme, China) on a Roche LightCycler 480 II instrument. Relative gene expression levels were analyzed using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

Differentially expressed genes and heatmap expression of FPKM of related genes

Differentially expressed genes (DEGs) were identified using the DESeq method based on negative binomial distribution with an absolute log2 (fold change) value ≥ 1 and the corresponding Q value ≤ 0.05 as the selecting thresholds. The Q value was a corrected p-value calculated using KOBAS 2.0 and the BH method (Mao et al., 2005). The FPKM values of the screened related genes were summarized as heatmap expression using TBtools (Chen et al., 2020).

Data analysis

Statistical analysis was performed using SPSS (version 20.0). Differential expression analysis was performed for each group of samples using the DESeq2 R package (1.20.0). One-way analysis of variance and the least significant difference test were used to detect significant differences between mutant and wild-type rice.

Results

Evolutionary analysis and tissue-specific expression profiles of the E2F family in *Arabidopsis* and rice

Six rice E2F genes were identified by screening and E2F candidate genes in the *Arabidopsis* and rice E2F transcription factor database using *Arabidopsis* and rice Pfam (PF02319) (Mariconti et al., 2002; Guo et al., 2007; Perrotta et al., 2021) and renamed according to the *Arabidopsis* nomenclature (Figure 1A). Their exon structure is similarly complex to that

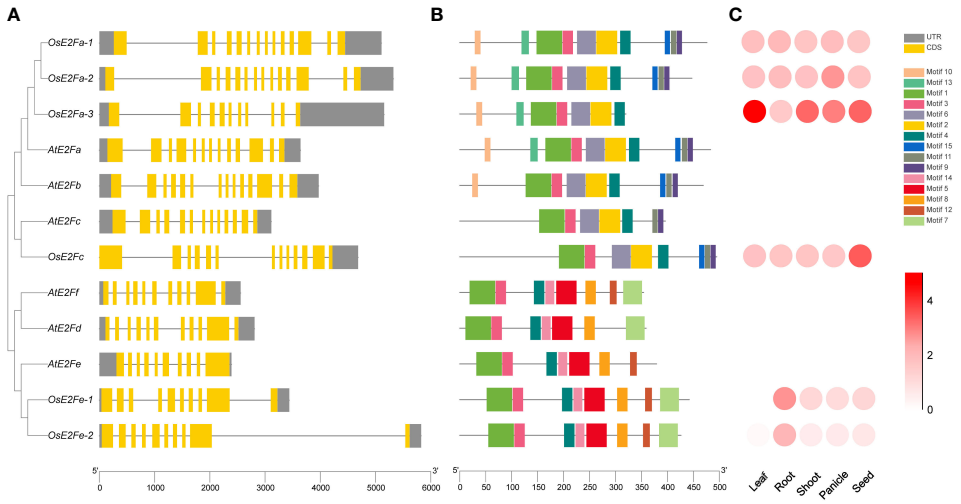


FIGURE 1
Phylogenetic tree, gene structure, conserved structural domains, and expression pattern analysis of rice and *Arabidopsis* E2Fs **(A)** On the left is the phylogenetic tree of the rice and *Arabidopsis* gene families. The gene structures are shown on the right. Exons and UTRs are indicated by boxes, while gray lines indicate introns of genes, respectively. **(B)** The distribution of motifs in the amino acid sequence. Patterns are indicated by different colored boxes. The scale at the bottom indicates the gene length. **(C)** Expression patterns of rice E2Fs in different tissues, with white indicating the lowest expression and red the highest.

of *Arabidopsis*, with the least number of *OsE2Fe-1* exons, 10. MEME(Motif-based sequence analysis tools) analysis of E2F proteins identified 15 conserved motifs, and their distributions in rice and *Arabidopsis* are shown in Figure 1B. Conservative motif analysis showed that most motifs clustered in the same phylogenetic taxon share a common motif composition, suggesting that most evolutionarily-related motifs may have the same physiological functions. The similarity of gene structures and motif compositions verifies that OsE2F branching and classification in this study are reliable. The expression patterns of each rice E2F gene in leaves, roots, stems, panicles and seeds were downloaded from the Rice Expression Database (http://expression.ic4r.org/) (Project ID: DRP000391) and presented as a heatmap (Figure 1C). It is obviously that the expression of *OsE2Fe-1* and *OsE2Fe-2* was extremely lower in rice leaves but very similar among other tissues. The amino acid sequence comparison revealed no significant difference in DNA binding-1, DNA binding-2, pRBR binding, leucine zipper, and marked box between *Arabidopsis* and rice E2F genes. Our findings (Figure S1) complement and refine the previous work, which only compared individual rice E2F genes Kosugi and Ohashi (2002).

E2F transcription factor can directly or indirectly activate MSH genes

To verify the activation status of MSHs gene by E2F transcription factor, possible transcription factor binding sites in the promoter regions of *OsMSH2*, *OsMSH3*, *OsMSH6*, and *OsMSH7* were analyzed and predicted using Plant Care. All binding sites were formed into a tandem sequence as the target promoter and constructed into LUC vectors to examine their binding with E2Fs in 62SK (Hellens et al., 2005). The results showed that all E2Fs, except OsE2Fe-2, could directly or indirectly activate the target genes (Figure 2A). The tandem element was then split into *OsMSH2* (TTTGCCGCTTTCCCGC), *OsMSH3* (TTTCCCGC), *OsMSH6* (TTTCCCGC), and *OsMSH7* (GCGGGAAATTTCCCGC), and their binding with each E2F was determined separately. The results showed that OsE2Fa-1 and OsE2Fa-3 could directly or indirectly activate *OsMSH2* (Figure 2B), OsE2Fa-2 and OsE2Fe-1 can activate *OsMSH3* directly or indirectly (Figure 2C), OsE2Fa-1 activates *OsMSH6* (Figure 2D), and OsE2Fa-1, OsE2Fc and OsE2Fe-1 activates *OsMSH7* (Figure 2E). Taken together, most E2F transcription factors activate MSH7, while OsE2Fa-1 activates most MSHs. These results are consistent with the previous prediction for the E2F binding sites (Vandepoele et al., 2005).

MSH genes are involved in cadmium stress response

Rice mutant materials *Osmsh2*, *Osmsh3*, *Osmsh6*, and *Osmsh7* were created using the CRISPR-Cas9 system. Figure 3A shows the gene structures and editing types. To ensure the accuracy of the experiment and the consistency of variables, the height of wild-type and mutant rice seedlings

without cadmium treatment were measured at the cycle point of rice growth till day 12 and showed no significant difference (Figure 3B). We also calculated the germination percentage of wild-type and mutant seeds before treatment. Treatment with 2 mg/L or 4 mg/L $CdCl_2$ significantly reduced the height, root length, fresh weight, and dry weight of wild-type and mutant rice seedlings (Supplemental Table 3). After 10 days of 2 mg/L $CdCl_2$ treatment, but not after 10 days of 4 mg/L $CdCl_2$ treatment, the height of rice seedlings was significantly different between wild-type seedlings and *Osmsh2*, *Osmsh3*, *Osmsh6*, and *Osmsh7* mutants. Moreover, the height of *Osmsh2* mutant seedlings, but not that of wild-type and other mutant seedlings, was significantly different after 2 mg/L and 4 mg/L $CdCl_2$ treatment (Figure 3C). Furthermore, the height of all mutant seedlings except *Osmsh3* without $CdCl_2$ treatment was significantly different from that after 4 mg/L $CdCl_2$ treatment (Figure 3D). There were no statistically significant changes in root length, fresh weight, and dry weight between wild-type and mutant seedlings after 2 mg/L and 4 mg/L $CdCl_2$ treatment (Figure S2D). These results indicate that *OsMSH2*, *OsMSH3*, *OsMSH6*, and *OsMSH7* are involved in cadmium stress response in rice.

MSHs affect cell cycle progression under cadmium stress

The effects of MSH genes on cell cycle progression under cadmium stress in rice leaves were determined by RNA-seq analysis to identify cell cycle genes with differential changes and further verified by qRT-PCR. The results showed that the expression of all MSH genes was upregulated in wild-type rice leaves after cadmium treatment. The expression of cell cycle-related genes (*CDKB2;1*, *ATR*, *MAD2*, *OsE2Fa-1*, *OsE2Fa-2*, *CycD5;2*, and *CDKA;1*) in rice leaves under increasing cadmium stress showed an inverted "U" pattern, with a 1.0-2.5fold upregulation at 2 mg/L $CdCl_2$ and a 0.8-1.9fold upregulation at 4 mg/L $CdCl_2$. The expression of *OsRBR1* was upregulated by 1.5-2.5 and 0.8-1.9 folds at 2 mg/L and 4 mg/L $CdCl_2$, respectively. Unlike other genes, *OsRBR1* showed a significant downregulation at 4 mg/L $CdCl_2$ and no significant difference at 2 mg/L $CdCl_2$. It is worth noting that *OsMSH2*, *OsMSH3*, *OsMSH6*, and *OsMSH7* showed a stepwise increase in expression at 2 mg/L and 4 mg/L $CdCl_2$. The expression of these genes was upregulated by 1.9-2.6 folds under 2 mg/L $CdCl_2$ and 2.2-3.3 folds under 4 mg/L $CdCl_2$, indicating that plants would activate the MMR system to avoid more serious damages when exposed to toxic cadmium (Figures S2A-C). Cadmium stress strongly activated the expression of *ATR*, *CDKB2;1*, *MAD2*, *OsE2Fa-1*, and *OsE2Fa-2* in rice leaves deficient in *OsMSH2*, *OsMSH3*, *OsMSH6*, or *OsMSH7* compared with controls, whereas the expression of *CycD5;2* and *OsRBR1* was dramatically reduced in leaves of cadmium-stressed rice (Figure 4). Expression of *CDKA;1* was activated in leaves of *Osmsh2* and *Osmsh6* mutant rice.

The expression of *OsMSH3* and *OsMSH6* was significantly downregulated while *OsMSH7* was upregulated in *Osmsh2* mutant leaves under cadmium stress; In *Osmsh3* mutant leaves, the expression of *OsMSH2* and *OsMSH6* was significantly downregulated and *OsMSH7* was similarly activated; in *Osmsh6* mutant, *OsMSH2* and *OsMSH7* expression was upregulated to some extent; in *Osmsh7* mutant, *OsMSH2* and *OsMSH6* expression was slightly suppressed. The results of qRT-PCR were consistent with the trend of transcriptome sequencing, where most genes related to the cell cycle in each mutant were repressed under cadmium stress.

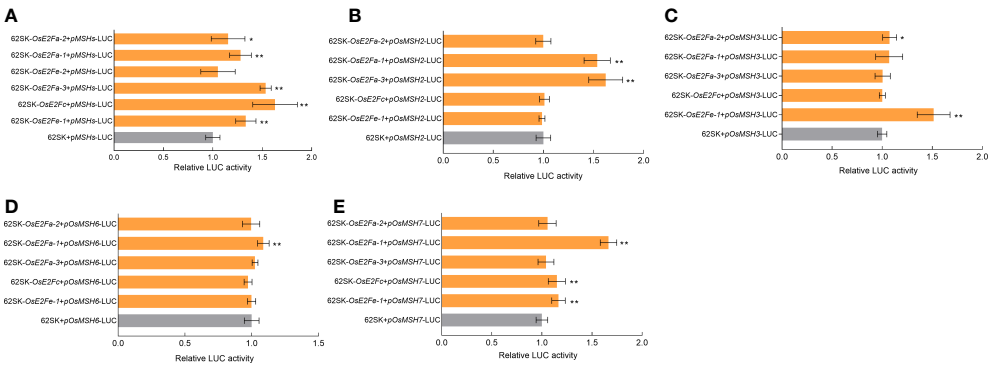


FIGURE 2
(A) LUC activation of OsE2Fs and MSHs. **(B-E)** LUC activation of OsE2Fs with OsMSH2,OsMSH3,OsMSH6 and OsMSH7, respectively. Transient dual luciferase reporter analysis indicates that E2Fs can activate MSHs. 62SK represents empty pGreenII 62-SK vector. 62SK-OsE2F represents the pGreenII 62-SK-OsE2F vector. pMSH-LUC represents pGreenII 0800-pMSH-LUC vector. Renilla luciferase (REN) was used for normalization. Results shown are means G SD (n = 9). Asterisks show significant differences from the control (Student's t-test, $p < 0.05$). * indicates $p<0.05$, ** indicates $p<0.01$.

10.3389/fpls.2022.1068769

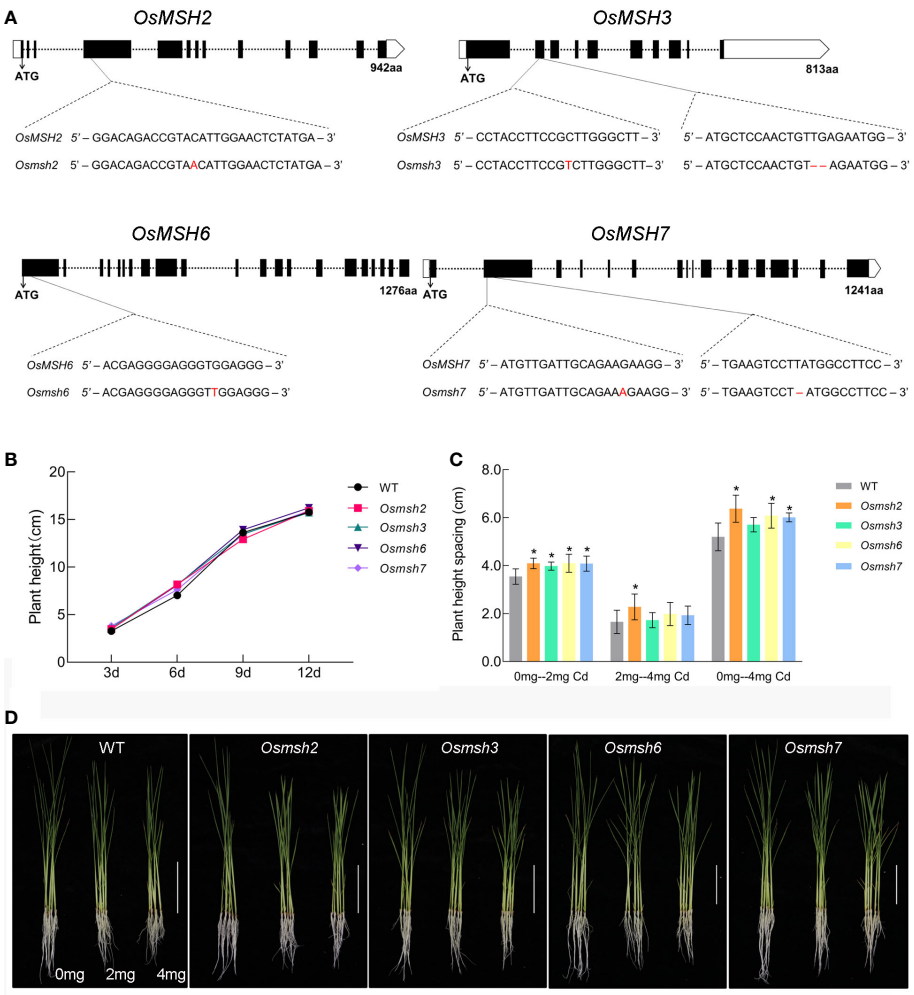


FIGURE 3
Phenotype of rice mutant seedlings under cadmium stress. **(A)** Gene structure of the rice MSH gene and editing loci for mutants. **(B)** Growth curves of rice msh mutant and wild-type seedlings from the start of germination to the 12-day stage. **(C)** Statistics of plant height differences between rice msh mutant and wild-type seedlings under 2 mg and 4 mg stress. (Student's t-test, $p < 0.05$). **(D)** Phenotypic photographs of rice msh mutant and wild-type seedlings under 2 mg and 4 mg stress(n = 5 Scale bars: 10 cm). * indicates $p<0.05$.

Interestingly, the genes mobilized by cadmium stress included the key component genes of G1/S and G2/M, suggesting that cadmium stress affects cell cycle in rice. Venn diagram plots of DEGs showed that *Osmsh2* mutant has the most downregulated genes under 4 mg/L $CdCl_2$ treatment, 1.3 times more than the wild-type (Figure S3A), of which cell cycle-specific genes account for 31% of the total downregulated genes. *Osmsh7* mutant has the most upregulated genes under 2 mg/L $CdCl_2$, 1.7 times more than the wild-type, of which cell cycle-specific genes account for 57% of the total upregulated genes. *Osmsh6* mutant has 1.3 times more cell cycle-specific upregulated genes under 4 mg/L $CdCl_2$ than the wild-type and 1.3 times more cell cycle-specific upregulated genes than the wild-type under 4 mg/L $CdCl_2$. The difference in number of DEGs between *Osmsh3* mutant and wild-type was insignificant (Figure S3B).

E2F transcription factors can respond to cadmium stress

Experiments with cadmium-sensitive yeast mutant strains showed that without cadmium, E2F expression did not affect the growth of yeast strains. However, under 5 μM cadmium stress, yeast strains expressing *OsE2Fa-1*, *OsE2Fa-2*, and *OsE2Fc* grew better than those expressing no or other E2F transcription factors. Similar results were observed under 10 μM cadmium stress, indicating that the yeast expression system is stable and the response trend of each E2F gene of Cd-sensitive strains to different cadmium stress is consistent. Understandably, this difference caused by gene expression was no longer significant at 20 μM cadmium stress (Figure 5).

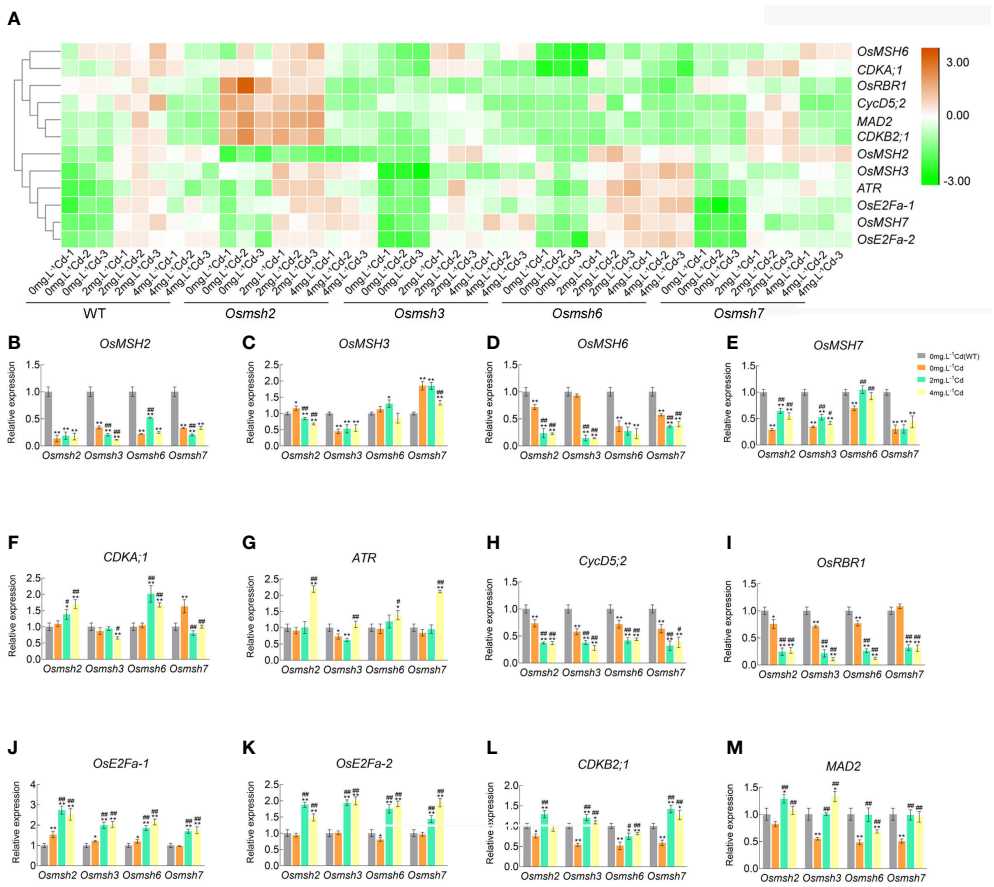


FIGURE 4
Differential expression profiles and quantitative validation of cell cycle-related genes in response to Cd stress in rice *Osmsh2*, *Osmsh3*, *Osmsh6* and *Osmsh7* mutant leaves. **(A)** Heat map representing some of the 12 cell cycle-related DEGs (P-adj≤ 0.05 and log2-fold change ≥ 1.5). Each row shows the relative expression levels of individual genes and each column shows the expression levels of individual samples. **(B–M)** The expression levels of the WT were set to 100% in the control by qRT-PCR analysis. Data were shown mean ± SD at least three independent experiments, and house-keeping gene Osactin was used as an internal control. * and # significantly statistical difference from the WT control and the corresponding mutant control, respectively (P < 0.05). ** and ## were significantly different from WT control and WT mutant control (P < 0.01).

MSHs are localized in the nucleus

To determine the subcellular localization of *OsMSH2*, *OsMSH3*, *OsMSH6*, and *OsMSH7*, *OsMSH2/3/6/7*-GFP fusion proteins driven by the CaMV 35S promoter were transiently co-expressed with the nuclear marker *AtWRKY25*-mCherry in tobacco. Confocal microscopic observation indicated that GFP signal co-existed with *AtWRKY25*-mCherry signals, indicated that *OsMSH2/3/6/7* proteins are localized in the nucleus (Figure 6).

Discussion

The E2F transcription factor is a key component of the RB/E2F pathway, which is controlled by cyclin-dependent kinases and regulates cell-cycle progression in plants and animals (Perrotta et al., 2021). Plant homologues of E2F have been tentatively identified in rice and *Arabidopsis* (Kosugi and Ohashi (2002); Mariconti et al., 2002), In this report, we extend the information on E2F factor in rice and conclusively confirmed the previous prediction (Ramirez-Parra, 2003; Vandepoele et al., 2005) that E2Fs can target and bind to MSHs, MutS is a key link in the MMR system. Plants lacking MutS will bypass MMR-mediated DDR, thereby reducing the tolerance of plants to cadmium. This has been shown in *soybeans* and *Arabidopsis* (Cui et al., 2017; Cao et al., 2018; Zhao et al., 2020), our study further revealed that E2F improves rice's tolerance to cadmium stress by binding to MSH, indicating that E2F together with MSHs responds to cadmium stress in the cell cycle. To the best of our knowledge, this is the first study on the relationship between E2F/MSH and cadmium stress in rice.

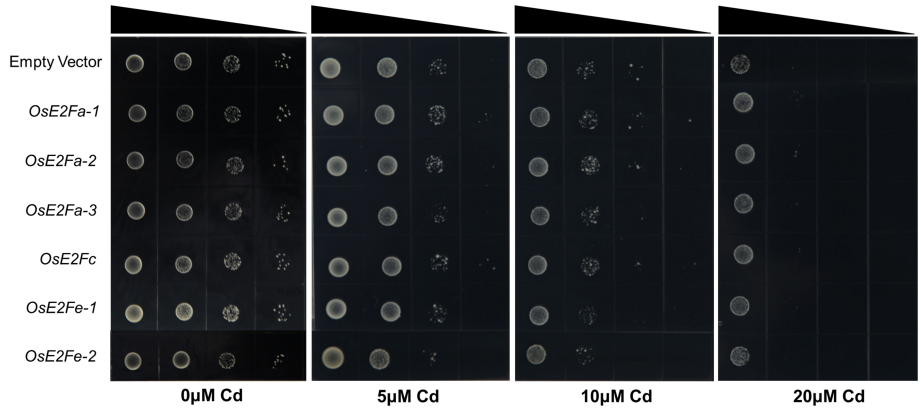


FIGURE 5
Expression of rice E2F members in Cd ion-sensitive mutant strain Δycf1. The empty vector, the control was pYES2, was grown in SD-Ura medium with the PYES2 vector linked with different E2F genes of rice. Induction of different levels of Cd stress was achieved using 0, 5, 10 and 20 μM of Cd.

The model developed here provides a preliminary explanation of the response mechanism of the E2F/MSHs pathway to cadmium stress in rice and a theoretical basis for creating cadmium-tolerant rice varieties.

Through the evolutionary tree analysis and amino acid sequence comparison of rice and *Arabidopsis* E2F transcription factors, we found that rice *OsE2Fa-1*, *OsE2Fa-2*, *OsE2Fa-3*, and *OsE2Fc* are functional transcription factors belonging to typical E2Fs (Figure 1), which, like *AtE2Fs*, have conserved C-terminal trans-activation domains and can specifically recognize E2F cis-elements and activate downstream target genes. By contrast, rice *OsE2Fe-1* and *OsE2Fe-2* are atypical E2Fs that possess DNA binding sites but lack transcriptional activation domains and cannot activate downstream gene expression. This further illustrates the complex regulatory mechanism of E2F transcription factors in rice cells (Rossi and Varotto, 2002).

Our LUC results showed that the tandem sequence consisting of cis-elements could bind to four typical transcription factors (*OsE2Fa-1*, *OsE2Fa-2*, *OsE2Fa-3*, and *OsE2Fc*) and atypical *OsE2Fe-1*, which also confirms that atypical E2F does not affect DNA binding in the absence of the transcriptional activation domain. Our activation assays showed that other rice E2F can directly or indirectly activate s *OsMSH2*, *OsMSH3*, *OsMSH6*, and *OsMSH7*, except *OsE2Fe-2*. These results confirmed the prediction that the cis-acting element of each gene could bind to 1-3 transcription factors and validated previous microarray analysis of the predicted target loci. In addition to the previously predicted *OsMSH3* and *OsMSH6*, our results suggest that *OsMSH2* and *OsMSH7* are also the targets of E2F (Vandepoele et al., 2005).

OsE2Fa-1 activates to OsMSH2, OsMSH6, and OsMSH7 (Figure 2), where OsMSH2 and OsMSH6 form MutSα heterodimers. MutSα recognizes single base mismatches such as polymerase errors, small insertion/deletion loops, and oxidative mismatches and methylation mismatches, which may be caused by translocation synthesis (TLS), (Zhang et al., 2002; Iyer et al., 2006; Campregher et al., 2008; Lippard, 2009). MSH2 and MSH7 form MutSγ heterodimers, which mainly recognize single base mismatches, including G/G, G/A, A/A, and C/A mismatches (Culligan and Hays, 2000; Wu et al., 2003; Gomez and Spampinato, 2013). In addition, MSH2 and MSH3 form MutSβ heterodimers, which sense insertion/deletion cycles and inter-strand cross-links (ICLs) (Marti et al., 2010). All of these suggest that *OsE2Fa-1* is a key transcription factor that activates MMR-related genes in canonical E2F. Other canonical E2F genes such as *OsE2Fa-2* activation of *OsMSH3* and *OsE2Fc* activation of *OsMSH7* also fully demonstrate that canonical E2F does not act alone and needs to cooperate with other MMR genes in MutSα, MutSβ, and MutSγ heterodimers.

The cadmium stress tests of the mutant and wild-type rice showed that the height of plants without cadmium stress and under 2 mg/L cadmium stress differed the most, possibly because rice leaves were most sensitive to 2 mg/L $CdCl_2$. At $CdCl_2$ > 2 mg/L, plants' defense system against cadmium stress was damaged, leading to no significant difference (Figure 3). The differences in plant height between wild-type and mutant rice under different cadmium stresses reflected the important roles of *OsMSH2*, *OsMSH3*, *OsMSH6*, and *OsMSH7* in rice response to cadmium stress, indicating that MutS is involved in rice cadmium stress response. In *Arabidopsis*, the root growth of MSH2 or MSH6-deficient seedlings under Cd stress (1.25-4.0 mg/L) was much more inhibited than that of wild-type seedlings, which was consistent with our results (Cao et al., 2018).

Yeast heterologous expression assays of E2F genes showed significant differences in the growth of E2F-expressing yeast

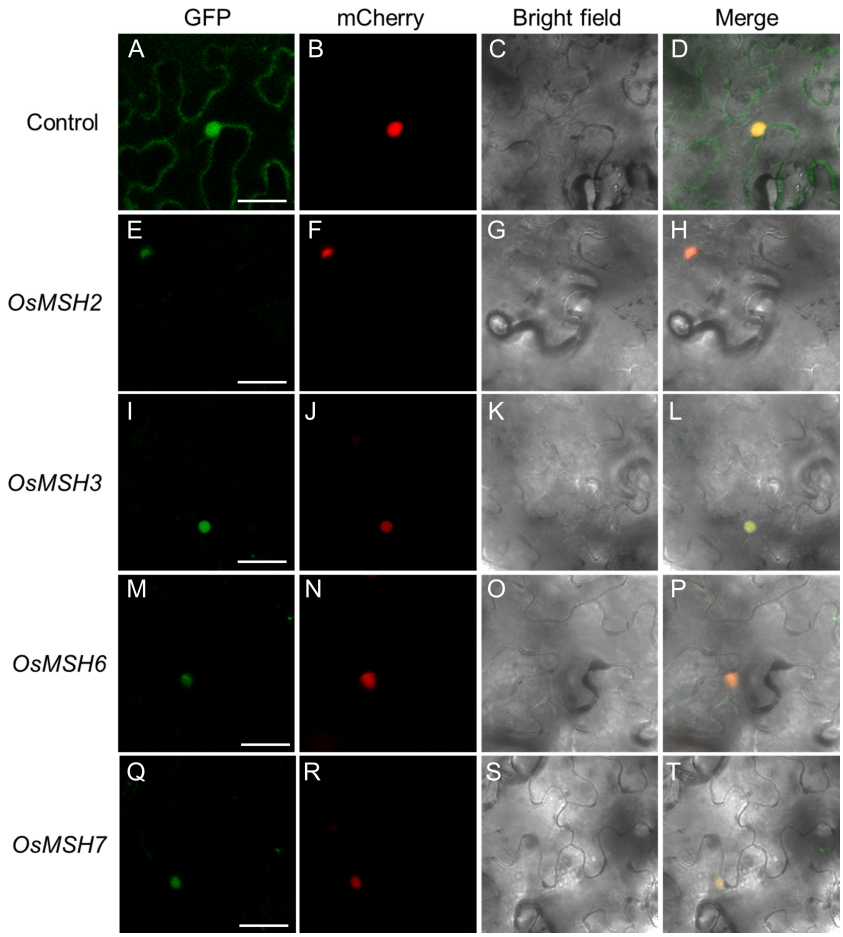


FIGURE 6
Subcellular localization of OsMSH2, OsMSH3, OsMSH6 and OsMSH7. **(A)** 35S::GFP control vector. **(E)** 35S::OsMSH2::GFP. **(I)** 35S:: OsMSH3::GFP. **(M)** 35S:: OsMSH6::GFP. **(Q)** 35S:: OsMSH7::GFP. **(B–R)** AtWRKY25-mCherry: a marker anchored in the cell nucleus. **(C–S)** Bright field. **(D–T)** Merged images. Scale bars: 20 μm.

strains under 5 μM and 10 μM cadmium treatments, with yeast strains expressing *OsE2Fa-1*, *OsE2Fa-2* and *OsE2Fc* growing slightly better than those expressing other genes, possibly because they are typical E2Fs with transcriptional activation domains that bind to downstream target genes (Figure 5). The results of LUC assay of E2F binding to the downstream *MSHs* gene and the phenotype of each *msh* mutant seedling under cadmium stress further indicate that E2F responds to cadmium stress in yeast by mobilizing its downstream MSH components.

Cadmium disturbs epigenetic modification and induces DNA damage in mouse preimplantation embryos and soybeans (Zhao et al., 2020; Cheng et al., 2021; Zhu et al., 2021). Cadmium exposure can affect oxidative stress, cell cycle arrest, DNA damage, and apoptosis in green crabs (*Scylla paramamosain*). Once cadmium-induced DNA damage is detected in plants, DDR is triggered. DDR is a cellular response to DNA damage, including regulation of DNA damage recognition and recruitment of DNA repair factors. Inhibition of cell cycle regulatory genes during DDR contributes to cell proliferation arrest (Jackson and Bartek, 2009; Schade et al., 2019). Therefore, genes related to mismatch repair and cell cycle regulation are, in principle, located in the nucleus. Our experiments demonstrated that *OsMSH2/3/6/7* are localized in the nucleus and function together with cell cycle regulatory genes in response to cadmium stress (Figure 6).

Adaptation to changes in cellular nutrition and external environments is a fundamental cellular property. This adaptation requires the coordination of multiple networks among metabolic, growth, and cell cycle regulators, including the CDKfamily, members of the RBfamily, and E2F transcription factors (Huber et al., 2021). In response to external growth stimuli, including plant hormones, the abundance of specific G1 cell cycle proteins increases (Riou-Khamlichi et al., 2000). The CYCD-CDKA;1 complex

phosphorylates RBR at multiple conserved sites, releasing activator E2F from bound RBR, thereby inducing the expression of downstream cell cycle genes (Nakagami et al., 2002; Magyar et al., 2012). Our qRT-PCR further confirmed the RNA-seq results, showing significant upregulation of *CDKA;1* and *CYCD* as the front end of G1/S activation, significant downregulation of *RBR* as the response to being phosphorylated, and significant upregulation of *OsE2Fa-1* and *OsE2Fa-2* due to being released from RBR (Figures S2B, C). All the results of LUC assay, yeast spot plate assay, and qRT-PCR indicated that the mobilization of E2F transcription factors is the key factor responsible for significant upregulation of *OsMSH2*, *OsMSH3*, *OsMSH6*, or *OsMSH7* under cadmium stress. *OsMSH2*, *OsMSH3*, *OsMSH6*, or *OsMSH7* were significantly downregulated in rice leaves in response to cadmium stress, contrary to their expression patterns in *Arabidopsis* roots. This discrepancy is possibly because cadmium stress induces different damages in rice leaves and *Arabidopsis* roots or because rice and *Arabidopsis* have different regulation patterns in response to cadmium stress. As shown in Figure S2A, the rice MMR system was not damaged, and *MSH2/3/6/7* was upregulated to repair DNA damage. A fully functional MMR system can regulate the G2/M phase by upregulating G2/M regulatory proteins and/or by activating *p53*, *ATM*, and *ATR* signaling pathways in human cells in response to exogenous and endogenous stresses. Therefore, we selected key G2/M genes (*CDKB* and *MAD2*) and other DNA stress sensors that coordinate stress responses with cell cycle checkpoint *ATR* genes (Wang et al., 2013). We found that both *CDKB* and *MAD2* were significantly downregulated after the knockdown of either MSH genes (Figures 4L-M) and significantly upregulated under cadmium stress, suggesting that MSH genes are also involved in the cell cycle regulation of G2/M. *OsMSH2* is significantly downregulated in *Osmsh3*, *Osmsh6*, and *Osmsh7* mutant rice, indicating an important role of *OsMSH2* in MutSα, MutSβ, and MutSγ complexes. *OsMSH3*, *OsMSH6*, and *OsMSH7* expression levels are altered in all mutant rice, indicating that the MutS system is a comprehensive complex involving the whole body (Figures 4B-E). Interestingly, *CycD5;2* expression was significantly upregulated in wild-type rice under cadmium stress and significantly downregulated in all mutants. Moreover, their expression was further significantly downregulated in the leaves after exposure of the mutants to cadmium stress. The expression of *CYCD-CDKA* genes was also changed under cadmium stress, as shown in transcriptome and qRT-PCR analyses. Therefore, we speculate that *CYCD-CDKA* may directly target and regulate MSH genes in G1/S for DNA damage repair in addition to activating downstream RBR phosphorylation (Figures 4B-E). However, this speculation needs to be proven by additional tests at a later stage. Based on our findings, we propose that the *CDKA-RBR-E2F-MSH* pathway may be the primary mechanism for enhancing cadmium stress tolerance in rice by mobilizing cell cycle factors of downstream key mismatch repair genes (Figure 7).

In conclusion, we systematically categorized and analyzed rice E2F genes by bioinformatics and demonstrated that rice E2F

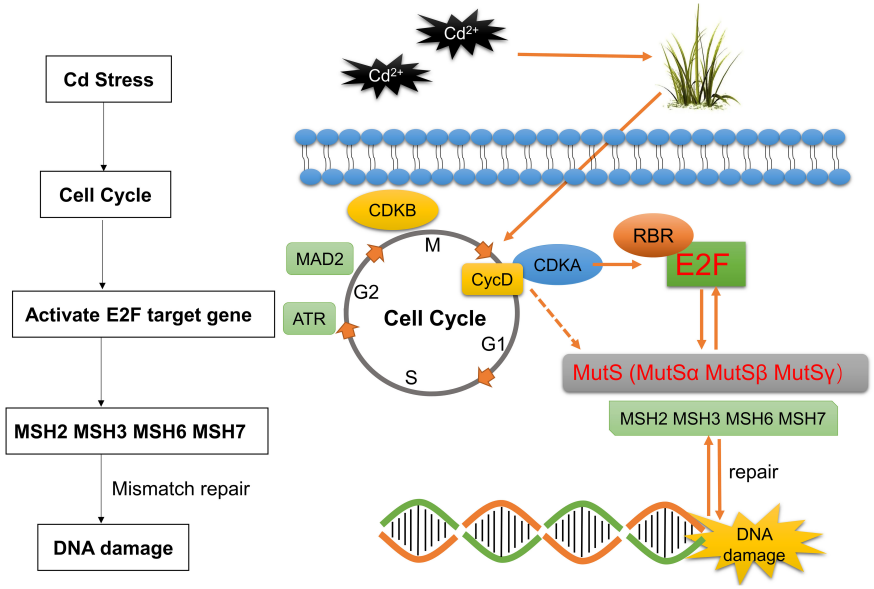


FIGURE 7
Model of Cd-induced cell cycle regulation in rice leaves. The solid line represents the critical pathway, the dashed line represents the possible pathway, and the technical route to the model is explained on the left.

transcription factor can activate MSHs target genes. The cadmium stress experiments in rice seedlings and the yeast heterologous expression system demonstrated that E2F transcription factors and MSHs can respond to cadmium stress and elucidated the specific link between E2F and MSHs from transcriptomic and quantitative perspectives. Furthermore, our study provided a preliminary theoretical basis for revealing the intrinsic mechanism of E2F transcription factors responding to cadmium stress by binding to MSH mismatch repair genes E2F in the cell cycle. Next, we will take this as a breakthrough to further explore the relationship between E2F factor and cadmium stress in rice, and study the relationship between E2F and MSH in rice and other crops from multiple perspectives and layers. Our work paved a novel way for expanding the theoretical basis of the plant cell cycle and provided a theoretical reference for the study of the mechanism responding to cadmium stress in rice.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI Sequence Read Archive under the BioProject identification number PRJNA890021.

Author contributions

B-GM and B-RZ designed the experiments. W-JZ carried out most of the experiments. W-QL, YP, YS, LT, C-TL, DZ, L-JZ and J-HL. assisted in phenotypic identification and protein interaction tests. W-QL and Z-CY assisted in field management. W-JZ wrote the manuscript. B-GM and B-RZ revised and approved the final version of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by National Natural Science Foundation of China (32172042), Natural Science Foundation of Hunan Province (2021JJ30487), the earmarked fund for China Agriculture Research System. The authors would like to thank TopEdit (www.topeditsci.com) for its linguistic assistance during the preparation of this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1068769/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Amino acid sequence analysis of rice and *Arabidopsis* E2Fs. According to amino acid similarity, gray, green and orange represent identical or similar amino acids. Highly homologous DNA-binding structural domains that are repeated twice in rice and *Arabidopsis* E2F family proteins, as well as conserved regions, are marked.

SUPPLEMENTARY FIGURE 2
Expression profiles of cell cycle-related genes in response to Cd stress in wild-type leaves of rice and statistics of rice root length differences after Cd stress. **(A–C)** Statistics of differential expression of cell cycle-related genes in rice wild-type leaves under 2 mg and 4 mg Cd stress. Data were shown mean ± SD at least three independent experiments, and house-keeping gene Osactin was used as an internal control. * and # significantly statistical difference from the WT control and the corresponding mutant control, respectively ($P < 0.05$). **(D)** Statistics of root length differences between rice msh mutant and wild-type seedlings under 2 mg and 4 mg stress. (Student's t-test, $p < 0.05$).

SUPPLEMENTARY FIGURE 3
Statistical analysis of the number of differenced genes in rice wild-type and MSH mutants under 2mg and 4mg cadmium stress **(A)** Overall statistics of the number of up and down-regulated genes **(B)** Venn diagram statistics of differential gene intersections and non-intersections between specific wild types and individual mutants

References

Campregher, C., Luciani, M. G., and Gasche, C. (2008). Activated neutrophils induce an hMSH2-dependent G2/M checkpoint arrest and replication errors at a (CA)13-repeat in colon epithelial cells. *Gut* 57, 780–787. doi: 10.1136/gut.2007.141556

Cao, X., Wang, H., Zhuang, D., Zhu, H., Du, Y., Cheng, Z., et al. (2018). Roles of MSH2 and MSH6 in cadmium-induced G2/M checkpoint arrest in arabidopsis roots. *Chemosphere* 201, 586–594. doi: 10.1016/j.chemosphere.2018.03.017

Chang, Y. M., Lin, H. H., Liu, W. Y., Yu, C. P., Chen, H. J., Wartini, P. P., et al. (2019). Comparative transcriptomics method to infer gene coexpression networks and its applications to maize and rice leaf transcriptomes. *Proc. Natl. Acad. Sci. U. S. A.* 116, 3091–3099. doi: 10.1073/PNAS.1817621116

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Cheng, C. H., Ma, H. L., Deng, Y. Q., Feng, J., Jie, Y. K., and Guo, Z. X. (2021). Oxidative stress, cell cycle arrest, DNA damage and apoptosis in the mud crab (Scylla paramamosain) induced by cadmium exposure. *Chemosphere* 263, 128277. doi: 10.1016/j.chemosphere.2020.128277

Cui, W., Wang, H., Song, J., Cao, X., Rogers, H. J., Francis, D., et al. (2017). Cell cycle arrest mediated by cd-induced DNA damage in arabidopsis root tips. *Ecotoxicol. Environ. Safety*. 145, 569–574. doi: 10.1016/j.ecoenv.2017.07.074

Culligan, K. M., and Hays, J. B. (2000). Arabidopsis MutS homologs-AtMSH2, AtMSH3, AtMSH6, and a novel AtMSH7-form three distinct protein heterodimers with different specificities for mismatched DNA. *Plant Cell.* 12, 991–1002. doi: 10.1105/tpc.12.6.991

del Pozo, J. C., Diaz-Trivino, S., Cisneros, N., and Gutierrez, C. (2006). The balance between cell division and endoreplication depends on E2FC-DPB, transcription factors regulated by the ubiquitin-SCFSKP2A pathway in arabidopsis. *Plant Cell.* 18, 2224–2235. doi: 10.1105/tpc.105.039651

Dimova, D. K., and Dyson, N. J. (2005). The E2F transcriptional network: old acquaintances with new faces. *Oncogene* 24, 2810–2826. doi: 10.1038/sj.onc.1208612

Filipic, M. (2012). Mechanisms of cadmium induced genomic instability. *Mutat. Res.* 733, 69–77. doi: 10.1016/j.mrfmmm.2011.09.002

Gomez, R., and Spampinato, C. P. (2013). Mismatch recognition function of arabidopsis thaliana MutSγ. *DNA Repair*. 12, 257–264. doi: 10.1016/j.dnarep.2013.01.002

Guo, J., Song, J., Wang, F., and Zhang, X. S. (2007). Genome-wide identification and expression analysis of rice cell cycle genes. *Plant Mol. Biol.* 64, 349–360. doi: 10.1007/s11103-007-9154-y

Gutierrez, C. (2009). The arabidopsis cell division cycle. *Arabidopsis Book.* 7, e0120. doi: 10.1199/tab.0120

Hellens, R. P., Allan, A. C., Friel, E. N., Bolitho, K., Grafton, K., Templeton, M. D., et al. (2005). Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* 1, 1–14. doi: 10.1186/1746-4811-1-13

Huber, K., Mestres-Arenas, A., Fajas, L., and Leal-Esteban, L. C. (2021). The multifaceted role of cell cycle regulators in the coordination of growth and metabolism. *FEBS J.* 288, 3813–3833. doi: 10.1111/febs.15586

Hu, Z., Cools, T., and De Veylder, L. (2016). Mechanisms used by plants to cope with DNA damage. *Annu. Rev. Plant Biol.* 67, 439–462. doi: 10.1146/annurev-arplant-043015-111902

Hu, Q., Li, Y., Wang, H., Shen, Y., Zhang, C., Du, G., et al. (2017). Meiotic chromosome association 1 interacts with TOP3α and regulates meiotic recombination in rice. *Plant Cell.* 29, 1697–1708. doi: 10.1105/tpc.17.00241

Ijsselsteijn, R., Jansen, J. G., and de Wind, N. (2020). DNA Mismatch repair-dependent DNA damage responses and cancer. *DNA Repair*. 93, 102923. doi: 10.1016/j.dnarep.2020.102923

Iyer, R. R., Pluciennik, A., Burdett, V., and Modrich, P. L. (2006). DNA Mismatch repair: functions and mechanisms. *Chem. Rev.* 106, 302–323. doi: 10.1021/cr0404794

Jackson, S. P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature* 461, 1071–1078. doi: 10.1038/nature08467

Jiang, M., Wu, X., Song, Y., Shen, H., and Cui, H. (2020). Effects of OsMSH6 mutations on microsatellite stability and homeologous recombination in rice. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00220

Karthika, V., Chandrashekar, B. K., Kiranmai, K., Ag, S., Makarla, U., Ramu, V.S, et al. (2021). Disruption in the DNA mismatch repair gene MSH2 by CRISPR-Cas9 in indica rice can create genetic variability. *J. Agric. Food Chem.* 69, 4144–4152. doi: 10.1021/acs.jafc.1c00328

Kosugi, S., and Ohashi, Y. (2002). E2F sites that can interact with E2F proteins cloned from rice are required for meristematic tissue-specific expression of rice and tobacco proliferating cell nuclear antigen promoters. *Plant. J.* 29 (1), 45–59. doi: 10.1046/j.1365-313x.2002.01196.x

Lippard, Z. S. J. (2009). Photoaffinity labeling reveals nuclear proteins that uniquely recognize cisplatin-DNA interstrand cross-links. *Biochemistry* 48, 4916–4925. doi: 10.1021/bi900389b

Liu, W., Zhou, Q., Li, P., Gao, H., Han, Y. P., Li, X. J., et al. (2009). DNA Mismatch repair related gene expression as potential biomarkers to assess cadmium exposure in arabidopsis seedlings. *J. Hazardous Mater.* 167, 1007–1013. doi: 10.1016/j.jhazmat.2009.01.093

Livak, K. J., and Schmittgen, T. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods* 25, 402–408. doi: 10.1016/j.jhazmat.2009.01.09310.1006/meth.2001.1262

Magyar, Z., Horváth, B., Khan, S., Mohammed, B., Henriques, R., De Veylder, L., et al. (2012). Arabidopsis E2FA stimulates proliferation and endocycle separately through RBR-bound and RBR-free complexes. *EMBO J.* 31, 1480–1493. doi: 10.1038/emboj.2012.13

Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.1093/bioinformatics/bti430

Mariconti, L., Pellegrini, B., Cantoni, R., Stevens, R., Bergounioux, C., Cella, R., et al. (2002). The E2F family of transcription factors from arabidopsis thaliana. novel and conserved components of the retinoblastoma/E2F pathway in plants. *J. Biol. Chem.* 277, 9911–9919. doi: 10.1074/jbc.M110616200

Marti, T. M., Kunz, C., and Fleck, O. (2010). DNA Mismatch repair and mutation avoidance pathways. *J. Cell. Physiol.* 191, 28–41. doi: 10.1016/s1369-5266(00)00134-5

Meijer, M., and Murray, J. A. (2001). Cell cycle controls and the development of plant form. *Curr. Opin. Plant Biol.* 4, 44–49. doi: 10.1016/s1369-5266(00)00134-5

Mironov, V., Montagu, M. V., and Inzé, D. (1997). Regulation of cell division in plants: An arabidopsis perspective. *Prog. Cell Cycle Res.* 3, 29–41. doi: 10.1007/978-1-4615-5371-7_3

Morshed, S., Shibata, T., Naito, K., Miyasato, K., Takeichi, Y., Takuma, T., et al. (2020). TORC1 regulates G1/S transition and cell proliferation *via* the E2F homologs MBF and SBF in yeast. *Biochem. Biophys. Res. Commun.* 529, 846–853. doi: 10.1016/j.bbrc.2020.05.122

Nakagami, H., Kawamura, K., Sugisaka, K., and Shinmyo, S. A. (2002). Phosphorylation of retinoblastoma-related protein by the cyclin d/cyclin-dependent kinase complex is activated at the G1/S-phase transition in tobacco. *Plant Cell.* 14, 1847–1857. doi: 10.1105/tpc.002550

Pena, L. B., Barcia, R. A., Azpilicueta, C. E., Méndez, A., and Gallego, S. M. (2012). Oxidative post translational modifications of proteins related to cell cycle are involved in cadmium toxicity in wheat seedlings. *Plant Sci.* 196, 1–7. doi: 10.1016/j.plantsci.2012.07.008

Perrotta, L., Giordo, R., Francis, D., Rogers, H. J., and Albani, D. (2021). Molecular analysis of the E2F/DP gene family of daucus carota and involvement of the DcE2F1 factor in cell proliferation. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.652570

Ponnamperuma, F. N. (1977). Screening rice for tolerance to mineral stresses. *Irri Res. Pap Ser. Int. Rice Res. Inst* 26.

Qi, F., and Zhang, F. (2019). Cell cycle regulation in the plant response to stress. *Front. Plant Sci.* 10 1765. doi: 10.3389/fpls.2019.01765

Ramirez-Parra, E. (2003). A genome-wide identification of E2F-regulated genes in arabidopsis. *Plant J.* 33, 801–811. doi: 10.1046/j.1365-313x.2003.01662.x

Riou-Khamlichi, C., Menges, M., Healy, J. M., and Murray, J. A. (2000). Sugar control of the plant cell cycle: Differential regulation of arabidopsis d-type cyclin gene expression. *Mol. Cell. Biol.* 20, 4513–4521. doi: 10.1128/MCB.20.13.4513-4521.2000

Rossi, V., and Varotto, S. (2002). Insights into the G1/S transition in plants. *Planta* 215, 345–356. doi: 10.1007/s00425-002-0780-y

Schade, A. E., Fischer, M., and DeCaprio, J. A. (2019). RB, p130 and p107 differentially repress G1/S and G2/M genes after p53 activation. *Nucleic Acids Res.* 47, 11197–11208. doi: 10.1093/nar/gkz961

Schofield, M. J., and Hsieh, P. (2003). DNA Mismatch repair: molecular mechanisms and biological function. *Annu. Rev. Microbiol.* 57, 579–608. doi: 10.1146/annurev.micro.57.030502.090847

Sparkes, I. A., Runions, J., Kearns, A., and Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.* 1, 2019–2025. doi: 10.1038/nprot.2006.286

Stals, H., and Inzé, D. (2001). When plant cells decide to divide. *Trends Plant Sci.* 6, 359–364. doi: 10.1016/s1360-1385(01)02016-7

Tan, L., Zhu, Y., Fan, T., Peng, C., Wang, J., Sun, L., et al. (2019). OsZIP7 functions in xylem loading in roots and inter-vascular transfer in nodes to deliver Zn/Cd to grain in rice. *Biochem. Biophys. Res. Commun.* 512, 112–118. doi: 10.1016/j.bbrc.2019.03.024

Uraguchi, S., and Fujiwara, T. (2012). Cadmium transport and tolerance in rice: perspectives for reducing grain cadmium accumulation. *Rice* 5, 5. doi: 10.1186/1939-8433-5-5

Vandepoele, K., Vlieghe, K., Florquin, K., Hennig, L., Beemster, G. T. S., Gruissem, W., et al. (2005). Genome-wide identification of potential plant E2F target genes. *Plant Physiol.* 139, 316–328. doi: 10.1104/pp.105.066290

Veylder, L. D., Joubès, J., and Inzé, D. (2003). Plant cell cycle transitions. *Curr. Opin. Plant Biol.* 6, 536–543. doi: 10.1016/j.pbi.2003.09.001

Wang, H., Cao, Q., Zhao, Q., Arfan, M., and Liu, W. (2020). Mechanisms used by DNA MMR system to cope with cadmium-induced DNA damage in plants. *Chemosphere* 246, 125614. doi: 10.1016/j.chemosphere.2019.125614

Wang, H., He, L., Song, J., Cui, W., Zhang, Y., Jia, C., et al. (2016). Cadmium-induced genomic instability in arabidopsis: Molecular toxicological biomarkers for early diagnosis of cadmium stress. *Chemosphere* 150, 258–265. doi: 10.1016/j.chemosphere.2016.02.042

Wang, J., Huang, S., Xing, L., Shen, H., Yan, X., Wang, J., et al. (2013). Role of hMLH1 in sterigmatocystin-induced G2 phase arrest in human esophageal epithelial het-1A cells *in vitro*. *Toxicol. Letters.* 217, 226–234. doi: 10.1016/j.toxlet.2012.12.020

Weintraub, S. J., Prater, C. A., and Dean, D. C. (1992). Retinoblastoma protein switches the E2F site from positive to negative element. *Nature* 358, 259–261. doi: 10.1038/358259a0

Wu, S. Y., Culligan, K., Lamers, M., and Hays, J. (2003). Dissimilar mispair-recognition spectra of arabidopsis DNA-mismatch-repair proteins MSH2*MSH6 (MutSalpha) and MSH2*MSH7 (MutSgamma). *Nucleic Acids Res.* 31, 6027–6034. doi: 10.1093/nar/gkg780

Xiang, Y., Laurent, B., Hsu, C. H., Nachtergaele, S., Lu, Z., Sheng, W., et al. (2017). m6A RNA methylation regulates the UV-induced DNA damage response. *Nature* 543, 573–576. doi: 10.1038/nature21671

Yoshioka, K. I., Yoshioka, Y., and Hsieh, P. (2006). ATR kinase activation mediated by MutSalpha and MutLalpha in response to cytotoxic O6-methylguanine adducts. *Mol. Cell.* 22, 501–510. doi: 10.1016/j.molcel.2006.04.023

Zhang, N., Lu, X., Zhang, X., Peterson, C. A., and Legerski, R. J. (2002). hMutSβ is required for the recognition and uncoupling of psoralen interstrand cross-links *In vitro*. *Mol. Cell. Biol.* 22, 2388–2397. doi: 10.1128/MCB.22.7.2388-2397.2002

Zhang, Y., Yuan, F., Presnell, S. R., Tian, K., Gao, Y., Tomkinson, A. E., et al. (2005). Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell* 122, 693–705. doi: 10.1016/j.cell.2005.06.027

Zhao, Q., Wang, H., Du, Y., Rogers, H., Wu, Z., Jia, S., et al. (2020). MSH2 and MSH6 in mismatch repair system account for soybean ((L.) merr.) tolerance to cadmium toxicity by determining dna damage response. *J. Agric. Food Chem.* 68, 1974–1985. doi: 10.1021/acs.jafc.9b06599

Zhu, J., Huang, Z., Yang, F., Zhu, M., Cao, J., Chen, J., et al. (2021). Cadmium disturbs epigenetic modification and induces DNA damage in mouse preimplantation embryos. *Ecotoxicol. Environ. Safety.* 219, 112306. doi: 10.1016/j.ecoenv.2021.112306

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 30 November 2022
DOI 10.3389/fpls.2022.1065253

OPEN ACCESS

EDITED BY
Xueyong Li,
Institute of Crop Sciences (CAAS), China

REVIEWED BY
Fuhao Cui,
China Agricultural University, China
Feng-Zhu Wang,
Sun Yat-sen University, China
Cai Lin Lei,
Chinese Academy of Agricultural Sciences (CAAS), China
Wen-Ming Wang,
Sichuan Agricultural University, China

*CORRESPONDENCE
Wenxue Zhai
wxzhai@genetics.ac.cn
Chunrong Li
lichunrong.0537@163.com

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Plant Bioinformatics, a section of the journal Frontiers in Plant Science

RECEIVED 09 October 2022
ACCEPTED 10 November 2022
PUBLISHED 30 November 2022

CITATION
Li Y, Liu P, Mei L, Jiang G, Lv Q, Zhai W and Li C (2022) Knockout of a papain-like cysteine protease gene *OCP* enhances blast resistance in rice. *Front. Plant Sci.* 13:1065253. doi: 10.3389/fpls.2022.1065253

Knockout of a papain-like cysteine protease gene *OCP* enhances blast resistance in rice

Yuying Li[1,2†], Pengcheng Liu[3†], Le Mei[1,2], Guanghuai Jiang[1], Qianwen Lv[1,2], Wenxue Zhai[1*] and Chunrong Li[1*]

[1]Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]College of Chemistry and Life Sciences, Zhejiang Normal University, Jinhua, China

Papain-like cysteine proteases (PLCPs) play an important role in the immune response of plants. In Arabidopsis, several homologous genes are known to be involved in defending against pathogens. However, the effects of PLCPs on diseases that afflict rice are largely unknown. In this study, we show that a PLCP, an oryzain alpha chain precursor (OCP), the ortholog of the Arabidopsis protease RD21 (responsive to dehydration 21), participates in regulating resistance to blast disease with a shorter lesion length characterizing the knockout lines (*ocp-ko*), generated *via* CRISPR/Cas9 technology. OCP was expressed in all rice tissues and mainly located in the cytoplasm. We prove that OCP, featuring cysteine protease activity, interacts with OsRACK1A (receptor for activated C kinase 1) and OsSNAP32 (synaptosome-associated protein of 32 kD) physically *in vitro* and *in vivo*, and they co-locate in the rice cytoplasm but cannot form a ternary complex. Many genes related to plant immunity were enriched in the *ocp-ko1* line whose expression levels changed significantly. The expression of jasmonic acid (JA) and ethylene (ET) biosynthesis and regulatory genes were up-regulated, while that of auxin efflux transporters was down-regulated in *ocp-ko1*. Therefore, OCP negatively regulates blast resistance in rice by interacting with OsRACK1A or OsSNAP32 and influencing the expression profiles of many resistance-related genes. Moreover, OCP might be the cornerstone of blast resistance by suppressing the activation of JA and ET signaling pathways as well as promoting auxin signaling pathways. Our research provides a comprehensive resource of PLCPs for rice plants in defense against pathogens that is also of potential breeding value.

KEYWORDS

blast resistance, oryzain alpha chain precursor (OCP), OsRACK1A, OsSNAP32, papain-like cysteine proteases (PLCPs), rice (*Oryza sativa*)

Introduction

The most destructive disease of cultivated rice is blast, caused by the fungus *Magnaporthe oryzae* (Zhai et al., 2022), which results in severe yield losses of about 30% (Yin et al., 2021). To defend against pathogens, plants have evolved complex immune systems, one is the basal defense defined by pattern recognition receptors, and the other is the immune response mediated by *Resistance* (*R*) genes (Yu et al., 2021). It is widely accepted that generating rice germplasm resources with *R* genes is the most economical and eco-friendly strategy to defend crops against blast (Xiao et al., 2020). Therefore, it is imperative we discover new *R* genes for controlling blast disease. Some blast-resistance genes have been cloned already and applied in crop breeding for disease resistance (Li et al., 2020). For example, *Pita* (*pyricularia oryzae resistance-ta*), encoding a major *R*-gene type, confers resistance to *M. oryzae* races containing the corresponding avirulence gene *AVR-Pita* (Bryan et al., 2000). The *pi21* (*pyricularia oryzae resistance 21*), encoding a proline-rich protein, is a non-race specific recessive gene that maintains resistance permanently, although this trait is incomplete in comparison with that triggered by *R* genes (Fukuoka et al., 2009; Nawaz et al., 2020; Tao et al., 2021). Moreover, there are some non-*R* genes in rice that can participate in blast resistance. Transgenic lines featured increased blast resistance when the SNAP25-type gene *OsSNAP32* was overexpressed, which encodes a soluble N-ethylmaleimide-sensitive-factor attachment protein receptor, whose expression is induced by blast fungus inoculation in rice seedlings (Bao et al., 2008; Luo et al., 2016). OsSYP121 (i.e., syntaxin of plants 121) can interact with OsSNAP32 and VAMP714/724 (vesicle-associated membrane protein 714/724) to form a complex that mediated host resistance to rice blast (Cao et al., 2019). Besides, many transcription factors and enzymes also are involved in blast resistance, such as *OsWRKY45* (*WRKY gene 45*) (Shimono et al., 2007) and *OsPAL1* (*phenylalanine ammonia-lyase 1*) (Zhou et al., 2018), to name two. Nevertheless, due to the diversity and complexity of pathogenic populations, it is difficult for rice blast resistance-related genes to effectively maintain their resistance or, if they do, it only applies to limited regions (Tao et al., 2021). Accordingly, it is necessary to continuously uncover more *R* genes to blast for rice crop improvement.

Cysteine proteases function critically in plant growth and development, of which the PLCPs are notable for being involved in protein maturation and degradation, plant senescence, seed germination, and programmed cell death (PCD) (Grudkowska and Zagdanska, 2004; Liu et al., 2018). Moreover, PLCPs play key roles in plant immune systems by inducing systemic immunity and degrading the pathogen-effector protein (Perez-Lopez et al., 2021). For example, in Arabidopsis, the *rd21a* mutants were more susceptible to the fungal pathogen *Botrytis cinerea* (Shindo et al., 2012), and knocking out *rd21a* inhibited flg22-triggered stomatal closure, which led to lowered resistance to *Pseudomonas syringae* (Liu et al., 2020). Null *XCP1* (*xylem cysteine peptidase 1*) or *XCP2* (*xylem cysteine peptidase 2*) mutants display increased resistance to pathogens (Zhang et al., 2014; Perez-Lopez et al., 2021). In tomato, C14/CYP1, targeted by the *Phytophthora infestans* effector AvrBlb2, plays a role in pathogen defense in that silencing *C14* increased susceptibility to *P. infestans* (Kaschani et al., 2010; Misas-Villamil et al., 2016). For rice, there are few reports about how PLCPs affect its growth and development. *OsCP1* (*cysteine protease 1*), a cysteine protease gene, was shown to influence pollen development and regulate PCD (Lee et al., 2004; Li et al., 2011). Yet whether and how the PLCPs function in rice immunity remains unclear.

Here we identified a PLCP, an oryzain alpha chain precursor (OCP), which is capable of negatively regulating rice blast resistance. Knocking out *OCP* resulted in the accumulation of mRNA for defense-related genes and shortened the lesion length of transgenic plants (*ocp-ko*) inoculated with blast isolates when compared with TP309. We find that *OCP* is highly conserved in plants and possesses the cysteine protease activity. By screening a yeast library, two rice proteins related to blast resistance were obtained, namely OsRACK1A and OsSNAP32. Many blast resistance genes are up-regulated in *ocp-ko* plants, which meant that OCP probably negatively regulates blast resistance by repressing the related gene expression. Therefore, this study not only fills the knowledge gap of PLCPs in disease resistance of rice but also provides effective and promising gene resources for use in future rice breeding.

Materials and methods

Plant materials and growth conditions

The *Japonica* rice cultivar TP309 was used for the transgenic experiments. All plants were cultivated in the experimental field of the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences. To generate gene overexpression, the coding sequence (CDS) of *OCP* was amplified from the cDNA of TP309 and cloned into the vector UBI-pCAMBIA1300; the knockout mutants were created using the CRISPR/Cas9 system. All the constructs were transformed into rice calli *via* *Agrobacterium*-mediated transformation.

Structural analysis and construction of the phylogenetic tree of orthologous proteins

Amino acid sequences of the plant species and corresponding accession numbers were retrieved from the

NCBI (https://www.ncbi.nlm.nih.gov/), RGAP (http://rice.uga.edu/index.shtml), and WheatOmics 1.0 (http://wheatomics.sdau.edu.cn/). The structures of orthologous proteins were drawn with SMART (http://smart.embl-heidelberg.de/). Based on the alignment of the amino acid sequences with the Muscle program, and using 1000 bootstrap replicates, a neighbor-joining tree was constructed in MEGA7 software. Multiple sequence alignment of the proteins was carried out by MEGA7 and the results were edited by GeneDoc software.

Blast inoculation assays

For the blast fungal spray inoculation assays, the plants were grown in a greenhouse at 28°C under a 12-h light/12-h dark photoperiod for 14 days. The assay was performed as described by Chen et al. (2018). The spore concentration was adjusted to 1×10^5 cfu/mL with 0.2% Tween-20, and the inoculated rice seedlings were kept in a dark chamber at 28°C for 24 h, and then moved into the greenhouse. The injection method for testing blast resistance in the field was followed (Lv et al., 2013). The seedlings at the tillering stage were injected with spore suspension, and the leaf status was observed 7 days after inoculation.

For the punch inoculation assays, 40-day-old rice plants were inoculated by following a previously described methodology (Park et al., 2012), albeit with slight modifications. The rice leaves were lightly wounded with 10-μL pipette tips on a 1.5-cm scale and put on the surface of 6-Benzylaminopurine, and the spore suspension was added onto the wound site. The ensuing lesion length was measured at 8 days post-inoculation in the greenhouse. The *M. oryzae* isolates 97-27-2, JL021605, and ZB13 were used in this study.

GUS staining

GUS activity was analyzed in transgenic plants ***via*** histochemical staining with 5-bromo-4-chloro-3-indolyl-b-Dglucuronicacid (X-Gluc), as described previously (Dong et al., 2021). The rice tissues were incubated for 16 h at 37°C in a staining buffer (100 mM sodium phosphate [pH 7.0], 10 mM EDTA, 0.5 mM $K_4Fe(CN)_6$, 0.5 mM $K_3Fe(CN)_6$, 0.1% [v/v] Triton X-100, and 1 mM X-Gluc), and then decolorized in 100% ethanol before photographing them.

Rice protoplast preparation and transformation

Rice protoplasts were prepared from 2-week-old seedlings of TP309 that had been grown in darkness. Protoplasts were transformed as described previously (Bart et al., 2006). Plasmid constructs were transformed into the rice protoplasts, which were then kept at 28°C for 16 or 18 h. After that, we detected the fluorescence or extracted proteins.

Subcellular localization

The coding region of *OCP* was fused with the green fluorescent protein (35S::OCP-GFP) and enhanced yellow fluorescent protein (35S::OCP-eYFP), and then transformed into TP309 calli and rice protoplasts to express the fusion proteins, respectively. For subcellular co-localization, we fused the coding region of OsRACK1A or OsSNAP32 with the mCherry tag and then transformed them into rice protoplasts alone or with 35S::OCP-eYFP. The fluorescent signal was visualized using the Zeiss LSM 710 NLO microscope (Carl Zeiss, Oberkochen, Germany) after incubation at 28°C for 16 h.

Yeast hybrid assays

The coding region of *OCP* was introduced into the pGBKT7 vector (BD-OCP) as bait and co-transformed with the rice cDNA library for the screening of interacting proteins on SD/-Leu-Trp-His-Ade selected plates. For specific interactions, the truncations of *OCP* were cloned into the pGBKT7 vector and the full-length coding region of *OsRACK1A* and *OsSNAP32* were cloned into the pGADT7 vector (respectively yielding AD-OsRACK1A and AD-OsSNAP32), and then co-transformed into the yeast strain Gold Y2H. The transformants were grown on SD/-Trp-Leu medium at 30°C for 3 to 5 days and the interaction was confirmed by colony growth on SD/-Ade-His-Leu-Trp (AD/-A-H-L-T) medium with X-α-gal.

For the yeast three-hybrid (Y3H) assays, the coding region of *OCP* was cloned into the multiple cloning site (MCS) 2 of pBridge (pBridge-OCP), and the coding region of *OsSNAP32* was introduced into MCS1 of pBridge (pBridge-OsSNAP32 and pBridge-OsSNAP32-OCP). Each of these vectors was then co-transformed into Gold Y2H with AD-OsRACK1A. The transformants were grown on (SD)/-Met-Trp-Leu medium at 30°C for about 5 days, and the interaction was confirmed by colony growth on SD/-Ade-His-Leu-Met-Trp (SD/A-H-L-M-T) medium with X-α-gal for about 5 days. Specific primers used are listed in Supplementary Table 1.

Bimolecular fluorescence complementation assays

The full-length coding region of *OCP* was cloned into the pVYCE vector (cYFP-OCP), and OsRACK1A and OsSNAP32 were introduced into the pVYNE vector (nYFP-OsRACK1A and nYFP-OsSNAP32). The ensuing constructs were transformed

into the *Agrobacterium tumefaciens* strain EHA105, and then allowed to infect 5-week-old *Nicotiana benthamiana* leaves. Fluorescent signals were detected and photographed using Zeiss LSM 710 NLO microscope (Carl Zeiss, Oberkochen, Germany) after infiltration for 3 or 4 days (Cao et al., 2019).

Co-immunoprecipitation assays

To verify the interaction of OCP with OsRACK1A and OsSNAP32 *in vivo*, the recombinant vectors OCP-Myc, OsRACK1A-mCherry, and OsSNAP32-mCherry were generated. These construct pairs were transiently co-expressed in rice protoplasts. After 18 h, the protoplasts were collected by centrifugation at 150 × *g* and rinsed three times with a wash buffer (50 mM Tris-HCl [pH 7.5], 150 mM NaCl). The total proteins were extracted using a lysis buffer (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 0.5% Triton X-100, 0.5% Nonidet P-40 [NP-40], 1 μM MG132, and 1 × Protein Inhibitor Cocktail). After its 1-h incubation on ice, the lysate was centrifuged at 4°C with 15 000 × *g*. Lysates containing the target proteins were incubated with 20 μL of Myc tag-Nanoab-Agarose Beads by tumbling them for 2 h at 4°C. Next, the beads were rinsed thrice with the wash buffer and boiled for 5 min with an SDS loading buffer. The proteins were analyzed by running a Western blot assay using anti-Myc and anti-mCherry.

Cysteine protease activity profiling

Procedures for cysteine protease activity profiling were largely followed as described previously (van der Hoorn et al., 2004). L-cysteine, E-64, and DCG-04 were purchased from Lablead (Beijing, China), Bioss (Beijing, China), and MedKoo (Morrisville, USA), respectively. Proteins were extracted from *Escherichia coli* and purified. About 30 μg of protein was mixed with 50 mM sodium acetate buffer (pH 6), to which was added 10 mM L-cysteine and 2 μM DCG-04; 0.4 mM E-64 was held in another tube as the control. The samples were incubated at room temperature for 5 h. Then proteins were precipitated by adding 1 mL of ice-cold acetone and collected by centrifugation (for 1 min, at 10 000 × *g*). Proteins were washed twice with 70% acetone and dissolved in 30 μL of TBS, boiled in 30 μL of SDS sample buffer, and finally separated on 10% SDS gels.

Protein stability tests

EGFP-His, GST-OsRACK1A, and GST-OsSNAP32 were extracted from *E. coli*. The proteins of the same quality were mixed and treated at room temperature for 30 min and 60 min. The fusion protein was detected by the corresponding antibody.

Transcriptome analysis

Total RNA was extracted from *ocp-ko1* and TP309 for transcriptome sequencing. The Volcano plot and KEGG enrichment analyses were completed using Majorbio (https://cloud.majorbio.com/). Heatmaps were generated using TBtools.

Total RNA isolation and qRT-PCR analysis

Total RNA was extracted using the TRIzol reagent (Invitrogen, Waltham, MA, USA) by following the manufacturer's protocol, after which cDNA was synthesized using the ReverTra Ace 1 qPCR RT Master Mix with gDNA Remover (TOYOBO, Osaka, Japan). The qRT-PCR was performed using 2 × T5 Fast qPCR Mix (SYBR Green I, Tsingke, Beijing, China) according to the manufacturer's instructions. The rice *OsActin* gene served as an internal control for the data normalization in the formal analysis. The results are presented as the mean ± SD in triplicate. Bar graphs were generated using GraphPad Prism 9. The specific primers used are listed in Supplementary Table 2.

Results

Structure and evolution analysis of OCP and its orthologous proteins in various plant species

The OCP contains four main domains, namely signal peptide, inhibitor I29, pept_C1, and GRAN (Figure 1A). The signal peptide started at position 1 aa (amino acid) and ended at position 21 aa. Inhibitor I29 was the cathepsin propeptide inhibitor domain (40–100 aa). Pept_C1 was the enzyme active domain, belonging to the papain family of cysteine proteases (129–344 aa). GRAN, granulin, is probably released by post-translational proteolytic processing carried out at 367 to 424 aa. Based on the consistency of amino acid sequence with OCP of at least 50% and all above four domains being present, we selected 28 amino acid sequences of five species to build a phylogenetic tree, as shown in Figure 1B. There were three orthologous proteins of OCP in rice, of which OsCP1 was reported to affect pollen development, seed germination, and plant height (Lee et al., 2004; Li et al., 2011). In *A. thaliana*, there were three proteins, including RD21B and RD21A, whose sequence alignment consistency with OCP was 68.13% and 67.59%, respectively (Figure 1C). Of three orthologous proteins in tomato, only CYP1 has been studied (Kaschani et al., 2010;

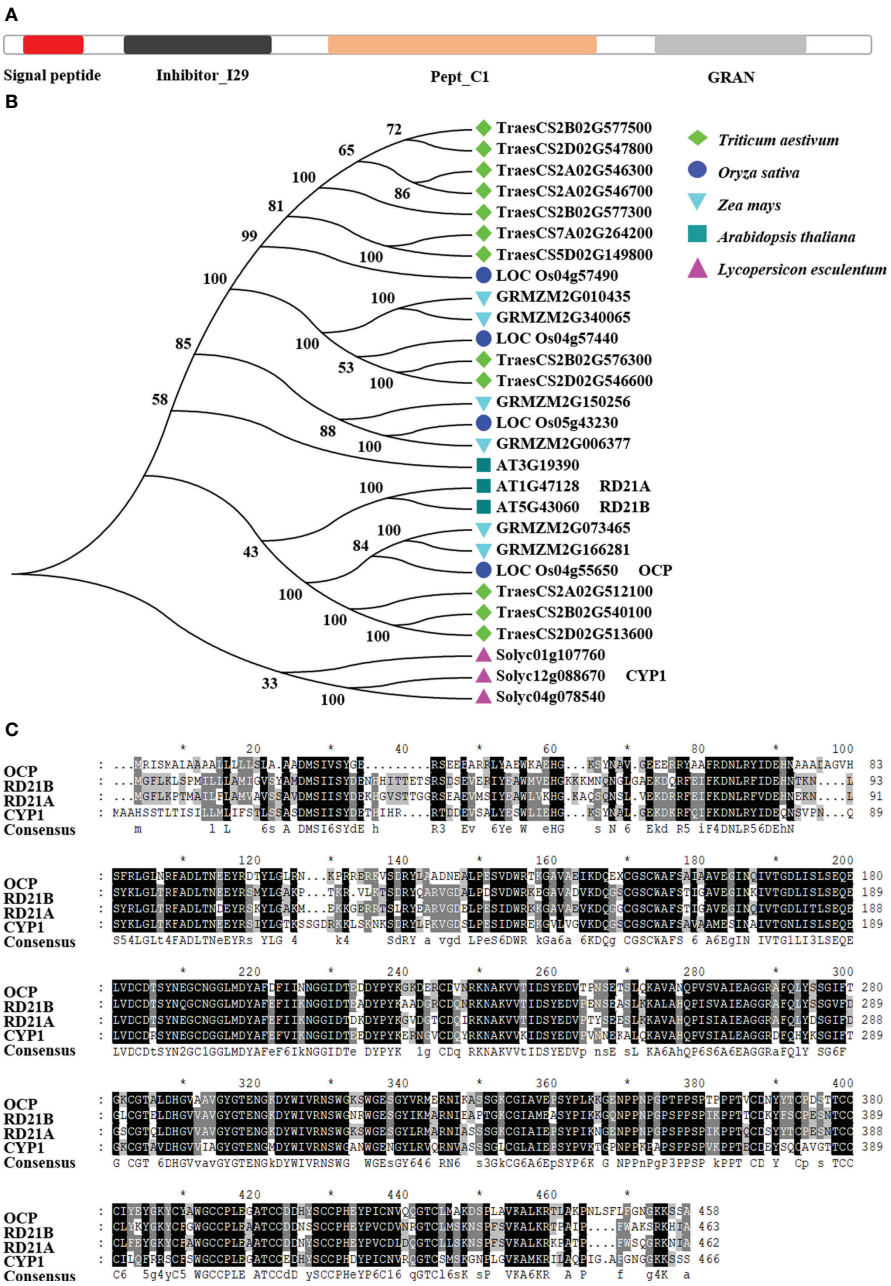


FIGURE 1
Evolutionary analysis of OCP and its orthologous proteins. **(A)** Structures of OCP. **(B)** Phylogenetic tree of OCP and its orthologous proteins in five plant species. Based on a shared amino acid sequence identity greater than 50% with OCP, 27 amino acid sequences were collected from NCBI. The neighbor-joining tree was built using the Muscle program, with 1000 bootstrap replicates, in MEGA7. Protein structure was analyzed by SMART. All the proteins contained the same four domains, signal peptide, inhibitor I29, pept_C1, and GRAN. **(C)** Comparison of the amino acid sequences of OCP with RD21B, RD21A, and CYP1. The black coloring shows the same amino acids present in the four proteins. The * indicates the amino acid position number, 10, 30, 50.....

Misas-Villamil et al., 2016), and its sequence alignment consistency with OCP was 62.5%. Six and 12 proteins were found in maize and wheat, respectively. Hence, we concluded OCP is highly conserved in plants.

Knockout of *OCP* enhances rice blast resistance

In order to explore the specific effects of *OCP* on rice growth and development, we constructed *OCP* knockout vectors with the sgRNA located at the first, second, and fourth exon, these distributed in the inhibitor I29, pept_C1, and GRAN domains, respectively, by using the CRISPR/Cas9 genome editing approach. The *OCP* overexpression vector was driven by the 35S promoter. These vectors were transformed into TP309 calli *via Agrobacterium*-mediated transformation. Through screening, 20, 11, and 5 edited plants of T0 progeny were obtained for three editing sites, respectively. Next, we found that the *ocp-ko* plants were mainly characterized by insertion or deletion of one or more bases. Most editing forms consisted of a single base insertion, and the insertion sites were unified. These mutant types led to an open-reading frame code shift for *OCP* and the premature termination of its translation (Figure 2A). To identify the overexpressing plants, we conducted quantitative real-time PCR (qRT-PCR) assays, by extracting total RNA from the T1 leaf. These results confirmed there were differences in expression levels among individual lines, with the highest level found for *OCP-OE#10* (~27-fold change) and the lowest level for *OCP-OE#1* (~1.8-fold change). Five independent transgenic lines had an expression level of *OCP* that was 15 times greater than that of wild-type plants (Figure 2B). Finally, *ocp-ko1*, *ocp-ko2*, *ocp-ko3*, *OCP-OE#7*, and *OCP-OE#10* lines were chosen for further study.

Cysteine proteases participate in plant immune responses (Shindo et al., 2012; Ormancey et al., 2019). To clarify the response of *OCP* to pathogens, we inoculated the wild-type TP309, *ocp-ko1*, *ocp-ko2*, *ocp-ko3*, and *OCP-OE#10* with six *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) isolates. These results revealed no significant difference in disease among these tested

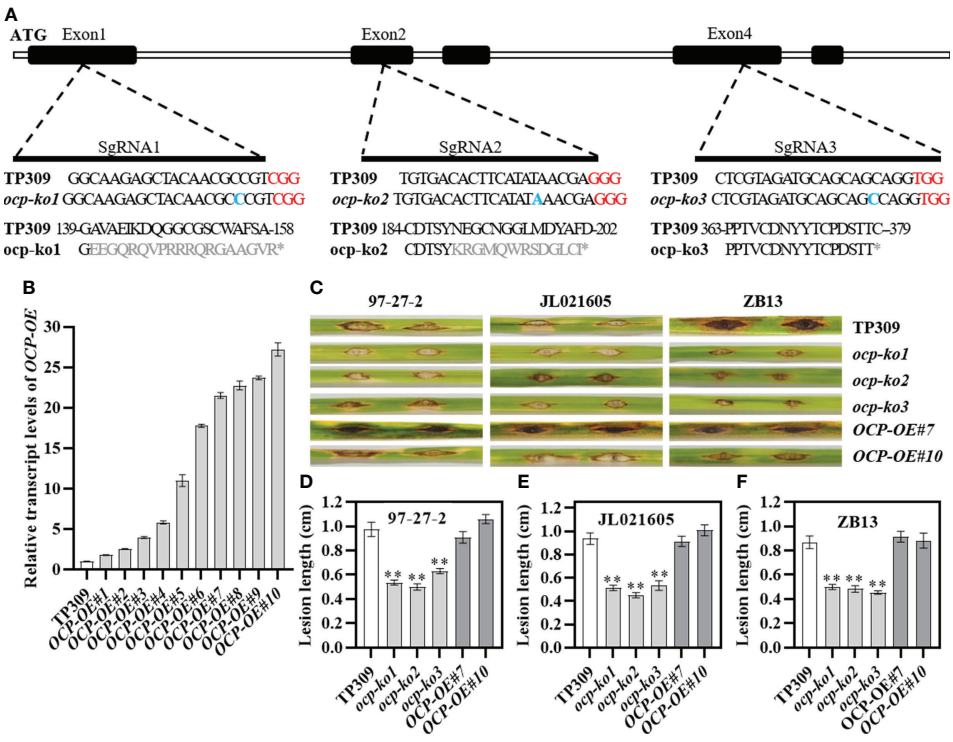


FIGURE 2
Mutant types and blast resistance identification of OCP. **(A)** Three kinds of allelic variations induced by CRISPR/Cas9 in different regions. CDS are shown in the black boxes. Untranslated regions (UTR) and introns correspond to the white sections. The sgRNA1, sgRNA2, and sgRNA3 were located in Exon 1, Exon 2, and Exon 4, respectively, corresponding to three domains of OCP, inhibitor I29, pept_C1, and GRAN. Red bases indicate the protospacer adjacent motif (PAM) recognition sites. Blue bases show the insert location. In gray are the altered amino acid residues due to mutation. The * indicates the terminate codon. **(B)** Relative transcript levels of 10 overexpression lines (*OCP-OE#1* to *OCP-OE#10*) in T1 of *OCP*. **(C)** Punch inoculation of wild-type rice TP309 and *OCP* mutant lines with the blast isolate 97-27-2, JL021605, and ZB13. This experiment was repeated twice. **(D–F)** Lesion lengths (mean ± SEM, n ≥ 8) of the tested lines according to the results in **(C)**. Asterisks indicate statistical significance compared with TP309 (** $P \leq 0.01$, *t* test).

plants (Supplementary Figure 1). Thus *OCP* was not involved in the resistance to bacterial blight. Then we inoculated these lines with *M. oryzae* isolates (97-27-2, JL021605, and ZB13) to assess their resistance to blast. Compared with TP309, the *ocp-ko* mutants exhibited a significantly shorter lesion length (Figure 2C). For all inoculated *M. oryzae* isolates, there was no significant size difference between the lesions of TP309 and *OCP-OE lines*, but those of *ocp-ko* lines were significantly reduced. For example, *ocp-ko1* plants had 0.54 ± 0.02 cm, 0.51 ± 0.02 cm, and 0.50 ± 0.02 cm lesions, corresponding to the three isolates, while TP309 had 0.98 ± 0.06 cm, 0.94 ± 0.05 cm, and 0.87 ± 0.05 cm, respectively (Figures 2D–F).

Then we validated the resistance to blast of the tested lines (TP309, *ocp-ko1*, and *OCP-OE#10*), by conducting spray inoculation assays. The lesion numbers of *ocp-ko1* were dramatically reduced with isolates 97-27-2, JL021605, and ZB13, whereas they were much more abundant for TP309 and *OCP-OE#10* (Supplementary Figure 2). In addition, we inoculated TP309, *ocp-ko1*, and *OCP-OE#10* with JL021605 in the field, finding that *ocp-ko1* plants featured a shorter lesion length when compared to TP309 (Supplementary Figure 3). Altogether, these results show that the *OCP* knockout enhances the resistance to blast but not bacterial blight in rice.

Expression pattern and protease activity analysis of OCP

The temporal and spatial expression pattern of the *OCP* gene was investigated in different tissues of TP309 by qRT-PCR. The gene was expressed in rice various tissues examined, albeit at a higher level in the seedling and at lower levels in the stem, panicle, and axillary bud, and at intermediate levels in the root, node, leaf, and sheath (Figure 3A). To further confirm the expression levels of *OCP*, transgenic plants were generated with the expression of a β-glucuronidase (GUS) driven by the

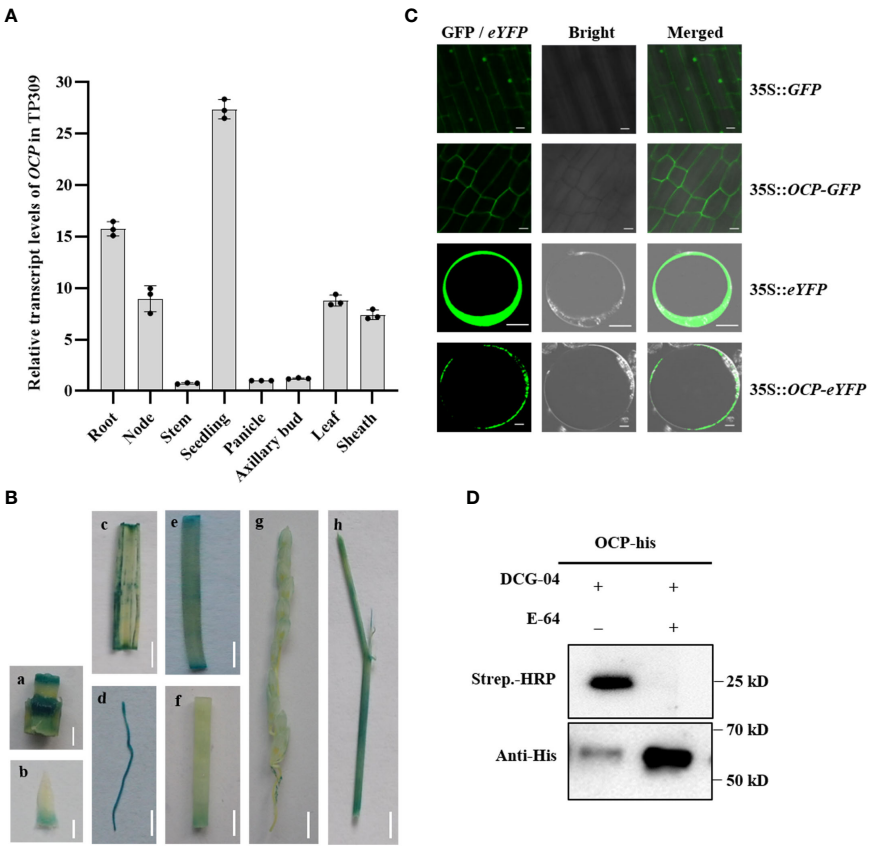


FIGURE 3
Expression pattern and cysteine protease activity analysis of OCP. **(A)** *OCP* RNA expression pattern obtained by qRT-PCR in rice TP309 (mean ± SEM, n = 3). **(B)** GUS activity of *OCP* in different rice tissues. a, node; b, axillary bud; c, leaf; d, root; e, sheath; f, stem; g, panicle; h, leaf and sheath of seedlings. Scale bars: a = 5 mm; b = 2 mm; c–h = 10 mm. **(C)** Subcellular localization of OCP in rice root tips and rice protoplasts. Scale bar = 10 μm. **(D)** Protease activity profiling of OCP *in vitro*. E-64 was an effective inhibitor of cysteine proteases, and DCG-04 was a biotinylated derivative of E-64. The biotinylated protease could be detected on a protein gel blot using a conjugate of streptavidin with HRP (strep.-HRP). OCP-His was detected with Anti-His.

promoter of *OCP* in TP309. Strong GUS activity was detected in the root, leaf, and sheath of seedling (Figure 3B), supporting well the qRT-PCR results.

To determine the subcellular localization of OCP, we fused the *OCP* coding region with the green fluorescent protein (GFP) and the enhanced yellow fluorescent protein (eYFP) driven by the cauliflower mosaic virus 35S promoter at the C terminus and expressed the fusion proteins in rice plants and protoplasts. Laser confocal microscopy revealed that the signal of the GFP-tagged OCP protein was excluded from the nucleus; meanwhile, the GFP signal alone was expressed in both the nucleus and cytoplasm in the roots of transgenic plants. In rice protoplasts, the green fluorescence emitted by the fusion protein was detected exclusively in the cytoplasm (Figure 3C). Therefore, OCP is located in the cytoplasm.

As the orthologous gene of *RD21*, *OCP* encodes a cysteine protease; hence, we further profiled protease activity by Western blotting *in vitro*. The method used followed one describe before (van der Hoorn et al., 2004). DCG-04 is a biotinylated derivative of the E-64 cysteine protease inhibitor. Active cysteine protease cleaves protein substrates through a covalent intermediate state, in that the biotinylation of active proteases by DCG-04 occurs because the cleavage mechanism is locked in a covalent intermediate state. Biotinylated proteases can be detected on SDS-PAGE gel using streptavidin conjugated to HRP. Firstly, the protein of OCP was expressed in *E. coli*, and we found that the protein mainly existed in the supernatant and only a small part occurred in the sediment (Supplementary Figure 4), which facilitated collection of the target protein. Then we treated the same content protein with DCG-04 and E-64 (as a control). These results showed that it was detected with streptavidin-HRP when OCP was treated with DCG-04 alone, and excess E-64 inhibited the binding of DCG-04 and OCP (Figure 3D). In addition, more protein remained in the sample with E-64 detected with anti-His. Taken together, our results demonstrated OCP has cysteine protease activity, and that E-64 can effectively delay the degradation of OCP.

OCP physically interacts with OsRACK1A and OsSNAP32

To identify the interaction partners of OCP, we performed a yeast two-hybrid (Y2H) screen using a cDNA library of rice. The coding region of *OCP* was cloned into the bait vector (BD-OCP), and no autoactivation and toxicity of OCP were proven (Supplementary Figure 5). Then, BD-OCP was co-transformed with a cDNA library into yeast cells. Forty-four clones were isolated from the quadruple dropout media (-L-T-H-A) (Supplementary Table 3), and their sequences were analyzed *via* amplifying and sequencing. Two possible interacting proteins, OsRACK1A and OsSNAP32, participating in rice blast resistance (Nakashima et al., 2008; Luo et al., 2016), were thus obtained. Next, the CDS of *OsRACK1A* and *OsSNAP32* were fused to the GAL4 activation domain (AD-OsRACK1A and AD-OsSNAP32), and each co-transformed with BD-OCP into yeast cells. These cells grew well on the quadruple dropout media containing BD-OCP and AD-OsRACK1A or AD-OsSNAP32, while those transformed with the corresponding control did not (Figure 4A). To clarify the domains interacting with OsRACK1A or OsSNAP32, we divided the protein of OCP into five segments (Figure 4A). Interestingly, BD-OCP-3 and BD-OCP-5 interacted with the empty vector AD, and 20 mM 3-AT was then used to verify the actual interaction (Supplementary Figure 6). Lastly, we found that BD-OCP-2 and BD-OCP-4 grew well on the screening medium with AD-OsRACK1A or AD-OsSNAP32, which demonstrated that pept_C1 was effective for their interaction.

To further verify the interaction between OCP and OsRACK1A or OsSNAP32, we conducted a bimolecular fluorescence complementation (BiFC) assay to produce the fusion proteins cYFP-OCP and nYFP-OsRACK1A or nYFP-OsSNAP32 in *N. benthamiana* leaves *via Agrobacterium*-mediated transformation. These results showed that OCP interacted with OsRACK1A or OsSNAP32, with the associated fluorescence detected in the cytoplasm predominantly (Figure 4B). Meanwhile, co-immunoprecipitation (Co-IP) assays to test the interaction *in vivo* were carried out, co-expressing the fusion proteins OCP-Myc and OsRACK1A-mCherry, OsSNAP32-mCherry, or mCherry alone in the rice protoplasts. A band was detected for OCP-Myc and OsRACK1A-mCherry or OsSNAP32-mCherry in their IP sample, but no band was discernible in the OCP-Myc and mCherry IP samples when using an anti-mCherry antibody (Figure 4C). Together, these results suggested that OCP interacts with OsRACK1A or OsSNAP32 physically, *in vitro* and *in vivo*.

OsRACK1A and OsSNAP32 cannot form a complex *via* OCP

To clarify the co-expression of OCP and its interacting proteins, subcellular localization assays were carried out. The results indicated that OsRACK1A was located in the cytoplasm, agreeing with a previous study (Zhang et al., 2018), whereas the OsSNAP32 was located in the cell membrane as well as the cytoplasm (Figure 5A). Subcellular co-localization showed that both OCP and OsRACK1A or OsSNAP32 are expressed in the cytoplasm (Figure 5B), which makes their interaction possible. Then the CDS of *OsSNAP32* was joined to pGBKT7 (BD-OsSNAP32) and co-transformed into yeast cells with AD-OsRACK1A *via* the Y2H system. These results indicated a non-interaction between OsRACK1A and OsSNAP32. Yet OCP interacted separately with both OsRACK1A and OsSNAP32, raising the question, could they form a complex? To answer this, yeast three-hybrid (Y3H) assays were conducted.

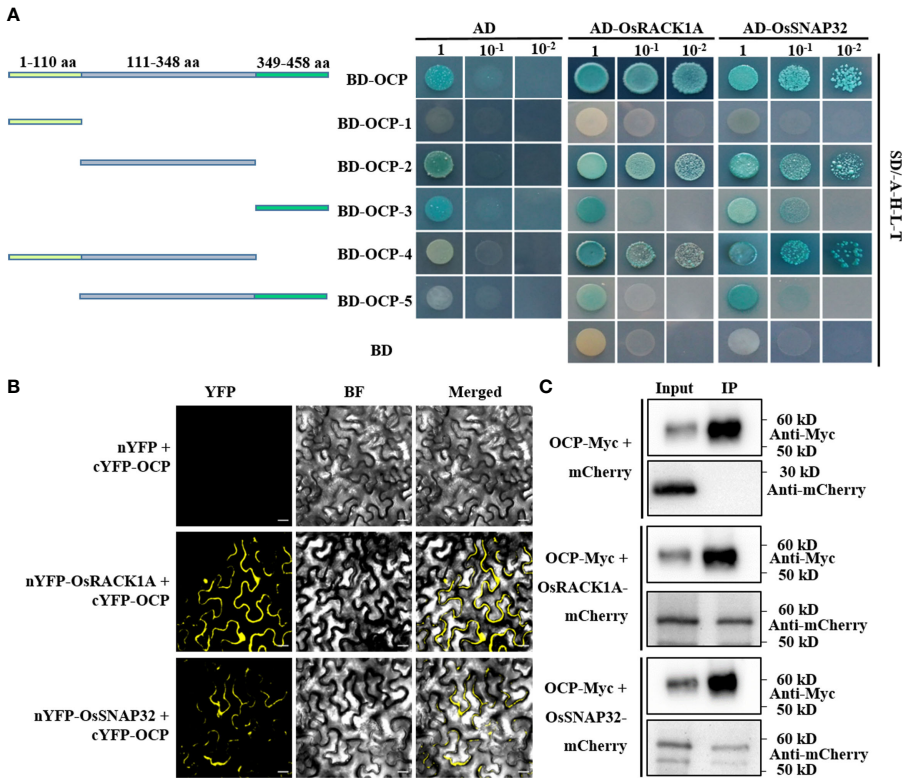


FIGURE 4
OCP physically interacts with OsRACK1A or OsSNAP32. **(A)** OCP and its truncations interact with OsRACK1A or OsSNAP32 in Y2H. AD, pGADT7; BD, pGBKT7; BD-OCP, full length of OCP; BD-OCP-1, 1–110 amino acid (aa) of OCP; BD-OCP-2, 111–348 aa; BD-OCP-3, 349–458 aa; BD-OCP-4, 1–348 aa; BD-OCP-5, 111–458 aa. 20 mM 3-amino-1,2,4-triazole (3-AT) was added to BD-OCP-3 and BD-OCP-5. **(B)** BiFC assay showing the interaction of OCP with OsRACK1A or OsSNAP32 in tobacco leaf epidermal cells. nYFP+cYFP-OCP was the negative control. Bar = 20 μm. **(C)** Co-immunoprecipitation assays to verify the interaction of OCP with OsRACK1A or OsSNAP32 in rice protoplasts. OCP-Myc +mCherry was the negative control. OsRACK1A-mCherry and OsSNAP32-mCherry were detected with Anti-mCherry.

OsSNAP32 and *OCP* were respectively inserted into the multiple cloning site (MCS) 1 (pBridge-OsSNAP32) and MCS2 (pBridge-OCP) of the Y3H vector pBridge, and also simultaneously (pBridge-OsSNAP32-OCP). We transformed the combinations with AD-OsRACK1A to yeast cells, which did not grow well on the SD/-Ade-His-Leu-Met-Trp medium. The results indicated that OsRACK1A and OsSNAP32 could not interact with each other when OCP acted as a bridge in the Y3H system (Figure 5C). Analyzing the transcript levels by qRT-PCR, we found that *OsSNAP32* mRNA accumulated to a significantly higher level (~10-fold) in *ocp-ko1* than in TP309, whereas the changed transcript levels of *OsRACK1A* was not significant (Figure 5D). We then examined the effect of OCP on the stability of OsRACK1A and OsSNAP32, finding it similar between the control and corresponding treatment, suggesting that OCP did not affect the stability of OsRACK1A and OsSNAP32 (Figure 5E). These results collectively show that OCP cannot form a complex with OsRACK1A and OsSNAP32, nor does it affect their stability, but it can suppress the expression of *OsSNAP32*.

Analysis of differentially expressed genes related to the immune response between *ocp-ko1* and TP309

To confirm *OCP*'s participation in the immune response of rice, we conducted RNA sequencing assays. The sequencing and mapping data are summarized in Supplementary Table 4. A total of 6956 DEGs were identified in the transcriptional profiles (having a fold-change ≥ 2 and p-adjusted value < 0.05), of which 4169 DEGs were up-regulated and 2787 DEGs were down-regulated in *ocp-ko1* vis-à-vis TP309 (Supplementary Figure 7). KEGG enrichment analysis revealed that the up-regulated DEGs were chiefly involved in phenylpropanoid biosynthesis, plant–pathogen interaction, flavonoid biosynthesis, MAPK signaling pathway-plant, and so on (Supplementary Figure 8). The down-regulated DEGs were enriched in terms of ribosome, DNA replication, plant hormone biosynthesis, signal transduction, and so on (Supplementary Figure 9). We found 71 DEGs responsive to plant-pathogen interaction and 47 DEGs functioning in the MAPK signaling pathway (Supplementary Figure 8), which are

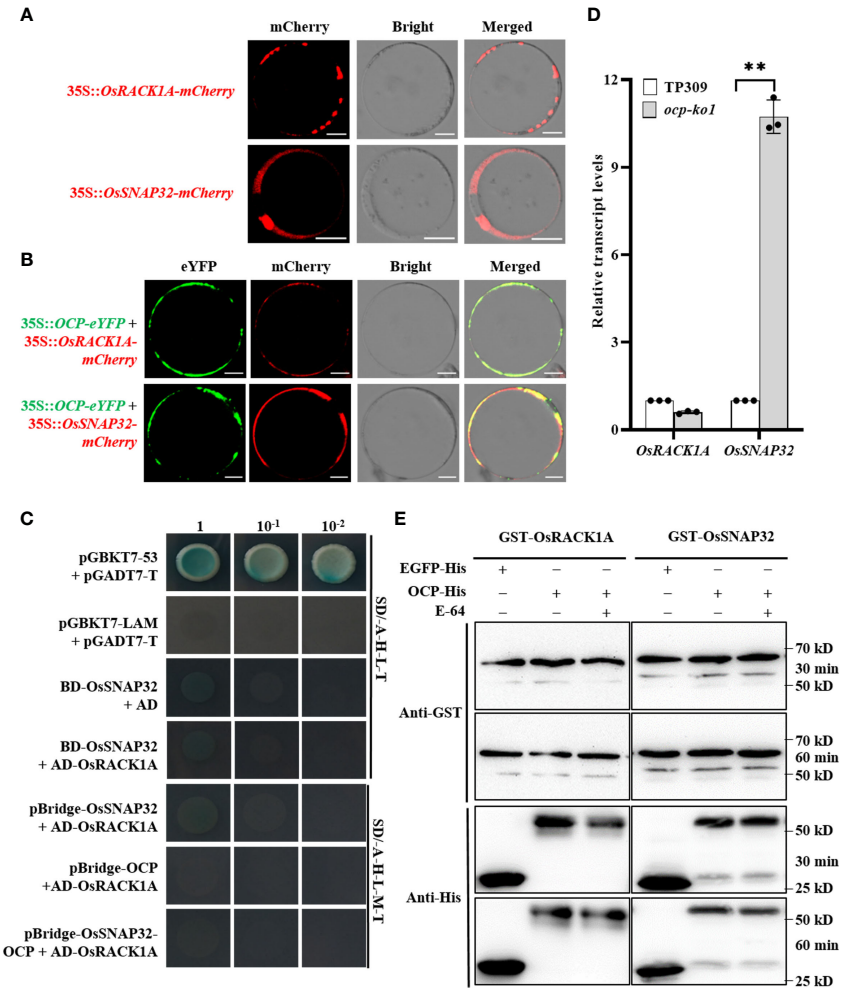


FIGURE 5
OCP is unable to form a complex with OsRACK1A or OsSNAP32, or affect the stability of either. **(A)** Subcellular localization of OsRACK1A and OsSNAP32 in rice protoplasts. Scale bar = 10 μm. **(B)** Subcellular co-localization of OCP and OsRACK1A or OsSNAP32 in rice protoplasts. Scale bar = 10 μm. **(C)** Yeast two-hybrid assays showing the non-interaction of OsRACK1A and OsSNAP32; pGBKT7-53 and pGADT7-T were used as positive controls, while pGBKT7-Lam and pGADT7-T were used as negative controls. Yeast three-hybrid assays showed that OCP, OsRACK1A, and OsSNAP32 could not form a complex. **(D)** Relative transcript levels of OsRACK1A and OsSNAP32 in *ocp-ko1* and TP309 were assessed by qRT-PCR (mean ± SD, n = 3). This experiment was repeated twice, ** $P \leq 0.01$, *t* test. **(E)** Detection of the influence of OCP on the stability of OsRACK1A and OsSNAP32. EGFP-His served as the negative control. GST-OsRACK1A and GST-OsSNAP32 were detected with Anti-GST.

related to plant tolerance of biotic stress (Xiong and Yang, 2003; Chang et al., 2022). We divided these DEGs into several categories, namely those related to disease, transcription factor, calmodulin, proteinase, hormone, and protein kinase (Figures 6A–F).

Phytohormones play pivotal roles in plant defense responses. In this study, we identified some genes related to jasmonic acid (JA), ethylene (ET), auxin, and abscisic acid (ABA) (Figure 6D). *OsAOS2* (*allene oxide synthase 2*), *OsAOS3* (*allene oxide synthase 3*), *OsOPR7* (*OPDA reductase 7*), and *JIOsPR10* (*jasmonate inducible pathogenesis-related class 10*), all of which figure prominently in the biosynthesis of JA, were up-regulated dramatically, which could be viewed as resistance response to pathogens (Wang et al., 2021). *SIT1* (*salt intolerance 1*), *OsERF101* (*ethylene response factor 101*), *OsERF62* (*ethylene response factor 62*), *Sub1C* (*submergence1C*), *OsETR2* (*ethylene response 2*), *OsACS2* (*ACC synthase 2*), and *OsACO5* (*ACC oxidase 5*), which participate in ethylene (ET) biosynthesis and response (Helliwell et al., 2016), were increased significantly at the transcript level. Auxin response factors (*OsARF2*, *auxin response factor 2*; *OsARF11 auxin response factor 11*; etc.) and auxin efflux transporters (*OsPIN1a*, *pin-formed1a*; *OsPIN1d*, *pin-formed1d*; etc.) were strongly down-regulated; in stark contrast, the auxin influx carrier *OsAUX1* (*auxin transporter 1*) and IAA synthetase gene *OsGH3-2* (*gretchen hagen 3-2*) were both up-regulated. Inducing the expression of *OsGH3-2* is known lower the auxin

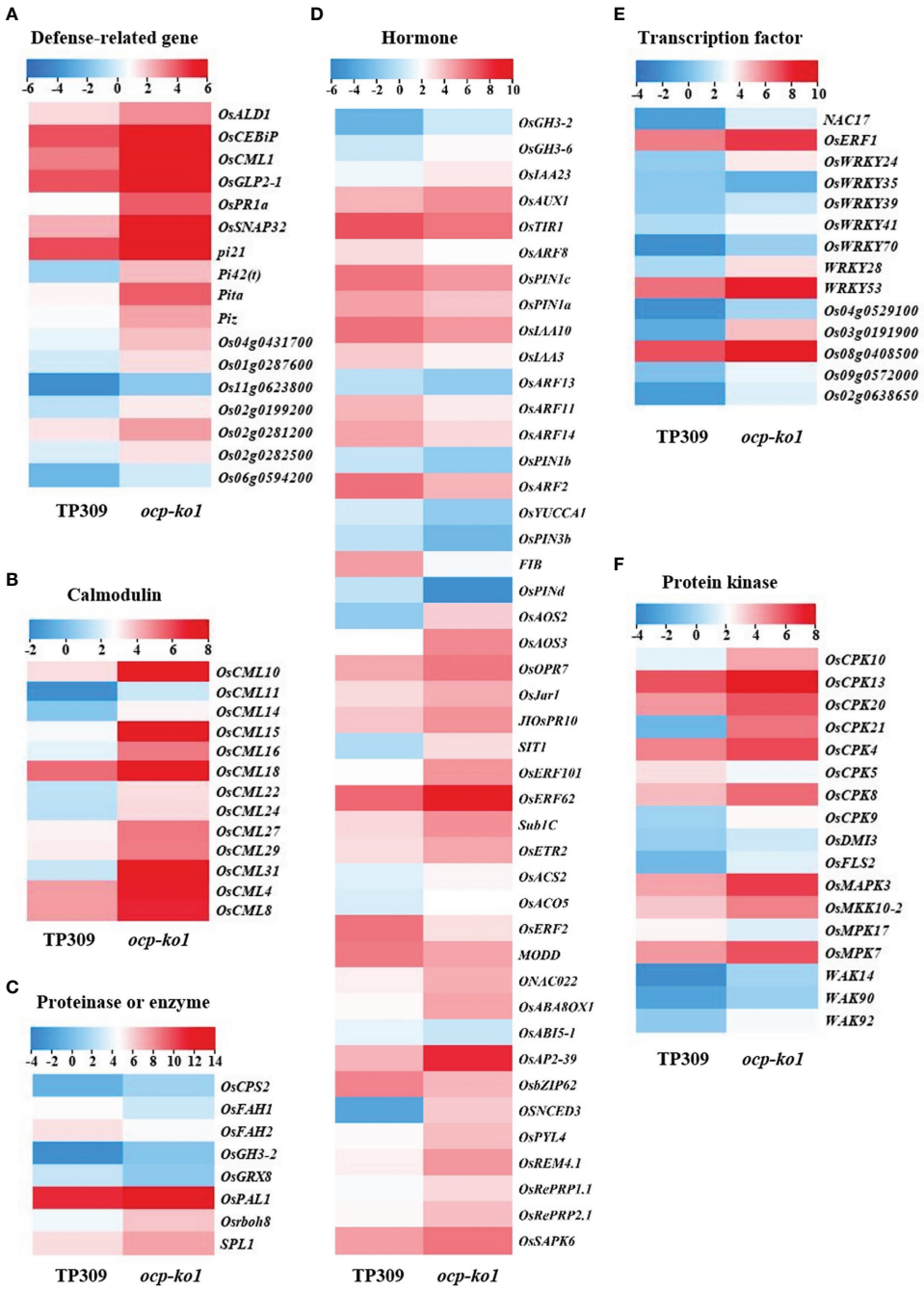


FIGURE 6
Detailed analysis of differentially expressed genes (DEGs) related to disease resistance in rice between *ocp-ko1* and TP309. **(A–F)**, heat map of DEGs. **(A)** Defense-related gene. **(B)** Calmodulin. **(C)** Proteinase or enzyme. **(D)** Hormone. **(E)** Transcription factor. **(F)** Protein kinase. The color legend insets indicate the $\log_2$ (FPKM) value.

content, leading to an auxin deletion phenotype and enhanced resistance to rice blast (Fu et al., 2011). Concerning ABA (Spence et al., 2015), some genes related to it were up-regulated; for instance, *ONAC022* (*NAC domain-containing protein 022*), *OsABA8OX1* (*ABA 8'-hydroxylase 1*), *OsAP2-39* (*APETALA-2-like transcription factor*), *OsNCED3* (*9-Cis-epoxycarotenoid dioxygenase 3*), and *OsREM4.1* (*remorin group 4 member 1*) were found up-regulated, while others were down-regulated, namely *MODD* (*mediator of OsbZIP46 deactivation and degradation*), *OsbZIP62* (*bZIP transcription factor 62*), and

OsABI5-1 (*abscisic acid insensitive 5*). Some genes can respond to more than one hormone at once; for example, *OsERF2* (*ethylene responsive factor 2*) (Xiao et al., 2016), an ethylene response factor, was evidently required for the control of the ET- and ABA-responses with its transcripts declining in *ocp-ko1*. The transcript levels of *pi21*, *Pi42(t)* (*Magnaporthe grisea resistance-42(t)*), *Pita* and *Piz* (*Magnaporthe grisea resistance-z*), the executive genes of rice blast resistance, were substantially increased (Figure 6A). Most transcription factors (*NAC17*, *NAC domain-containing protein 17*; *OsWRKY53*, *WRKY gene 53*; etc.), proteinases (*OsPAL1*; *SPL1*, *sphingosine-1-phosphate lyase 1*; etc.), and protein kinases (*OsFLS2*, *flagellin sensitive 2*; *OsMAPK3*, *mitogen-activated protein kinase*; etc.) were up-regulated in *ocp-ko1*, when compared with TP309 (Figures 6C, E, F). Intriguingly, the transcript levels of many calmodulin genes (*OsCML4*, *calmodulin-like 4*; *OsCML8*; *OsCML10*; etc.) and calcium-dependent protein kinase genes (*OsCPK4*, *calcium-dependent protein kinase 4*; *OsCPK5*; *OsCPK8*; etc.) were drastically increased (Figures 6B, F). These results suggested that *OCP* could regulate blast resistance by influencing the expression of defense-related genes.

To verify the RNA sequencing results of *ocp-ko1* and TP309, we designed primers for qRT-PCR to compare the turnover rate of genes' mRNA related to blast (Figures 7A–C). All tested genes presented a difference and reached a significant level. The response of many genes response to auxin changed significantly (Figure 7A). The blast resistance genes *pi21* and *Pita* were up-regulated more than twice and eight-fold, respectively. The mRNA of *OsPAL1* and *OsMAPK3*, which enhance resistance to blast when overexpressed, accumulated in *ocp-ko1*. Both *OsPR1a* (*pathogenesis-related 1a*) and *OsFLS2*, which trigger an immune response in response to pathogen inoculation, were up-regulated about 13 and 8 times, respectively. Moreover, *OsCPK4*, *OsCPK10*, *OsCPK20*, and *OsCPK21*, all of which encode calcium-dependent protein kinases, were expressed more in *ocp-ko1*, and these genes' up-regulation conferred enhanced immunity upon pathogen infection (Figure 7B). Finally, *OsABA8ox1*, *OsAP2-39*, and

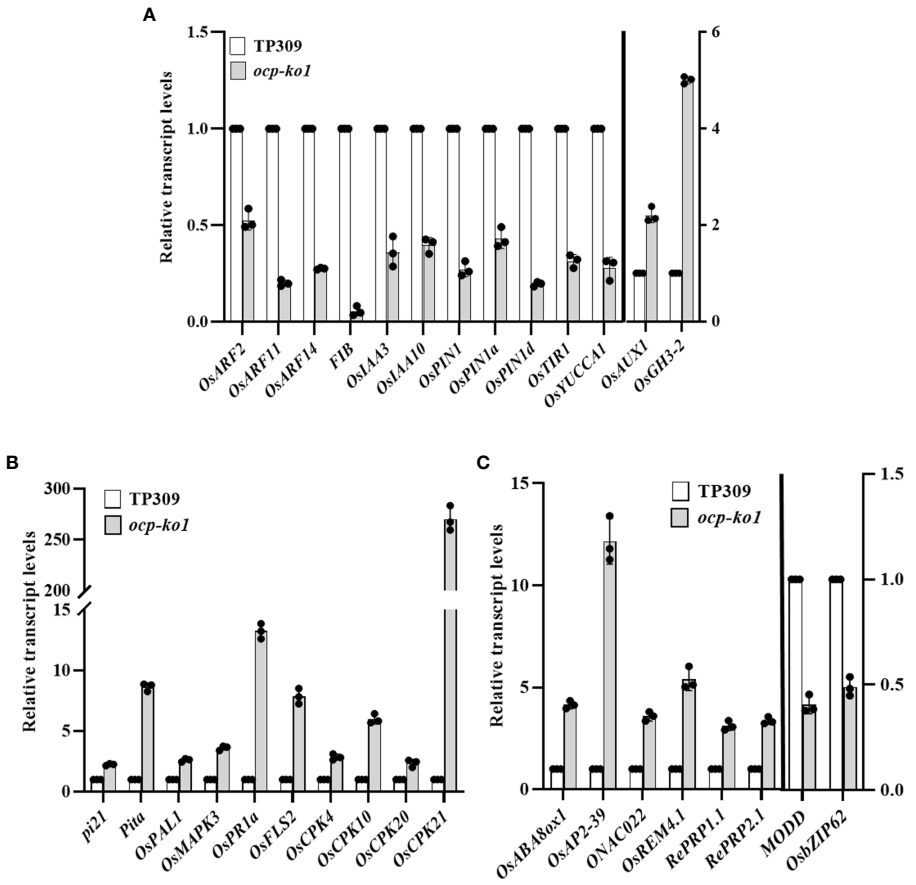


FIGURE 7
Relative transcript levels of (DEGs) related to disease resistance in rice between *ocp-ko1* and TP309 validated by qRT-PCR. **(A)** Relative transcript levels of auxin-related genes. **(B)** Relative transcript levels of blast-related genes. **(C)** Relative transcript levels of ABA-related genes (mean ± SD, n = 3). This experiment was repeated twice.

OsREM4.1 were up-regulated while both *MODD* and *OsbZIP62* were down-regulated (Figure 7C). The above qRT-PCR results were consistent with those of the transcriptome analysis, which suggested OCP may regulate the blast resistance by affecting various pathways in rice.

Discussion

Our study focused on the PLCP gene *OCP*, which was involved in regulating the rice response to the blast-causing fungus *M. oryzae*. We selected three loci distributed in different domains of *OCP* for editing, and the pathogen inoculation results showed that the *OCP* knockout lines presented resistance to blast isolates 97-27-2, JL021605, and ZB13, whereas the *OCP* overexpression lines and TP309 did not and responded similarly (Figures 2C-F). OCP possessed cysteine protease activity, and it interacted with OsRACK1A and OsSNAP32 physically *in vitro* and *in vivo*. Accordingly, it is worthwhile to study the genetic relationship between OCP and OsRACK1A or OsSNAP32. Further, *OCP* influenced the expression of some genes related to blast resistance.

OCP has pleiotropic effects on rice development and resistance

In the study, *ocp-ko* lines showed increased resistance to *M. oryzae*, but not to *Xoo*, whereas *OCP-OE* was susceptible to both phytopathogens. Therefore, OCP negatively influences plant defense against fungal pathogens. In Arabidopsis, the *rd21* null mutants were more susceptible to the necrotrophic fungal pathogen (Shindo et al., 2012); conversely, null *XCP1* or *XCP2* mutants displayed enhanced resistance (Zhang et al., 2014; Perez-Lopez et al., 2021). Therefore, genes harboring the same functional domains do not necessarily function in the same way.

The homologous gene *RD21* of *OCP* in *A. thaliana* responds to biotic and abiotic stressors (Kikuchi et al., 2008; Rustgi et al., 2017), and *OsCP1* in rice is known to affect pollen development (Lee et al., 2004). Here, we found that knocking out *OCP* led to shorter plant height and lower fertility than TP309. The plant height of *ocp-ko1* was significantly reduced to 103.9 ± 0.89 cm, while TP309 was taller, at 119.5 ± 0.79 cm (Supplementary Figure 10). Compared with TP309, the panicle length of *ocp-ko1* was shorter and it had more empty grains, mainly due to the abnormal pollen development of *ocp-ko1* that resulted in its significantly decreased seed setting rate (Supplementary Figure 11). Therefore, OCP is a pleiotropic gene, which modulated blast resistance yet also influenced plant height and pollen development. In further research, we will aim to identify the mechanism by which OCP regulates plant height and fertility.

OCP negatively regulates blast resistance *via* multiple pathways

PLCPs play key roles in the growth and development of plants, as well as the immune responses to pathogens (Avci et al., 2008; Liu et al., 2018). Yet, further investigation is required to uncover the protease substrates and functional pathways (Demir et al., 2018). We found OCP located in the cytoplasm interspersed with OsRACK1A and OsSNAP32 (Figure 5A); not surprisingly, perhaps, these two proteins physically interacted with OCP (Figure 4). Nevertheless, the yeast hybrid results showed that OsRACK1A and OsSNAP32 did not interact with each other, and they could not form a ternary complex with OCP (Figure 5C). *In vitro*, we proved that OCP possessed cysteine protease activity (Figure 3D). Expression analysis found that many genes related to disease resistance, such as *pi21*, *Pi42(t)*, *OsSNAP32*, and *OsMAPK3* (among others), were up-regulated in *ocp-ko1*. Therefore, it is quite plausible that *OCP* negatively regulates blast resistance by influencing the expression of *OsSNAP32* and other disease-resistance genes.

During immune responses, phytohormones act as signals to trigger and mediate defense responses in plants against enemies (Yang et al., 2013). JA functions critically in the basal defense of rice, especially against necrotrophic pathogens (Browse, 2009). ET regulates disease resistance positively or negatively depending on the different pathogens and local environmental conditions (Broekaert et al., 2006; van Loon et al., 2006), and exogenous application of an ET generator could increase rice blast disease resistance (Singh et al., 2004). Auxin, being a widespread important hormone in plants, is involved in almost all developmental processes. The accumulation of auxin content in the model plants Arabidopsis and rice leads to their increased susceptibility to disease (Yang et al., 2013). Regarding ABA, its application to rice suppresses resistance to blast (Koga et al., 2004; Yang et al., 2013). In our study, the expression of particular genes known to participate in hormone synthesis and metabolism was changed in *ocp-ko1* plants (Figure 6B, Figures 7A, C). Often, hormones interplay and engage in hormonal crosstalk to defend plants against pathogens (Kazan and Lyons, 2014). In addition to the above, many other DEGs, transcription factors, protein kinases, and so forth, were proven to respond to *M. oryzae*. Importantly, some calmodulin genes were up-regulated significantly; hence, it is possible that *OCP* suppresses blast resistance *via* multiple pathways, wherein calmodulin might play a crucial role. To better understand the molecular mechanisms of OCP-mediated blast resistance in rice, further investigations are needed to clarify the signaling pathways of *OCP* vis-à-vis other factors in the plant immune response.

Conclusion

OCP negatively regulates blast resistance in rice, because all *ocp-ko1*, *ocp-ko2*, and *ocp-ko3* lines have enhanced resistance to *M. oryzae*. OCP is expressed in all rice tissues and located mainly in the cytoplasm, interacting with OsRACK1A and OsSNAP32 *in vivo* and *in vitro*, but they could not form a complex. The transcriptome analysis shows that the expression of many factors responsive to *M. oryzae* are changed in *ocp-ko1* significantly, including the phytohormones JA, ET, auxin, and ABA, which suggests that OCP could affect host resistance to rice blast in multiple ways and plays a fundamental role. Therefore, this study's findings provide the basis for exploring the molecular mechanism of cysteine protease in the disease resistance of rice. Further, in screening many potential interaction proteins of OCP, this work can assist in comprehensively investigating the effects of *OCP* on rice growth and development.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI, PRJNA855166. Sequence data from this article can be found in the GenBank database under the following accession numbers: OCP, LOC_Os04g55650; OsRACK1A, Os01g0686800; OsSNAP32, Os02g0437200.

Author contributions

YL, WZ, and CL conceived the study and designed the experiments. YL, PL, LM, GJ, and QL performed the experiments. YL and CL wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (grant nos. 31900383, 31971911), and China Postdoctoral Science Foundation (grant no. 2019M660853).

Acknowledgments

We thank Dr. Zhuangzhi Zhou and Dr. Minxiang Yu for supplying the *M. oryzae* isolates 97-27-2, JL021605 and ZB13.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1065253/full#supplementary-material

References

Avci, U., Earl Petzold, H., Ismail, I. O., Beers, E. P., and Haigler, C. H. (2008). Cysteine proteases XCP1 and XCP2 aid micro-autolysis within the intact central vacuole during xylogenesis in arabidopsis roots. *Plant J.* 56, 303–315. doi: 10.1111/j.1365-313X.2008.03592.x

Bao, Y. M., Wang, J. F., Huang, J., and Zhang, H. S. (2008). Molecular cloning and characterization of a novel SNAP25-type protein gene *OsSNAP32* in rice (*Oryza sativa* l.). *Mol. Biol. Rep.* 35, 145–152. doi: 10.1007/s11033-007-9064-8

Bart, R., Chern, M., Park, C. J., Bartley, L., and Ronald, P. C. (2006). A novel system for gene silencing using siRNAs in rice leaf and stem-derived protoplasts. *Plant Methods* 2, 13. doi: 10.1186/1746-4811-2-13

Broekaert, W. F., Delaure, S. L., De Bolle, M. F., and Cammue, B. P. (2006). The role of ethylene in host-pathogen interactions. *Annu. Rev. Phytopathol.* 44, 393–416. doi: 10.1146/annurev.phyto.44.070505.143440

Browse, J. (2009). Jasmonate passes muster: a receptor and targets for the defense hormone. *Annu. Rev. Plant Biol.* 60, 183–205. doi: 10.1146/annurev.arplant.043008.092007

Bryan, G. T., Wu, K. S., Farrall, L., Jia, Y. L., Hershey, H. P., McAdams, S. A., et al. (2000). A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. *Plant Cell* 12, 2033–2045. doi: 10.1105/tpc.12.11.2033

Cao, W. L., Yu, Y., Li, M. Y., Luo, J., Wang, R. S., Tang, H. J., et al. (2019). OsSYP121 accumulates at fungal penetration sites and mediates host resistance to rice blast. *Plant Physiol.* 179, 1330–1342. doi: 10.1104/pp.18.01013

Chang, M., Chen, H., Liu, F., and Fu, Z. Q. (2022). PTI and ETI: convergent pathways with diverse elicitors. *Trends Plant Sci.* 27, 113–115. doi: 10.1016/j.tplants.2021.11.013

Chen, Z. X., Zhao, W., Zhu, X. B., Zou, C. D., Yin, J. J., Chern, M. S., et al. (2018). Identification and characterization of rice blast resistance gene *Pid4* by a combination of transcriptomic profiling and genome analysis. *J. Genet. Genomics* 45, 663–672. doi: 10.1016/j.jgg.2018.10.007

Demir, F., Niedermaier, S., Villamor, J. G., and Huesgen, P. F. (2018). Quantitative proteomics in plant protease substrate identification. *New Phytol.* 218, 936–943. doi: 10.1111/nph.14587

Dong, J. F., Zhou, L., Feng, A., Zhang, S., Fu, H., Chen, L., et al. (2021). The *OsOXO2*, *OsOXO3* and *OsOXO4* positively regulate panicle blast resistance in rice. *Rice* 14, 51. doi: 10.1186/s12284-021-00494-9

Fukuoka, S., Saka, N., Koga, H., Ono, K., Shimizu, T., Ebana, K., et al. (2009). Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science* 325, 998–1001. doi: 10.1126/science.1175550

Fu, J., Liu, H., Li, Y., Yu, H., Li, X., Xiao, J., et al. (2011). Manipulating broad-spectrum disease resistance by suppressing pathogen-induced auxin accumulation in rice. *Plant Physiol.* 155, 589–602. doi: 10.1104/pp.110.163774

Grudkowska, M., and Zagdanska, B. (2004). Multifunctional role of plant cysteine proteinases. *Acta Biochim. Pol.* 51, 609–624. doi: 10.18388/abp.2004_3547

Helliwell, E. E., Wang, Q., and Yang, Y. N. (2016). Ethylene biosynthesis and signaling is required for rice immune response and basal resistance against *magnaporthe oryzae* infection. *Mol. Plant Microbe Interact.* 29, 831–843. doi: 10.1094/MPMI-06-16-0121-R

Kaschani, F., Shabab, M., Bozkurt, T., Shindo, T., Schornack, S., Gu, C., et al. (2010). An effector-targeted protease contributes to defense against *Phytophthora infestans* and is under diversifying selection in natural hosts. *Plant Physiol.* 154, 1794–1804. doi: 10.1104/pp.110.158030

Kazan, K., and Lyons, R. (2014). Intervention of phytohormone pathways by pathogen effectors. *Plant Cell* 26, 2285–2309. doi: 10.1105/tpc.114.125419

Kikuchi, Y., Saika, H., Yuasa, K., Nagahama, M., and Tsuji, A. (2008). Isolation and biochemical characterization of two forms of RD21 from cotyledons of daikon radish (*Raphanus sativus*). *J. Biochem.* 144, 789–798. doi: 10.1093/jb/mvn132

Koga, H., Dohi, K., and Mori, M. (2004). Abscisic acid and low temperatures suppress the whole plant-specific resistance reaction of rice plants to the infection of *Magnaporthe grisea*. *Physiol. Mol. Plant P.* 65, 3–9. doi: 10.1016/j.pmpp.2004.11.002

Lee, S., Jung, K. H., An, G. H., and Chung, Y. Y. (2004). Isolation and characterization of a rice cysteine protease gene, *OSCP1*, using T-DNA gene-trap system. *Plant Mol. Biol.* 54, 755–765. doi: 10.1023/B:PLAN.0000040904.15329.29

Li, W., Deng, Y., Ning, Y., He, Z., and Wang, G. L. (2020). Exploiting broad-spectrum disease resistance in crops: from molecular dissection to breeding. *Annu. Rev. Plant Biol.* 71, 575–603. doi: 10.1146/annurev-arplant-010720-022215

Li, X. W., Gao, X. Q., Wei, Y., Deng, L., Ouyang, Y. D., Chen, G. X., et al. (2011). Rice APOPTOSIS INHIBITOR5 coupled with two DEAD-box adenosine 5'-triphosphate-dependent RNA helicases regulates tapetum degeneration. *Plant Cell* 23, 1416–1434. doi: 10.1105/tpc.110.082636

Liu, H. J., Hu, M. H., Wang, Q., Cheng, L., and Zhang, Z. B. (2018). Role of papain-like cysteine proteases in plant development. *Front. Plant Sci.* 9, 1717. doi: 10.3389/fpls.2018.01717

Liu, Y., Wang, K. R., Cheng, Q., Kong, D. Y., Zhang, X. Z., Wang, Z. B., et al. (2020). Cysteine protease RD21A regulated by E3 ligase SINAT4 is required for drought-induced resistance to *Pseudomonas syringae* in arabidopsis. *J. Exp. Bot.* 71, 5562–5576. doi: 10.1093/jxb/eraa255

Luo, J., Zhang, H., He, W. W., Zhang, Y., Cao, W. L., Zhang, H. S., et al. (2016). *OsSNAP32*, a SNAP25-type SNARE protein-encoding gene from rice, enhanced resistance to blast fungus. *Plant Growth Regul.* 80, 37–45. doi: 10.1007/s10725-016-0152-4

Lv, Q. M., Xu, X., Shang, J. J., Jiang, G. H., Pang, Z. Q., Zhou, Z. Z., et al. (2013). Functional analysis of Pid3-A4, an ortholog of rice blast resistance gene Pid3 revealed by allele mining in common wild rice. *Phytopathology* 103, 594–599. doi: 10.1094/PHYTO-10-12-0260-R

Misas-Villamil, J. C., van der Hoorn, R. A., and Doehlemann, G. (2016). Papain-like cysteine proteases as hubs in plant immunity. *New Phytol.* 212, 902–907. doi: 10.1111/nph.14117

Nakashima, A., Chen, L., Thao, N., Fujiwara, M., Wong, H., Kuwano, M., et al. (2008). RACK1 functions in rice innate immunity by interacting with the Rac1 immune complex. *Plant Cell* 20, 2265–2279. doi: 10.1105/tpc.107.054395

Nawaz, G., Usman, B., Peng, H. W., Zhao, N., Yuan, R. Z., Liu, Y. G., et al. (2020). Knockout of *Pi21* by CRISPR/Cas9 and iTRAQ-based proteomic analysis of mutants revealed new insights into *M. oryzae* resistance in elite rice line. *Genes* 11, 735. doi: 10.3390/genes11070735

Ormancey, M., Thuleau, P., van der Hoorn, R. A. L., Grat, S., Testard, A., Kamal, K. Y., et al. (2019). Sphingolipid-induced cell death in arabidopsis is negatively regulated by the papain-like cysteine protease RD21. *Plant Sci.* 280, 12–17. doi: 10.1016/j.plantsci.2018.10.028

Park, C. H., Chen, S., Shirsekar, G., Zhou, B., Khang, C. H., Songkumarn, P., et al. (2012). The *Magnaporthe oryzae* effector AvrPiz-t targets the RING E3 ubiquitin ligase APIP6 to suppress pathogen-associated molecular pattern-triggered immunity in rice. *Plant Cell* 24, 4748–4762. doi: 10.1105/tpc.112.105429

Perez-Lopez, E., Hossain, M. M., Wei, Y., Todd, C. D., and Bonham-Smith, P. C. (2021). A clubroot pathogen effector targets cruciferous cysteine proteases to suppress plant immunity. *Virulence* 12, 2327–2340. doi: 10.1080/21505594.2021.1968684

Rustgi, S., Boex-Fontvieille, E., Reinbothe, C., von Wettstein, D., and Reinbothe, S. (2017). Serpin1 and WSCP differentially regulate the activity of the cysteine protease RD21 during plant development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 114, 2212–2217. doi: 10.1073/pnas.1621496114

Shimono, M., Sugano, S., Nakayama, A., Jiang, C. J., Ono, K., Toki, S., et al. (2007). Rice WRKY45 plays a crucial role in benzothiadiazole-inducible blast resistance. *Plant Cell* 19, 2064–2076. doi: 10.1105/tpc.106.046250

Shindo, T., Misas-Villamil, J. C., Horger, A. C., Song, J., and van der Hoorn, R. A. (2012). A role in immunity for arabidopsis cysteine protease RD21, the ortholog of the tomato immune protease C14. *PloS One* 7, e29317. doi: 10.1371/journal.pone.0029317

Singh, M. P., Lee, F. N., Counce, P. A., and Gibbons, J. H. (2004). Mediation of partial resistance to rice blast through anaerobic induction of ethylene. *Phytopathology* 94, 819–825. doi: 10.1094/PHYTO.2004.94.8.819

Spence, C. A., Lakshmanan, V., Donofrio, N., and Bais, H. P. (2015). Crucial roles of abscisic acid biogenesis in virulence of rice blast fungus *Magnaporthe oryzae*. *Front. Plant Sci.* 6, 1082. doi: 10.3389/fpls.2015.01082

Tao, H., Shi, X. T., He, F., Wang, D., Xiao, N., Fang, H., et al. (2021). Engineering broad-spectrum disease-resistant rice by editing multiple susceptibility genes. *J. Integr. Plant Biol.* 63, 1639–1648. doi: 10.1111/jipb.13145

van der Hoorn, R. A., Leeuwenburgh, M. A., Bogyo, M., Joosten, M. H., and Peck, S. C. (2004). Activity profiling of papain-like cysteine proteases in plants. *Plant Physiol.* 135, 1170–1178. doi: 10.1104/pp.104.041467

van Loon, L. C., Geraats, B. P., and Linthorst, H. J. (2006). Ethylene as a modulator of disease resistance in plants. *Trends Plant Sci.* 11, 184–191. doi: 10.1016/j.tplants.2006.02.005

Wang, Y., Duan, G., Li, C., Ma, X., and Yang, J. (2021). Application of jasmonic acid at the stage of visible brown necrotic spots in *Magnaporthe oryzae* infection as a novel and environment-friendly control strategy for rice blast disease. *Protoplasma* 258, 743–752. doi: 10.1007/s00709-020-01591-0

Xiao, G. Q., Qin, H., Zhou, J. H., Quan, R. D., Lu, X. Y., Huang, R. F., et al. (2016). OsERF2 controls rice root growth and hormone responses through tuning expression of key genes involved in hormone signaling and sucrose metabolism. *Plant Mol. Biol.* 90, 293–302. doi: 10.1007/s11103-015-0416-9

Xiao, N., Wu, Y. Y., and Li, A. L. (2020). Strategy for use of rice blast resistance genes in rice molecular breeding. *Rice Sci.* 27, 263–277. doi: 10.1016/j.rsci.2020.05.003

Xiong, L. Z., and Yang, Y. N. (2003). Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acid-inducible mitogen-activated protein kinase. *Plant Cell* 15, 745–759. doi: 10.1105/tpc.008714

Yang, D. L., Yang, Y., and He, Z. (2013). Roles of plant hormones and their interplay in rice immunity. *Mol. Plant* 6, 675–685. doi: 10.1093/mp/sst056

Yin, J. J., Zou, L. J., Zhu, X. B., Cao, Y. Y., He, M., and Chen, X. W. (2021). Fighting the enemy: How rice survives the blast pathogen's attack. *Crop J.* 9, 543–552. doi: 10.1016/j.cj.2021.03.009

Yu, M. X., Zhou, Z. Z., Liu, X., Yin, D. D., Li, D. Y., Zhao, X. F., et al. (2021). The OsSPK1-OsRac1-RAI1 defense signaling pathway is shared by two distantly related NLR proteins in rice blast resistance. *Plant Physiol.* 187, 2852–2864. doi: 10.1093/plphys/kiab445

Zhai, K. R., Liang, D., Li, H. L., Jiao, F. Y., Yan, B. X., Liu, J., et al. (2022). NLRs guard metabolism to coordinate pattern-and effector-triggered immunity. *Nature* 601, 245–251. doi: 10.1038/s41586-021-04219-2

Zhang, B., Tremousaygue, D., Denance, N., van Esse, H. P., Horger, A. C., Dabos, P., et al. (2014). PIRIN2 stabilizes cysteine protease XCP2 and increases susceptibility to the vascular pathogen *Ralstonia solanacearum* in arabidopsis. *Plant J.* 79, 1009–1019. doi: 10.1111/tpj.12602

Zhang, D. P., Wang, Y. Z., Shen, J. Y., Yin, J. F., Li, D. H., Gao, Y., et al. (2018). *OsRACK1A*, encodes a circadian clock-regulated WD40 protein, negatively affect salt tolerance in rice. *Rice* 11, 45. doi: 10.1186/s12284-018-0232-3

Zhou, X. G., Liao, H. C., Chern, M., Yin, J. J., Chen, Y. F., Wang, J. P., et al. (2018). Loss of function of a rice TPR-domain RNA-binding protein confers broad-spectrum disease resistance. *Proc. Natl. Acad. Sci. U.S.A.* 115, 3174–3179. doi: 10.1073/pnas.1705927115

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 16 January 2023
DOI 10.3389/fpls.2022.1041081

OPEN ACCESS

EDITED BY
Xueyong Li,
Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, China

REVIEWED BY
Jian Sun,
Shenyang Agricultural University, China
Tao Guo,
South China Agricultural University, China
Di Cui,
Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, China

*CORRESPONDENCE
Xi Yuan
xiyuan@zjnu.edu.cn
Yuchun Rao
ryc@zjnu.cn
Yuexing Wang
wangyuexing@caas.cn

SPECIALTY SECTION
This article was submitted to Plant Bioinformatics, a section of the journal Frontiers in Plant Science

RECEIVED 10 September 2022
ACCEPTED 28 December 2022
PUBLISHED 16 January 2023

CITATION
Yin W, Lu T, Chen Z, Lu T, Ye H, Mao Y, Luo Y, Lu M, Zhu X, Yuan X, Rao Y and Wang Y (2023) Quantitative trait locus mapping and candidate gene analysis for salt tolerance at bud stage in rice. *Front. Plant Sci.* 13:1041081. doi: 10.3389/fpls.2022.1041081

Quantitative trait locus mapping and candidate gene analysis for salt tolerance at bud stage in rice

Wenjing Yin[1], Tianqi Lu[1], Zhengai Chen[1], Tao Lu[1], Hanfei Ye[1], Yijian Mao[2], Yiting Luo[1], Mei Lu[1], Xudong Zhu[2], Xi Yuan[1]*, Yuchun Rao[1]* and Yuexing Wang[2]*

[1]College of Chemistry and Life Sciences, Zhejiang Normal University, Jinhua, Zhejiang, China, [2]State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou, Zhejiang, China

Soil salinization has a serious influence on rice yield and quality. How to enhance salt tolerance in rice is a topical issue. In this study, 120 recombinant inbred line populations were generated through nonstop multi-generation selfing using a male indica rice variety Huazhan (*Oryza sativa* L. subsp. *indica* cv. 'HZ') and a female variety of Nekken2 (*Oryza sativa* L. subsp. *japonica* cv. 'Nekken2') as the parents. Germination under 80 mM NaCl conditions was measured and analyzed, and quantitative trait locus (QTL) mapping was completed using a genetic map. A total of 16 salt-tolerance QTL ranges were detected at bud stage in rice, which were situated on chromosomes 3, 4, 6, 8, 9, 10, 11, and 12. The maximum limit of detection was 4.69. Moreover, the *qST12.3* was narrowed to a 192 kb region on chromosome 12 using map-based cloning strategy. Statistical analysis of the expression levels of these candidate genes under different NaCl concentrations by qRT-PCR revealed that *qST12.3* (*LOC_Os12g25200*) was significantly down-regulated with increasing NaCl concentration, and the expression level of the chlorine-transporter-encoding gene *LOC_Os12g25200* in HZ was significantly higher than that of Nekken2 under 0 mM NaCl. Sequencing analysis of *LOC_Os12g25200* promoter region indicated that the gene expression difference between parents may be due to eight base differences in the promoter region. Through QTL mining and analysis, a plurality of candidate genes related to salt tolerance in rice was obtained, and the results showed that *LOC_Os12g25200* might negatively regulate salt tolerance in rice. The results provide the basis for further screening and cultivation of salt-tolerant rice varieties and have laid the foundation for elucidating further molecular regulation mechanisms of salt tolerance in rice.

KEYWORDS

rice, germination rate, salt stress, QTL mapping, recombinant inbred line population

Introduction

Soil salinization is a global problem that severely restricts the cultivation of rice and is one of the predominant abiotic stresses affecting the rice crop yield (Munns and Gilliham, 2015; Yang and Guo, 2018). According to reported statistics, about 100 million hm^2 of the world's land is affected by salinization. About 10 million hm^2 of this land is in China, accounting for about 10% of global soil salinization regions, which severely limits food productivity in China (Hu et al., 2018; Liu et al., 2021; Zou et al., 2021). Rice (*Oryza sativa* L.), a salt-sensitive crop, is a gramineous plant and one of the three most essential staple crops in China and throughout the world (Tian et al., 2011; Ganie et al., 2019). It is most susceptible to salt stress at the seedling stage and heading stage (Zou et al., 2021). Salt tolerance is a complicated physiological process. When under salt stress, the rice in cellular salt content disturbs the cells' ionic and osmotic stability, having an effect on the health and productivity of plants. At this time, plant cells try to maximize the dynamic stability of ions, and thus plant health, by inducing gene expression (Wu et al., 2013; Singh et al., 2021). Those attempting to increase the yield and quality of rice face new challenges (Rao et al., 2020), and efforts to enhance the salt tolerance of rice will assist in maximizing the use of saline-alkali land, and hence, improve crop yields (Radanielson et al., 2018; Wang et al., 2019). Therefore, exploring the mechanisms of rice salt tolerance is anticipated to provide a theoretical groundwork that will lead to a realistic way to cultivate salt-tolerant rice types, enhancing rice yield in saline land and, consequently, securing food production in China and other parts of the world.

Seed germination is critical, in which the catabolism of proteins, starches or oils accumulated by seeds during maturation supports the growth of early seedlings. Therefore, seed germination is also the basis for the total biomass and total production of plants throughout their life cycles (Bewley, 1997; Penfield, 2017). In the process of seed germination, the high salt environment will lead to the disorder of various metabolic pathways in plants, change the enzyme activities, and the leakage of high solutes, resulting in the reduction of K^+ efflux, and hinder the activity of α-amylase (Gomes-Filho et al., 2008). Some studies identified about 50 loci have been identified for seed germination under salt stress by genome-wide association studies (GWAS) (Shi et al., 2017; Yu et al., 2018). He et al. (2019) located a QTL *qSE3* during the germination and seedling stage of rice seeds under salt stress, cloned and isolated *qSE3* and found that it encoded the K^+ transporter *OsHAK21*. *qSE3* significantly enhanced the uptake of Na^+ and K^+ by germinating seeds under salt stress. Zeng et al. (2021) identified 13 QTLs from the local cultivar Wujiaozhan, and a major salt-tolerance specific QTL *qGR6.2* was precisely mapped. These studies provided information on the genetic basis of marker-assisted selection to improve rice seed germination under salt stress.

Salt tolerance in rice is a complex quantitative trait regulated by multiple genes, and the expression of these related genes is closely related to the environment (Qi et al., 2008; Tiwari et al., 2016). Salt stress can cause changes to the activity of electron transport chains in mitochondria and chloroplasts and destroy the dynamic balance between the generation and removal of reactive oxygen species (ROS), leading to protein oxidation, membrane damage, and DNA fragmentation, and other disruptions (Mishra et al., 2013). At present, QTL mapping studies have found the highest number of genes related to salt tolerance on chromosomes 2 and 6, and the fewest genes related to salt tolerance on chromosomes 10, 11, and 12 (Hu et al., 2018). Abdur et al. (Razzaque et al., 2011) found that high salt stress affected the balance of Na^+, K^+, and Cl^- in the roots of rice plants. Na^+ was the main ion that caused harm to rice under salt stress (Munns and Tester, 2008), and Na^+ transporters such as HKTs, NHXs and ROS scavengers play key roles in maintaining cell homeostasis under salt stress and positively affect the ability of plants to cope with increased salinity. High concentrations of Na^+ prevent the normal functioning of rice biofilms, affect the normal absorption of nutrients, and lead to physiological metabolic disorders (Bharathkumar et al., 2016). A high concentration of Cl^- was shown to affect the normal growth and development of the rice plant. The influencing mechanisms for this might be anion mediation of the periodic regulation and inhibition of ribosomal enzymes that catalyze protein synthesis by cells (Geilfus, 2018). Joo et al. (2014) found that CatA, CatB, and CatC were all involved in the regulation of plant abiotic stress. Some results indicated that the Cat protein was also involved in the response of plants to salt stress (Nagamiya et al., 2007). Zhou et al. (2018) found that the receptor-like cytoplasmic kinase STRK1 phosphorylated and activated CatC which then positively regulated the salt tolerance of rice, which has a practical significance in rice yield and quality.

In this study, recombinant inbred line (RIL) populations constructed using Huazhan (HZ) and Nekken2 rice varieties as parents were used to locate and analyze QTL for salt tolerance. This was achieved by examining the response of rice seeds to salt stress at the germination stage and creating a high-density single nucleotide polymorphism (SNP)-based molecular marker map. Candidate genes were screened and expression was analyzed for QTL clusters with high effector values to ascertain the identity of additional QTL and genes related to the regulation of salt stress and provide a molecular research basis and theoretical support for the screening of salt-tolerant rice germplasm resources.

Materials and methods

Experimental materials

Huazhan (*Oryza sativa* L. subsp. *indica* cv. 'HZ') is a kind of super rice variety, which has strong adaptability, better resistance to adversity and stress than ordinary varieties, and better resistance to diseases and pests. It belongs to the varieties with high and stable yield. Nekken2 (*Oryza sativa* L. subsp. *japonica* cv. 'Nekken2') is a kind of rice with strong affinity and strong spectral affinity, and its F_1 seed setting rate is up to 80%, which has important utilization value in Chinese rice breeding of super high yield (Sun et al., 1992). In this experiment, an HZ plant was used as the male parent to hybridize with Nekken2, used as the female parent, and obtain the F_1 generation. The F_1 generation was bagged and selfed *via* the single-seed descent method, and 120 stably inherited recombinant inbred lines were obtained after continuous selfing for 12 generations to form the RIL population (Wang et al., 2020).

Rice cultivation and management

For each strain of rice seed, 60 seeds were taken from each of the 120 strains of F_{12} and both parents, rinsed 2-3 times with 70% alcohol, and surface-disinfected with 10% sodium hypochlorite solution. The seeds were then rinsed well with deionized water. After 1 month, 24 seedlings of each strain were selected and transplanted in the experimental field, with each strain planted in four rows of six plants. Routine water and fertilizer management and insect pest and weed control strategies were carried out during this period.

Salt stress tolerant treatments

Sixty full-seeded seeds of HZ, Nekken2, and each hybrid strain were selected, hulled, disinfected with 70% alcohol for 2 minutes, disinfected with 30% sodium hypochlorite solution for 20-30 minutes, rinsed with distilled water, inoculated onto 1/2 MS medium containing 80 mM NaCl (Zhang, 2012), and incubated in an artificial climate incubator for 10 days. The germination rate of each strain was determined (germination standard: Refer to GB 5520-85 for the germination standard of rice seeds: the young roots should reach the seed length, and the young buds should reach at least half the length of the grain) (Meng et al., 2022). Three replicates were set up for each group. The incubator conditions were light/dark (14 h/10 h), 30°C, and 70% humidity.

Genetic mapping construction

DNA from the biparental HZ and Nekken2 and 120 recombinant self-inbred line populations were extracted and the genomes re-sequenced. The genome sequencing results were sorted to obtain 4858 molecular markers evenly distributed throughout 12 chromosomes for the construction of the genetic map (Wang et al., 2020).

QTL localization

Using a high-density SNP molecular marker linkage map method already established in the laboratory, interval mapping was used to analyze the QTL for seed germination rate separately using Mapmaker/QTL1.1B software. An LOD setting of 2.5 was applied as the threshold value to determine the presence of the QTL and we following the rules proposed by Mccouch et al. (1997) to name the QTL.

The development of CSSLs and fine mapping for *qST12.3*

To obtain Chromosome Segment Substitution Lines (CSSL) for *qST12.3*, the plants carrying HZ genotype at the flanking region of *qST12.3* were selected to backcross with Nekken2 for 3 rounds. The SSR markers RM101 and RM1337 were simultaneously used to identify the plants containing HZ genotype in backcross lines. A set of 124 SSR markers (Supplementary Table S1) uniformly distributed on a previous linkage map were used to select the individual plants containing the least HZ DNA in BC_4F_1 lines. It contained a small amount of HZ DNA in its genetic background, and carried a homozygous introgression across the entire *qST12.3* region and without any introgression across *qST3*, *qST4.1*, *qST4.2*, *qST6*, *qST8*, *qST9*, *qST10.1*, *qST10.2*, *qST11.1*, *qST11.2*, *qST11.3*, *qST11.4*, and *qST11.5* regions on chromosomes 3, 4, 6, 6, 9, 10, and 11 respectively. Then, a total of 186 selected plants from 770 BC_4F_2 progenies were cultivated in the paddies to gain enough seeds for the salt stress tolerant evaluation and further fine mapping.

Candidate gene expression analysis

Using the Total Plant RNA Extraction Kit (Axygen), total RNA was extracted from HZ and Nekken2 leaves, and 1 μg of total RNA was aspirated to reverse transcribe into cDNA using the cDNA Reverse Transcription Kit (Toyobo). Candidate genes were predicted based on the QTL localization results combined with the Rice Genome Annotation Database (http://rice.plantbiology.msu.edu/). In qRT-PCR, the SYBR Green Realtime PCR Master Mix (Toyobo) and the primers (Supplementary Table S2) were used to detect the expression levels of candidate salt-tolerance-related genes, with rice *OsActin* as the internal reference gene.

The total volume of the qRT-PCR reaction system was 10 μL, including 2 μL of cDNA template, 5 μL of SYBR qPCR Mix (Toyobo), 1 μL each of upstream and downstream primer (10 μmol/L^{-1}), and 1 μL of ddH_2O.

The qRT-PCR reaction program was 95°C for 30 s, 95°C for 5 s, 55°C for 10 s, 72°C for 15 s, and 45 cycles. Three replicates were set up for each reaction and relative quantification was performed using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). The data obtained were analyzed for significant differences using Excel for t-tests and GraphPad Prism 6 software for graphical analysis. $P<0.05$ indicates a significant difference, while $P<0.01$ indicates a highly significant difference.

Salt gradient stress treatment

Sixty seeds of each parent, HZ and Nekken2, were sterilized and inoculated onto 1/2 MS medium containing 0 mM, 60 mM, 80 mM, or 100 mM NaCl for 14 days. Candidate genes were analyzed for expression. Based on difference in the expression of candidate genes between the parents, candidate genomic DNA fragments were amplified from both parents using high-fidelity DNA polymerase KOD-Plus-Ver.2 (Toyobo) and the PCR products were sequenced.

Results and analysis

Phenotypes of both parents and RIL groups

The results of salt-tolerant treatments of both parents showed that there was a significant difference between HZ and Nekken2 (Figure 1). The germination rates (GR) of seeds after salt treatment

10.3389/fpls.2022.1041081

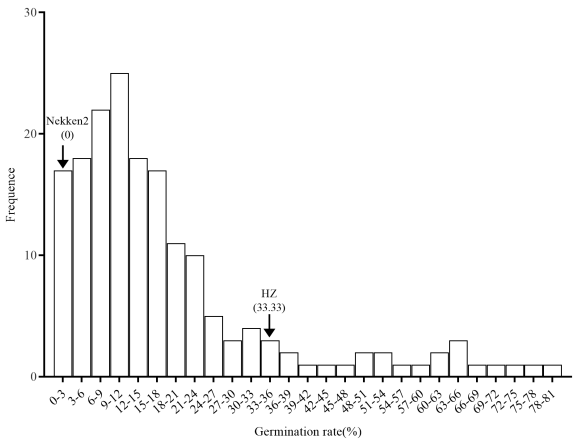


FIGURE 1
Germination rate distribution of rice recombinant inbred lines under salt treatment.

were found to be 0% for Nekken2 and approximately 33.33% for HZ, the latter of which was presumed to be more salt tolerant than Nekken2 at a concentration of 80 mM NaCl (Figures 2, S1). The germination rate data for each strain of the RIL population under salt stress showed a continuous normal distribution, and several strains showed super parental germination rates. In addition, the genetic characteristics of the quantitative traits were consistent with the requirements of QTL interval mapping.

Salt tolerance QTL localization analysis

A total of 16 QTLs related to salt tolerance were detected on chromosomes 3, 4, 6, 8, 9, 10, 11, and 12 using the molecular linkage map established in our laboratory. The highest LOD value was 4.69 and was located with the genetic distances 23.86 cM and 24.48 cM on chromosome 10 (Table 1; Figure 3).

Functional analysis of candidate genes

Based on the results of the QTL localization analysis and fine mapping, the location of the salt tolerance interval was determined, and the functions of the genes in each QTL interval were analyzed in conjunction with the rice gene database (http://rice.plantbiology.msu.edu/). The functions of and physiological information on these genes were briefly summarized (Table 2). Eight of these genes were cloned, including *LOC_Os03g12730* (*OsRLK1*), *LOC_Os03g12820* (*OsSRO1c*), and *LOC_Os03g12840* (*OsITPK2*) on chromosome 3, *LOC_Os09g37949* (*OsRPK1*) and *LOC_Os09g38000* (*OsNAC109*) on chromosome 9; *LOC_Os10g01480* (*OsITPK1*) on chromosome 10, *LOC_Os11g10590* (*OsDT11*) on chromosome 11, *LOC_Os12g12860* (*OSIPK*) and *LOC_Os12g24800* (*OsNCED2*) on chromosome 12 (Table 2; Figure 3). Moreover, the *qST12.3* was fine-mapped to a 192 kb region on rice chromosome 12 using map-based cloning strategy.

Candidate gene expression level analysis

Analysis of salt tolerance candidate genes by qRT-PCR after 14 days of treatment with 0, 60, 80 or 100 mM salt stress of both parents, HZ and Nekken2, revealed (Figures 4A–D) that, at 0 mM NaCl, *LOC_Os03g12730*, *LOC_Os03g12820*, *LOC_Os03g12840*, *LOC_Os10g01480*, *LOC_Os11g10590*, *LOC_Os11g10640*, *LOC_Os12g10740*, *LOC_Os12g12860*, *LOC_Os12g13570*, *LOC_Os12g24800*, and *LOC_Os12g25200* were significantly differentially expressed in the two parents; at 60 mM NaCl, *LOC_Os03g12820* and *LOC_Os12g24800* were significantly differentially expressed in the two parents, and *LOC_Os03g12840*, *LOC_Os09g37949*, *LOC_Os09g38000*, *LOC_Os10g01760*, *LOC_Os11g10590*, *LOC_Os12g10740*, *LOC_Os12g12860*, and *LOC_Os12g13570*, *LOC_Os12g25200* were very significantly differently expressed in the two parents; at 80 mM NaCl concentration, *LOC_Os03g12730*, *LOC_Os12g12860* had significantly different in expression in the two parents, and *LOC_Os03g12820*, *LOC_Os03g12840*, *LOC_Os09g37949*, *LOC_Os09g38000*, *LOC_Os10g01480*, *LOC_Os10g01760*, *LOC_Os11g10640*, *LOC_Os12g10660*, *LOC_Os12g24800*, and *LOC_ Os12g25200* were very significantly different in expression in the two parents; and at 100 mM NaCl concentration, *LOC_Os09g37949*, *LOC_Os10g01480*, *LOC_Os12g12860*, and *LOC_Os12g13570* were expressed significantly differently between the two parents, and *LOC_Os03g12730*, *LOC_Os03g12820*, *LOC_Os03g12840*, *LOC_Os11g10590*, *LOC_Os11g10640*, *LOC_Os12g10660*, *LOC_Os12g25200* were significantly differentially between the two parents.

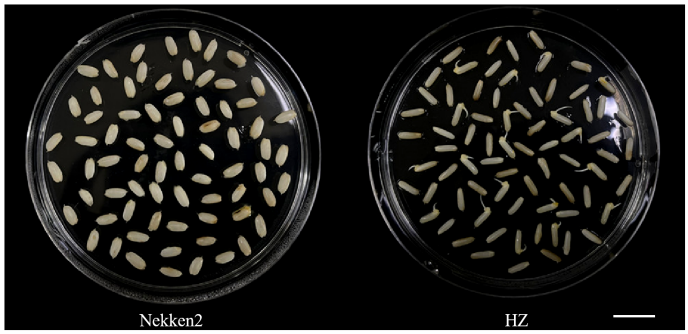


FIGURE 2
Phenotype diagram of germination rate of Huazan and Nekken2 under 80mM salt stress, bar=5cm.

10.3389/fpls.2022.1041081

TABLE 1 Mapping of QTL for salt tolerance in rice.

QTL	Chr	Physical distance (bp)	Position of support (cM)	LOD
qST3	3	6743718~7034710	28.90~30.16	3.69
qST4.1	4	17123204~17600552	73.40~75.45	3.32
qST4.2	4	19482788~19642144	83.51~84.20	3.16
qST6	6	8917905~8938061	38.22~38.32	2.61
qST8	8	3764126~4411594	16.13~18.92	2.94
qST9	9	21667334~21932045	92.88~94.02	2.80
qST10.1	10	57733~1093124	0.24~4.69	3.05
qST10.2	10	5567413~5709027	23.86~24.48	4.69
qST11.1	11	5588569~5889092	23.95~25.25	2.59
qST11.2	11	6190775~6314547	26.53~27.44	2.82
qST11.3	11	8013951~8692604	34.35~37.26	2.67
qST11.4	11	9438113~9808325	40.45~42.05	2.82
qST11.5	11	19809355~19870739	84.91~85.18	3.16
qST12.1	12	5494931~6224960	23.55~26.68	2.59
qST12.2	12	6959601~8245853	29.83~35.35	2.83
qST12.3	12	14141879~15247966	60.62~65.36	3.00

Under 0 mM NaCl, the expression of *LOC_Os12g25200* in HZ was significantly higher than that in Nekken2. After treatment with 60, 80, or 100 mM NaCl, the expression of this gene decreased significantly in HZ. The expression of this gene decreased sequentially with the increase in salt concentration and reached the lowest expression at 100 mM NaCl. The quantitative results showed that, with increasing salt concentration, the resistance of the plants to salt stress was maintained by a decrease in *LOC_Os12g25200* expression in HZ, which maintained the normal growth and development of the plants.

Sequence analysis of the candidate gene *LOC_Os12g25200*

Analysis of the effects of stress treatment with the different concentrations of salt *via* quantitative qRT-PCR showed that the expression of *LOC_Os12g25200* was significantly higher in HZ than in Nekken2 in the absence of salt stress, while expression of this gene gradually decreased with increasing salt concentration, suggesting this may be a novel gene that negatively regulates salt tolerance in rice.

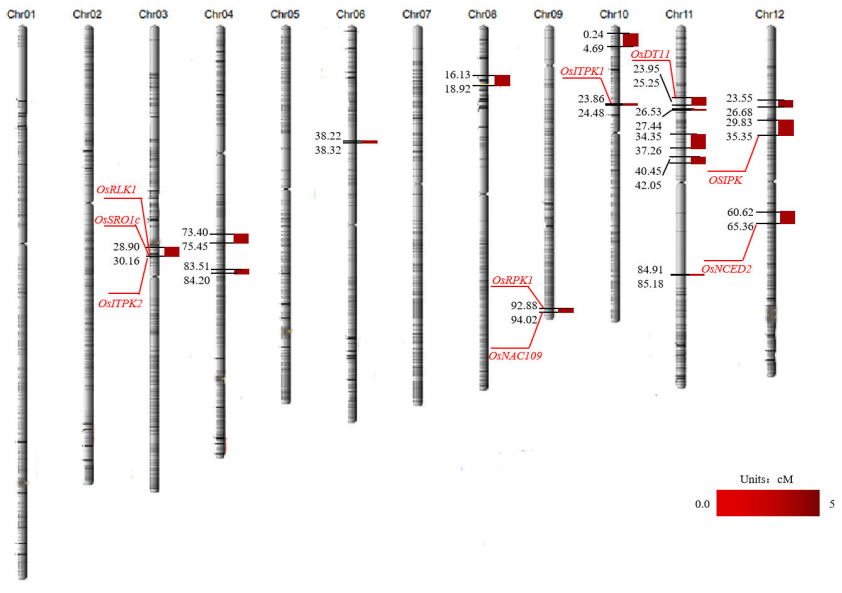


FIGURE 3
Mapping of QTL for salt tolerance in rice. The red gene number is the cloned salt tolerance gene screened in the salt tolerance interval.

TABLE 2 Functional analysis of candidate genes in QTL mapping interval.

Gene ID	Chr	Function	Cloned or not
LOC_Os03g12730	3	Leucine-rich repeat receptor-like protein kinases	Have been cloned
LOC_Os03g12820	3	(SRO) SRO protein; SNAC1 target gene; Callus Browning	Have been cloned
LOC_Os03g12840	3	1,3,4-phosphoinositol 5/6-kinase gene	Have been cloned
LOC_Os09g37949	9	Receptor protein kinase	Have been cloned
LOC_Os09g38000	9	NAC transcription factor	Have been cloned
LOC_Os10g01480	10	1,3,4-phosphoinositol 5/6-kinase gene	Have been cloned
LOC_Os10g01760	10	Peroxidase precursors	Not cloning
LOC_Os11g10590	11	Drought tolerance	Have been cloned
LOC_Os11g10640	11	A protein containing a protein kinase domain	Not cloning
LOC_Os12g10660	12	B-box zinc finger protein	Not cloning
LOC_Os12g10740	12	Leucine-rich repeating family of proteins	Not cloning
LOC_Os12g12860	12	Calcium-dependent protein kinase	Have been cloned
LOC_Os12g13570	12	MYB family transcription factors	Not cloning
LOC_Os12g24800	12	9-cis-epoxy carotenoid dioxygenase gene *OsNCED2*	Have been cloned
LOC_Os12g25200	12	Chlorine transporters, a family of chlorine channels	Not cloning

Therefore, the *LOC_Os12g25200* gene in HZ and Nekken2 was sequenced and analyzed for locus differences, and a total of four single-base differences were found in the exons (Figures 5A–E), located at 1073 bp (G in HZ vs T in Nekken2), 1139 bp (G in HZ vs T in Nekken2), 1959 bp (G in HZ va C in Nekken2), and at 2352 bp (G in HZ vs A in Nekken2). Among these, the single-base differences at 1073 bp and 1139 bp caused amino acid changes, located at position 60 (glycine in HZ vs valine in Nekken2) and position 82 (tryptophan in HZ vs leucine in Nekken2), respectively, while in the protein, the other two single-base differences did not cause amino acid changes. Therefore, it was speculated that the difference in the expression of the *LOC_Os12g25200* gene in HZ and Nekken2 may be caused by differences in the amino acid sequences encoded by their genes, which in turn leads to differences in salt tolerance between HZ and Nekken2. In order to further explore the reasons for the difference in the expression level of this gene between parents, we sequenced the first 2200 bp promoter sequence of *LOC_Os12g25200* initiation codon and a total of eight single-base differences were found in the exons (Supplementary Figures S2A–H), located at 733 bp (C in HZ vs A in Nekken2), 937 bp (A in HZ vs T in Nekken2), 1005-1006 bp (Nekken2 is deficiency G and A), 1058 bp (C in HZ vs T in Nekken2), 1119 bp (G in HZ vs A in Nekken2), 1174 bp (C in HZ vs T in Nekken2), 1605 bp (C in HZ vs T in Nekken2) and at 1979 bp (A in HZ vs G in Nekken2). We speculated that the single base differences in the eight promoter regions above might lead to the differences in gene expression levels between parents.

Substitution mapping of *qST12.3* and analyzing candidate genes

Three lines from the RIL population were selected to backcross with Nekken2 for 4 rounds. According to the QTL analysis and a previous genetic linkage map, the SSR markers RM101 and RM1337 were used in marker-assisted selection for segregating the progenies carrying *qST12.3* during every generation of backcross. After 4 rounds of backcrosses with Nekken2, the BC_4F_1 and BC_4F_2 generations were scanned with a set of 124 SSR markers, which were uniformly distributed on a previous linkage map (Supplementary Table S1). The plant CSSL12 containing a small amount of HZ DNA was selected. It carried a homozygous introgression across the entire *qST12.3* region, but without any introgression across *qST3*, *qST4.1*, *qST4.2*, *qST6*, *qST8*, *qST9*, *qST10.1*, *qST10.2*, *qST11.1*, *qST11.2*, *qST11.3*, *qST11.4*, and *qST11.5* regions on chromosomes 3, 4, 6, 8, 9, 10, and 11, respectively (Figure 6). The GR of CSSL12 was significantly higher than that of its recurrent parent Nekken2, but was much more similar to that of HZ (Figure 7).

The target region contains 2 predicted genes (*LOC_Os12g24800* and *LOC_Os12g25200*) based on the Rice Genome Annotation Website (http://rice.plantbiology.msu.edu/). The *LOC_Os12g24800* (*OsNCED2*) has been demonstrated by Sandra et al. (Oliver et al., 2007) to increase the ABA level in anthers of Doongara rice after low temperature treatment, which proved that *LOC_Os12g24800* can regulate endogenous ABA in plants under cold stress. We performed a quantitative reverse transcription-PCR (qRT-PCR) analysis, the results show that *LOC_Os12g24800* and *LOC_Os12g25200* indicated significantly different expression levels in HZ and CSSL12 compared with Nekken2 (Figures 8A–D). The expression levels for *LOC_Os12g25200* decreased with the increase of salt stress concentration. In conclusion, *LOC_Os12g25200* is likely to be the candidate genes for *qST12.3*.

Discussion

Salt tolerance in rice is a quantitative trait regulated by multiple genes. After treating HZ and Nekken2 with different concentrations of

10.3389/fpls.2022.1041081

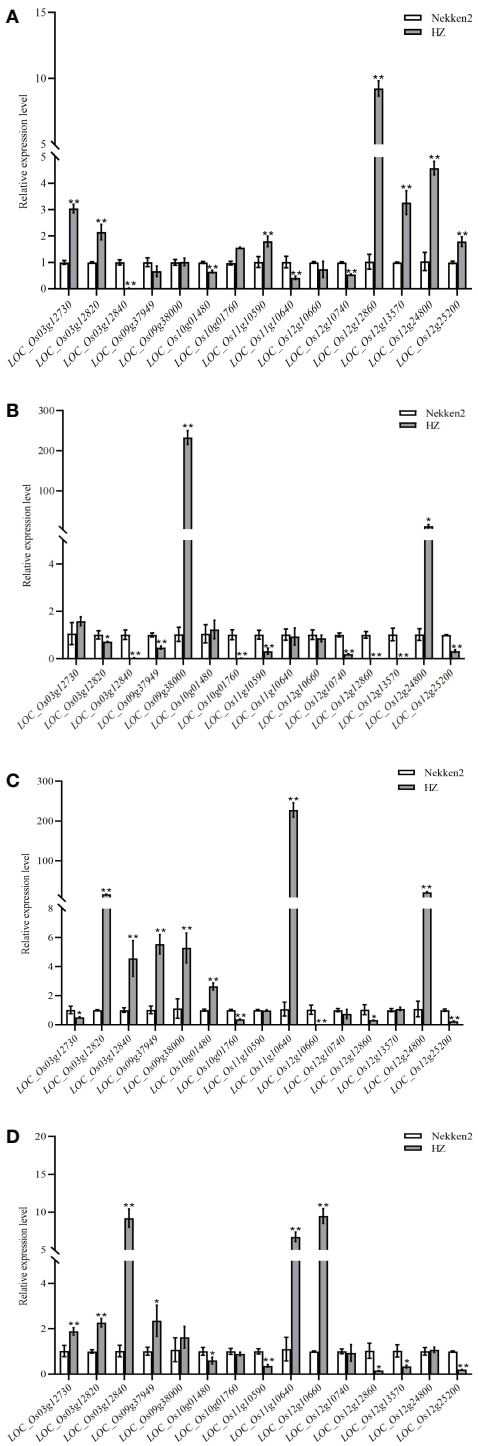


FIGURE 4
Expression of candidate genes for salt tolerance in rice under different concentrations of NaCl. **(A)** Expression of candidate genes at 0 mM NaCl concentration. **(B)** Expression of candidate genes at 60 mM NaCl concentration. **(C)** Expression of candidate genes at 80 mM NaCl concentration. **(D)** Expression of candidate genes at 100 mM NaCl concentration. * indicates significance at $P \leq 0.05$ by Student's *t* test, ** indicates significance at $P \leq 0.01$ by Student's *t* test.

salt, analysis using qRT-PCR revealed significant differences in the expression of candidate genes in the two parents. Different treatments, such as the different growing environments of rice, differences between parents, and the determination of the germination rate at each salt concentration could affect the results of the QTL localization to varying degrees. In this experiment, we used salt stress tolerance as the only indicator for QTL detection. Based on the results of QTL interval localization, we combined with qRT-PCR and mapping of substitution lines, we explored and analyzed possible genes involved in regulating salt stress tolerance in rice to provide a theoretical basis for expanding the use of land with high soil salinity and improving the yield and quality of rice under this condition.

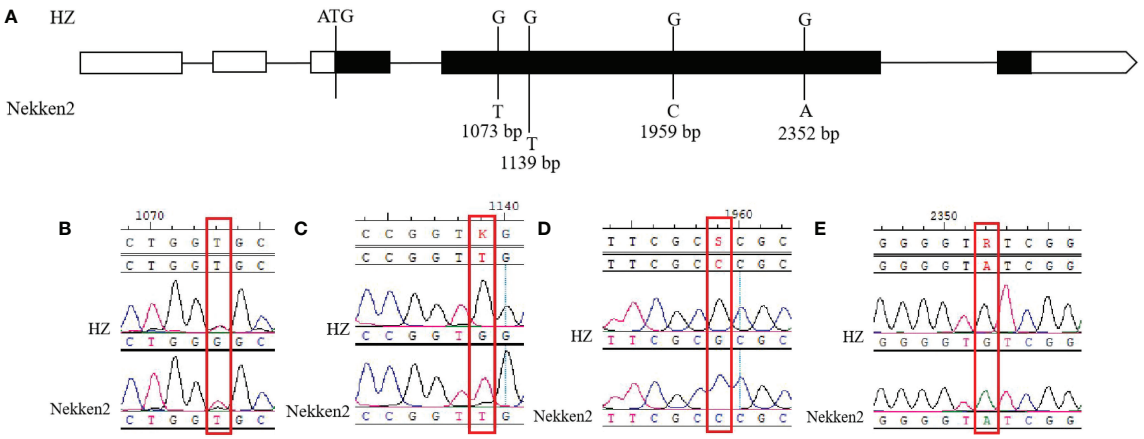


FIGURE 5
Sequence analysis of *LOC_Os12g25200* gene. The black region is the exon region, where single base mutations occur. **(A)** DNA loci difference analysis of *LOC_Os12g25200* gene. **(B)** Site difference at 1073 bp. **(C)** Site difference at 1139 bp. **(D)** Site difference at 1959 bp. **(E)** Site difference at 2352 bp.

In this study, QTL localization was performed using a population of RILs composed of HZ and Nekken2 hybrids based on germination rate statistics under salt treatment conditions, resulting in 16 QTL intervals associated with salt tolerance, of which six QTL loci overlapped with those described by previous studies. Singh et al. (2021) located QTL intervals associated with salt tolerance at the seedling and reproductive stages of rice by meta-QTL, *mQTL4.7* on chromosome 4 at 83.2 cM, *mQTL11.2* on chromosome 11 at 40.96 cM, and *mQTL12.3* on chromosome 12 at 60.06 cM are similar to the loci on chromosomes 4, 11, and 12 found in this study. Gimhani et al. (2016) showed *mQTL12.3* which was associated with the salinity survival index (SSI) and percent damaged shoots (PDS) under salt stress, was located at the uppermost 0.63 cM region of chromosome 10, and this partially overlaps with *qST4.2*, *qST11.4*, and *qST12.3* which were respectively localized to chromosomes 4, 11, and 12 in this study. The loci *qSSI10* and *qPDS10* associated with the SSI and PDS respectively, partially overlap with *qST10.1* on chromosome 10 in this study. Ziyan Xie (2020) used the BC_2F_7 backcross introgression population constructed from MingChuan 63 (salt-tolerant, recurrent parent) and 02428 (salt-sensitive) to characterize chromosome 12 *qSHL12.2*, *qSHW12.4*, and *qRW12.2* and metabolite-related *qM16-12.1*, *qM27-12*, *qM28-12*, *qM39-12*, *qM47-12.1*, *qM57-12*, *qM58-12.1*, *qM61-12*, *qM72-12*, *qM76-12.1*, *qM79-12.1*, *qM85-12*, and *qM87-12* intervals. Of these, *qM47-12.1* and *qM58-12.1* contributed 43.0% and 34.7%, respectively, and these partially overlap with the *qST12.2* interval localized on chromosome 12 in this study. Tiwari et al. (2016) used a population of RILs composed of CSR11/M148 to identify a salt tolerance gene encoding a rice MYB transcription factor within the *qSSIGY8.1* interval localized to chromosome 8, which is adjacent to the *qST8* interval localized to chromosome 8 in the present study.

Based on the previous studies, the present study has yielded several new loci for salt tolerance, including *qST3*, which is localized in the interval 28.90–30.16 cM on chromosome 3; *qST4.1* in the interval 73.40–75.45 cM on chromosome 4; *qST9* in the interval 92.88–94.02 cM on chromosome 9; *qST10.2*, in the interval 23.86–24.48 cM on chromosome 10, with an LOD value of 4.69, indicating

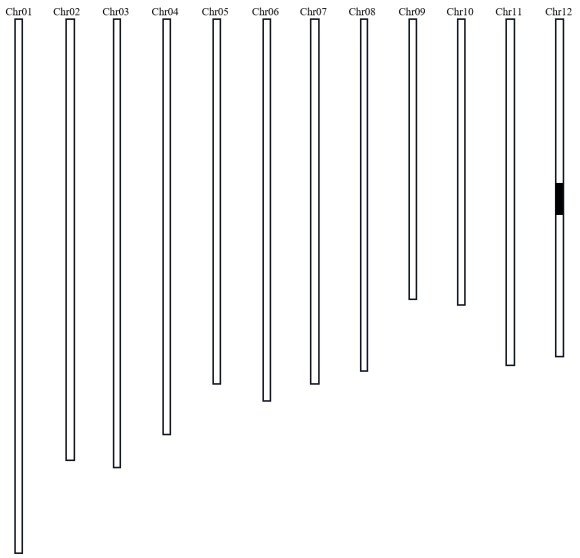


FIGURE 6
Graphical genotype of CSSL12 (a substitution line of chromosome 12). Black bar indicates the genome fragment from HZ; the other parts were from Nekken2.

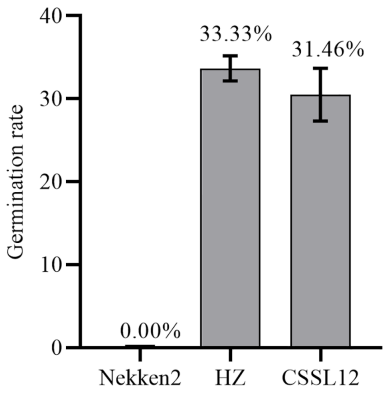


FIGURE 7
The germination rate for Nekken2, HZ and CSSL12.

10.3389/fpls.2022.1041081

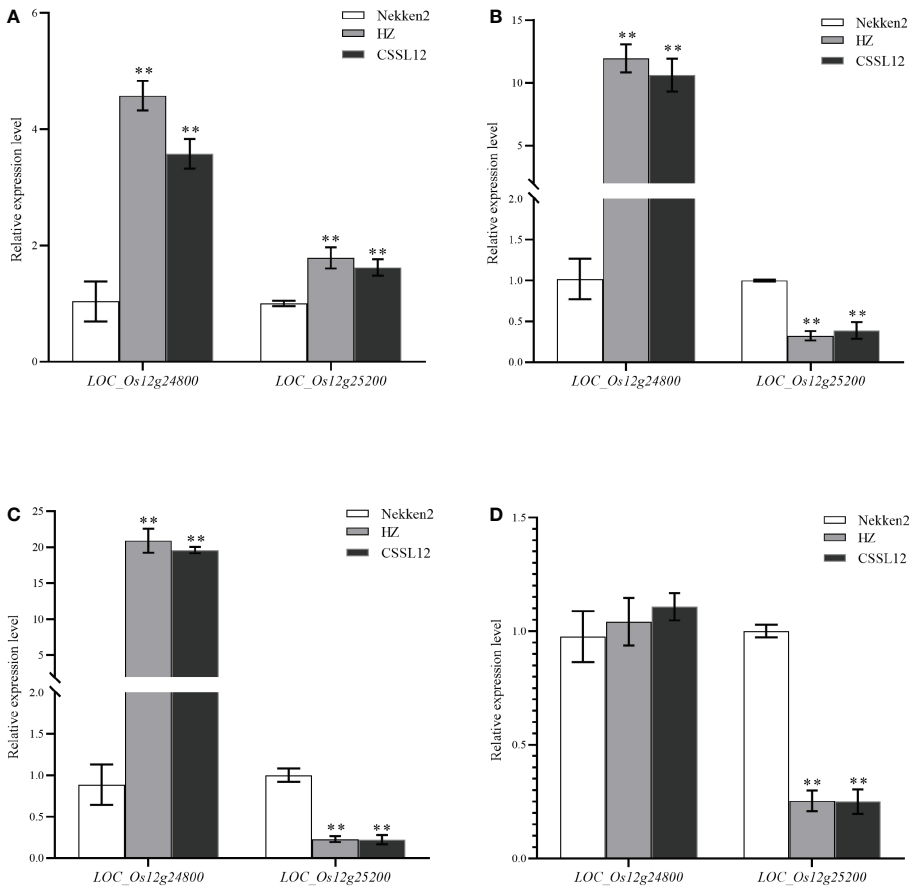


FIGURE 8
The qRT-PCR analysis of predicted genes in Nekken2, HZ and CSSL12. **(A)** Expression of candidate genes at 0 mM NaCl concentration. **(B)** Expression of candidate genes at 60 mM NaCl concentration. **(C)** Expression of candidate genes at 80 mM NaCl concentration. **(D)** Expression of candidate genes at 100 mM NaCl concentration. * indicates significance at $P \leq 0.05$ by Student's *t* test, ** indicates significance at $P \leq 0.01$ by Student's *t* test.

the likely presence of salt tolerance-related master genes; *qST11.1*, *qST11.2*, *qST11.3*, and *qST11.5* on chromosome 11; and *qST12.1* on chromosome 12 in the interval of 23.55–26.68 cM.

Analysis in conjunction with the Rice Gene Database (http://rice.plantbiology.msu.edu/) revealed that *LOC_Os12g25200*, a functional gene contained on QTL interval *qST12.3* encoding a chloride transporter protein of the chloride channel protein family overlaps with the interval locus reported by Singh et al. (Tian et al., 2011) The *LOC_Os12g25200* gene, which may be involved in osmotically regulating the transport of anionic chloride ions in rice and is associated with the regulation of salt tolerance in rice. Combined with qRT-PCR analysis of the expression of this gene in HZ and Nekken2 under different salt stress conditions, our results suggested that *LOC_Os12g25200* showed significant differences in expression between the two parents, and the expression of *LOC_Os12g25200* in HZ was significantly higher than that in Nekken2 under 0 mM NaCl stress. After treatment with 60, 80, or 100 mM NaCl stress gradients, the expression of *LOC_Os12g25200* decreased significantly in HZ and declined sequentially with increasing salt stress concentrations, reaching the lowest expression at 100 mM NaCl. Golldack et al. (2003) compared transcriptional changes in the salt-sensitive rice strain IR29 and the salt-tolerant strain Pokkali and identified the expression of *OsCLC1* in rice as a means to investigate Cl^- homeostasis under salt stress conditions. Analysis of the transcripts revealed that anion channels may also be involved in the uptake of anions by cells. The above findings indicate that *OsCLC1* plays an important role in the coordinated regulation of anion and cation homeostasis in rice under abiotic salt stress, suggesting an osmoregulatory function for *OsCLC1* in high salt stress environments. This has some similarities with the function of the protein encoded by *LOC_Os12g25200* found in our study, suggesting that *LOC_Os12g25200* may be involved in a similar mechanism in HZ, reduction in the expression of this gene improved the salt tolerance of rice. The above findings need to be further verified by subsequent relevant experiments.

To further verify the molecular mechanisms of *LOC_Os12g25200* in improving salt stress tolerance in rice, we sequenced the gene and analyzed changes in the gene sequence. Sequencing *LOC_Os12g25200* revealed four single-base differences between HZ and Nekken2 at exons 1073 bp, 1139 bp, 1959 bp, and 2352 bp, respectively, with "G to T" changes occurred at 1073 bp and 1139 bp, respectively. The amino acids encoded at positions 60 and 82 both showed "glycine to valine" changes, while the remaining two single-base differences did not cause amino acid changes. Meanwhile, we sequenced the promoter region of this gene and found that there were 8 base differences between HZ and Nekken2 at 733 bp, 937 bp, 1005–1006 bp, 1058 bp, 1119 bp, 1174 bp, 1605 bp and 1979 bp within the first 2200 bp of the start codon. It is speculated that these differences may lead to the difference in the expression of this gene in parents. The above sequencing results indicate that single-base differences the DNA sequences and

promoter sequences of the *LOC_Os12g25200* candidate gene indeed lead to changes in the encoded amino acids, and these differences may be important to differences in gene expression and thus the salt stress tolerance in HZ and Nekken2. Previous studies have indicated that Yin et al. (2022) measured the QTL interval localization related to root traits without NaCl treatment, and the QTL interval that they mapped did not overlap with the one in this study. These results indicated that the QTL interval mapped in this study was caused by the treatment with 80 mM NaCl, and further ruled out the difference in interval location caused by the genotype. This also laid the foundation for the subsequent cloning of salt tolerance genes.

In this study, *qST12.1*, *qST12.2* and *qST12.3* related to salinity tolerance at three seed bud stages were reported, which a potential *qST12.3* controlled GR. The QTL on chromosome 12 partially overlaps with results from previous studies (Livak and Schmittgen, 2001). Based on the QTL interval mapped by RIL population, candidate genes were screened out for real-time fluorescence quantitative analysis of their expression changes under salt stress. Furthermore, the CSSL12 substitution line was constructed based on *qST12.3*, and the results showed that it was a salt tolerance QTL controlling seed germination on chromosome 12 of rice. At the same time, we found that *LOC_Os12g25200* was a novel gene, and the expression of CSSL12 was different between the parents and the substitution line according to the gene transcription level. A candidate gene *LOC_Os12g25200* for a chloride transporter in a family of chloride channels was identified. Therefore, additional tests are needed to test whether *LOC_Os12g25200* is a major gene for *qST12.3* tolerance in seed germination.

The reaction mechanism of rice plants to salt stress is complex. Osmotic stress is one of the main stresses that plants face in the early stage of salt stress (Ahmed et al., 2015). Although some QTLs have been applied to high-yield rice varieties to improve salt tolerance, salt tolerance in rice is a complex trait controlled by multiple genes and affected by environmental factors. At present, there is no explanation for this complex regulatory mechanism, and improvement progress is slow. Therefore, the identification and cloning of major genes related to salt tolerance in rice will help to explain the unknown mechanism of salt tolerance and apply it to breeding.

Conclusion

Using a population of RILs composed of HZ and Nekken2 hybrids in this study, several new salt tolerance loci were located, among which a gene encoding a chloride transporter protein, with chloride channel family function, was found in the *qST12.3* interval, a locus that overlapsd with those described in previous studies. It was found that the expression of *LOC_Os12g25200* decreased significantly with increasing salt concentration and was less experessed in Nekken2 than HZ, by constructing the CSSL12 substitution line population, we further proved that there was a major gene controlling salt tolerance in the major effect interval of *qST12.3*, suggesting that this gene may play an important role in negatively regulating salt stress tolerance in rice and may be the main reason for the difference in salt stress tolerance between the two strains. The results of this experiment further uncovered QTL loci related to salt tolerance and confirmed the functions of genes related to salt stress tolerance in rice, the study provides a new theoretical basis and has application value for elucidating the mechanism of salt stress tolerance in rice. The results of this research may be applicable to practical rice molecular breeding and could provide key genetic loci that can be used to improve the ability of rice to tolerate salt stress. This can ultimately greatly icrease the land area that can be used for food production, greatly enhance the yield and quality of rice, and has important practical significance for the screening and breeding of excellent rice varieties.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

WY, TiL, ZC, TaL, and HY conceived and designed the experiment. YM, YL performed data collection. WY, ML, and XZ produced data analysis and figures. WY and TiL wrote the original manuscript with input from XY, YR and YW. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the National Key R&D Program of China (2021YFA1300703), the Zhejiang Provincial Natural Science Outstanding Youth Fund under (Grant No. LR20C130001), the National Natural Science Foundation of China under (Grant No. 31971921, U20A2030), the Zhejiang Provincial Ten Thousand People Plan for Young Top Talents under (Grant No. ZJWR0108023), and the National (202010345067; 202010345051).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1041081/full#supplementary-material

References

Ahmed, A. A. M., Roosens, N., Dewaele, E., Jacobs, M., and Angenon, G. (2015). Overexpression of a novel feedback-desensitized Δ^1-pyrroline-5-carboxylate synthetase increases proline accumulation and confers salt tolerance in transgenic *Nicotiana plumbaginifolia*. *Plant Cell Tiss Organ Cult.* 122, 383–393. doi: 10.1007/s11240-015-0776-5

Bewley, J. D. (1997). Seed germination and dormancy. *Plant Cell.* 9 (7), 1055–1066. doi: 10.1105/tpc.9.7.1055

Bharathkumar, S., Jena, P. P., Kumar, J., Baksh, S. K. Y., Samal, R., Gouda, G., et al. (2016). Identification of new alleles in salt tolerant rice germplasm lines through phenotypic and genotypic screening. *Int. J. Agric. Biol.* 18 (2), 441–448. doi: 10.17957/IJAB/15.0112

Ganie, S. A., Molla, K. A., Henry, R. J., Bhat, K. V., and Mondal, T. K. (2019). Advances in under-standing salt tolerance in rice. *Theor. Appl. Genet.* 132 (4), 851–870. doi: 10.1007/s00122-019-03301-8

Geilfus, C. M. (2018). Review on the significance of chlorine for crop yield and quality. *Plant Sci.* 270, 114–122. doi: 10.1016/j.plantsci.2018.02.014

Gimhani, D. R., Gregorio, G. B., Kottearachchi, N. S., and Samarasinghe, W. L. G. (2016). SNP-based discovery of salinity-tolerant QTLs in a bi-parental population of rice (*Oryza sativa*). *Mol. Genet. Genomics* 291 (6), 2081–2099. doi: 10.1007/s00438-016-1241-9

Golldack, D., Quigley, F., Michalowski, C. B., Kamasani, U. R., and Bohnert, H. J. (2003). Salinity stress-tolerant and -sensitive rice (*Oryza sativa* l.) regulate AKT1-type potassium channel transcripts differently. *Plant Mol. Biol.* 51 (1), 71–81. doi: 10.1023/a:1020763218045

Gomes-Filho, E., Lima, C. R. F. M., Costa, J. H., da Silva., A. C. M., da Guia Silva Lima, M., de Lacerda., C. F., et al. (2008). Cowpea ribonuclease: properties and effect of NaCl-salinity on its activation during seed germination and seedling establishment. *Plant Cell Rep.* 27 (1), 147–157. doi: 10.1007/s00299-007-0433-5

He, Y. Q., Yang, B., He, Y., Zhan, C. F., Cheng, Y. H., Zhang, J. H., et al. (2019). A quantitative trait locus, *qSE3*, promotes seed germination and seedling establishment under salinity stress in rice. *Plant J.* 97 (6), 1089–1104. doi: 10.1111/tpj.14181

Hu, T., Zhang, G. X., Zheng, F. C., and Cao, Y. (2018). Research progress in plant salt stress response. *Mol. Plant Breed.* 16 (9), 3006–3015. doi: 10.13271/j.mpb.016.003006

Joo, J., Lee, Y. H., and Song, S. I. (2014). Rice CatA, CatB, and CatC are involved in environmental stress response, root growth, and photorespiration, respectively. *J. Plant Biol.* 57 (6), 375–382. doi: 10.1007/s12374-014-0383-8

Liu, K., Zhu, J. W., Wan, B. J., Dai, J. Y., Tang, H. S., and Sun, M. F. (2021). Research progress on molecular genetics of rice salt tolerance. *J. Plant Genet. Resour.* 22 (4), 881–889. doi: 10.13430/j.cnki.jpgr.20201121002

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\triangle\triangle CT}$ method. *Methods* 25 (4), 402–408. doi: 10.1006/meth.2001.1262

Mccouch, S., Cho, Y. G., Yano, M., Paul, E., Blinstrub, M., Morishima, H., et al. (1997). Report on QTL nomenclature. *Rice Genet. Newsl.* 14, 11–13.

Meng, Y., Xu, Y. Z., Tang, A., Liu, Y. H., Dong, W. J., Wng, L. Z., et al. (2022). Identification and grading evaluation of saline-alkali tolerance of rice at gernination stage in songnen plain. *Heilongjiang Agric. Sci.* 8), 1–9. doi: 10.11942/j.issn1002-2767.2022.08.0001

Mishra, P., Bhoomika, K., and Dubey, R. S. (2013). Differential responses of antioxidative defense system to prolonged salinity stress in salt-tolerant and salt-sensitive indica rice (*Oryza sativa* l.) seedlings. *Protoplasma* 250 (1), 3–19. doi: 10.1007/s00709-011-0365-3

Munns, R., and Gilliham, M. (2015). Salinity tolerance of crops–what is the cost? *New Phytol.* 208 (3), 668–673. doi: 10.1111/nph.13519

Munns, R., and Tester, M. (2008). Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* 59, 651–681. doi: 10.1146/annurev.arplant.59.032607.092911

Nagamiya, K., Motohashi, T., Nakao, K., Prodhan, S. H., Hattori, E., Hirose, S., et al. (2007). Enhancement of salt tolerance in transgenic rice expressing an escherichia coli catalase gene, katE. *Plant Biotechnol. Resports.* 1 (1), 49–55. doi: 10.1007/s11816-007-0007-6

Oliver, S. N., Dennis, E. S., and Dolferus, R. (2007). ABA regulates apoplastic sugar transport and is a potential signal for cold-induced pollen sterility in rice. *Plant Cell Physiol.* 48 (9), 1319–1330. doi: 10.1093/pcp/pcm100

Penfield, S. (2017). Seed dormancy and germination. *Curr. Biol.* 27 (17), R874–R878. doi: 10.1016/j.cub.2017.05.050

Qi, D. L., Guo, G. Z., Lee, M. C., Zhang, J. G., Cao, G. L., Zhang, S. Y., et al. (2008). Identification of quantitative trait loci for the dead leaf rate and the seedling dead rate under alkaline stress in rice. *J. Genet. Genomics* 35 (5), 299–305. doi: 10.1016/S1673-8527(08)60043-0

Radanielson, A. M., Gaydon, D. S., Li, T., Angeles, O., and Roth, C. H. (2018). Modeling salinity effect on rice growth and grain yield with ORYZA v3 and APSIM-oryza. *Eur. J. Agron.* 100, 44–55. doi: 10.1016/j.eja.2018.01.015

Rao, Y. C., Lin, H., Xiao, S. Q., Wu, Y., Zhang, Y., and Wang, S. (2020). Identifying of QTL forresistance to submergence in rice. *J. Zhejiang Normal Univ. (Natural Sciences).* 43 (3), 312–319. doi: 10.16218/j.issn.1001-5051.2020.03.011

Razzaque, M. A., Talukder, N. M., Islam, M. T., and Dutta, R. K. (2011). Salinity effect on mineral nutrient distribution along roots and shoots of rice (*Oryza sativa* l.) genotypes differing in salt tolerance. *Arch. Agron. Soil Science.* 57 (1), 33–45. doi: 10.1080/03650340903207923

Shi, Y. Y., Gao, L. L., Wu, Z. C., Zhang, X. J., Wang, M. M., Zhang, C. S., et al. (2017). Genome-wide association study of salt tolerance at the seed germination stage in rice. *BMC Plant Biol.* 17 (1), 92. doi: 10.1186/s12870-017-1044-0

Singh, R. K., Kota, S., and Flowers, T. J. (2021). Salt tolerance in rice: seedling and reproductive stage QTL mapping come of age. *Theor. Appl. Genet.* 134 (11), 3495–3533. doi: 10.1007/s00122-021-03890-3

Sun, M., Zhang, P. J., Bai, Y. S., and Li, C. Q. (1992). The compatibility of Nekken1, Nekken2 with rice in China. *J. Anhui Agric. Sci.* 20 (3), 198–201. doi: 10.13989/j.cnki.0517-6611.1992.03.002

Tian, L., Tan, L. B., Liu, F. X., Cai, H. W., and Sun, C. Q. (2011). Identification of quantitative trait loci associated with salt tolerance at seedling stage from *Oryza rufipogon*. *J. Genet. Genomics* 38 (12), 593–601. doi: 10.1016/j.jgg.2011.11.005

Tiwari, S., SL, K., Kumar, V., Singh, B., Rao, A. R., Amitha Mithra, S. V., et al. (2016). Mapping QTLs for salt tolerance in rice (*Oryza sativa* l.) by bulked segregant analysis of recombinant inbred lines using 50K SNP chip. *PloS One* 11 (4), e0153610. doi: 10.1371/journal.pone.0153610

Wang, W. C., Lin, T. C., Kieber, J., and Tsai, Y. C. (2019). Response regulators 9 and 10 negatively regulate salinity tolerance in rice. *Plant Cell Physiol.* 60 (11), 2549–2563. doi: 10.1093/pcp/pcz149

Wang, Y. X., Shang, L. G., Yu, H., Zeng, L. J., Hu, J., Ni, S., et al. (2020). A strigolactone biosynthesis gene contributed to the green revolution in rice. *Mol. Plant* 13 (6), 923–932. doi: 10.1016/j.molp.2020.03.009

Wu, D. Z., Cai, S. G., Chen, M. X., Ye, L. Z., Chen, Z. H., and Zhang, H. T. (2013). Tissue metabolic responses to salt stress in wild and cultivated barley. *PloS One* 8 (1), e55431. doi: 10.1371/journal.pone.0055431

Xie, Z. Y. (2020). Identification QTL and candidate genes using multi-omics for salt tolerance in rice. *CAAS*, 000807. doi: 10.27630/d.cnki.gznky.2020.000807

Yang, Y. Q., and Guo, Y. (2018). Elucidating the molecular mechanisms mediating plant salt-stress responses. *New Phytol.* 217 (2), 523–539. doi: 10.1111/nph.14920

Yin, W. J., Chen, Z. G., Gao, P. H., Lu, T., Ye, H. F., Ye, R. L., et al. (2022). QTL mapping and expression analysis of candidate genes for root traits in rice. *J. Zhejiang Normal Univ. (Natural Sciences).* 45 (4), 419–426. doi: 10.16218/j.issn.1001-5051.2022.007

Yu, J., Zhao, W. G., Tong, W., He, Q., Yoon, M. Y., Li, F. P., et al. (2018). A genome-wide association study reveals candidate genes related to salt tolerance in rice (*Oryza sativa*) at the germination stage. *Int. J. Mol. Sci.* 19 (10), 3145. doi: 10.3390/ijms19103145

Zeng, P., Zhu, P. W., Qian, L. F., Qian, X. M., Mi, Y. X., Lin, Z. F., et al. (2021). Identification and fine mapping of *qGR6.2*, a novel locus controlling rice seed germination under salt stress. *BMC Plant Biol.* 21 (1), 36. doi: 10.1186/s12870-020-02820-7

Zhang, J. (2012). Effect of salt stress on seed germination of rice. *Seed* 31 (1), 98–100. doi: 10.16590/j.cnki.1001-4705.2012.01.057

Zhou, Y. B., Liu, C., Tang, D. Y., Yan, L., Wang, D., Yang, Y. Z., et al. (2018). The receptor-like cytoplasmic kinase STRK1 phosphorylates and activates CatC, thereby regulating H_2O_2 homeostasis and improving salt tolerance in rice. *Plant Cell.* 30 (5), 1100–1118. doi: 10.1105/tpc.17.01000

Zou, X., Liu, L., Hu, Z. B., Wang, X. K., Zhu, Y. C., Zhang, J. L., et al. (2021). Salt-induced inhibition of rice seminal root growth is mediated by ethylene-jasmonate interaction. *J. Exp. Bot.* 72 (15), 5656–5672. doi: 10.1093/jxb/erab206

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 24 March 2023
DOI 10.3389/fpls.2023.1152196

OPEN ACCESS

EDITED BY
Jinsong Bao,
Zhejiang University, China

REVIEWED BY
Xu Feifei,
Zhejiang University, China
Divya Balakrishnan,
Indian Institute of Rice Research (ICAR), India
Xiaoli Fan,
Chengdu Institute of Biology (CAS), China

*CORRESPONDENCE
Xiaojin Luo
luoxj@fudan.edu.cn
Ying Wang
wang_y@fudan.edu.cn

†These authors contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Plant Bioinformatics, a section of the journal Frontiers in Plant Science

RECEIVED 27 January 2023
ACCEPTED 13 March 2023
PUBLISHED 24 March 2023

CITATION
Lan D, Cao L, Liu M, Ma F, Yan P, Zhang X, Hu J, Niu F, He S, Cui J, Yuan X, Yang J, Wang Y and Luo X (2023) The identification and characterization of a plant height and grain length related gene *hfr131* in rice. *Front. Plant Sci.* 14:1152196. doi: 10.3389/fpls.2023.1152196

The identification and characterization of a plant height and grain length related gene *hfr131* in rice

Dengyong Lan[1,2†], Liming Cao[3†], Mingyu Liu[1], Fuying Ma[1], Peiwen Yan[1], Xinwei Zhang[1], Jian Hu[1], Fuan Niu[1,3], Shicong He[1], Jinhao Cui[1], Xinyu Yuan[1], Jinshui Yang[1], Ying Wang[1,2*] and Xiaojin Luo[1,4*]

[1]State Key Laboratory of Genetic Engineering and Engineering Research Center of Gene Technology (Ministry of Education), School of Life Sciences, Fudan University, Shanghai, China, [2]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Department of Ecology and Evolutionary Biology, School of Life Sciences, Fudan University, Shanghai, China, [3]Institute of Crop Breeding and Cultivation, Shanghai Academy of Agricultural Sciences, Shanghai, China, [4]Ministry of Education, Key Laboratory of Crop Physiology, Ecology and Genetic Breeding College of Agronomy, Jiangxi Agricultural University, Nanchang, China

Plant height and grain size are important agronomic traits affecting rice yield. Various plant hormones participate in the regulation of plant height and grain size in rice. However, how these hormones cooperate to regulate plant height and grain size is poorly understood. In this study, we identified a brassinosteroid-related gene, *hfr131*, from an introgression line constructed using *Oryza longistaminata*, that caused brassinosteroid insensitivity and reduced plant height and grain length in rice. Further study showed that *hfr131* is a new allele of *OsBRI1* with a single-nucleotide polymorphism (G to A) in the coding region, leading to a T988I conversion at a conserved site of the kinase domain. By combining yeast one-hybrid assays, chromatin immunoprecipitation-quantitative PCR and gene expression quantification, we demonstrated that OsARF17, an auxin response factor, could bind to the promoter region of *HFR131* and positively regulated *HFR131* expression, thereby regulating the plant height and grain length, and influencing brassinosteroid sensitivity. Haplotype analysis showed that the consociation of $OsAFR17^{Hap1}$/$HFR131^{Hap6}$ conferred an increase in grain length. Overall, this study identified *hfr131* as a new allele of *OsBRI1* that regulates plant height and grain length in rice, revealed that brassinosteroid and auxin might coordinate through *OsARF17–HFR131* interaction, and provided a potential breeding target for improvement of rice yield.

KEYWORDS

plant height, grain length, brassinosteroid, auxin, *OsARF17*, *OsBRI1*, rice

Introduction

Rice is among the most important food crops worldwide, providing staple food for more than half of the world's population (Qiao et al., 2021; Huang et al., 2022). However, the global population is growing rapidly and is forecast to attain 9.7 billion by 2050 and 10.4 billion by 2100 (Economic and Social Affairs, 2022). Furthermore, with ongoing economic development, the environment is deteriorating, and the arable land area is decreasing. Therefore, it is particularly important to increase the rice yield per unit land area to keep pace with the projected population growth (McClung, 2014).

Plant height and grain size are two important factors affecting rice yield. An excessively high or low plant height will affect rice yield. A plant that is too tall has poor lodging resistance, whereas if the plant is too short, the plant will produce smaller grains, an increased number of ineffective tillers, and have poor resistance to disease (Liu et al., 2018). Therefore, an appropriate plant height is essential to improve rice yield. This is comparable to the "Green Revolution" wave of semi-dwarf breeding in the 1960s, which doubled food production in much of the world (Hargrove and Cabanilla, 1979; Khush, 1999). The yield per plant of rice is directly affected by three factors: panicle number, number of grains per panicle, and thousand-grain weight. The thousand-grain weight is positively associated with grain size, including grain length, grain width, grain thickness, and degree of grain filling (Xing and Zhang, 2010). Thus, grain size is among the most agronomically important traits in rice breeding.

Plant hormones play crucial roles in the regulation of plant height and grain size in rice, among which brassinosteroids (BRs) are particularly important. Similar in structure to mammalian steroid hormones, BRs are a class of polyhydroxysteroid plant hormones and are involved in multiple biological processes in growth and development (Clouse, 2011). Many BR-related genes have been reported in rice, and most affect both plant height and grain size. The BR synthesis-associated mutants *brd2*, *d11*, *d2*, and *brd1* have a reduced plant height and grain size to different degrees (Mori et al., 2002; Hong et al., 2003; Hong et al., 2005; Tanabe et al., 2005). *OsBRI1*, the rice ortholog of *Arabidopsis BRI1*, encodes the receptor of BR in rice. Loss-of-function of *OsBRI1* results in dwarf culms, reduced grain size, erect leaves, and insensitivity to BR (Yamamuro et al., 2000; Morinaka et al., 2006). OsBAK1 interacts with OsBRI1 to promote BR signal transduction downstream. Overexpression of *OsBAK1* increases sensitivity to BR, however, reduces plant height and grain size, which may be caused by repressing GA level (Li et al., 2009; Tong et al., 2014). *OsGSK1–OsGSK4*, which are homologs of *AtBIN2*, belong to the GSK3-like kinase family and play an important role in the regulation of BR signaling (Yoo et al., 2006). Loss-of-function of one or several *OsGSK* genes has revealed that the *gsk2*, *gsk1,2*, *gsk2,3*, and *gsk2,4* mutants show increased plant height, whereas the *gsk2,3,4* mutant has decreased plant height and that of the *gsk1,2,3,4* mutant is not changed significantly. However, the grain length and thousand-grain weight of all types of mutants is significantly increased (Liu et al., 2021). The mutant or knockout lines of the BR-associated transcription factors *OsBZR1*, *OsLIC*, and *DLT*/*GS6* all show significant dwarfing; however, the grain size and yield of the mutant or knockout lines of *OsBZR1* and *OsLIC* are significantly decreased, whereas those of the *gs6* mutant are significantly increased (Bai et al., 2007; Wang et al., 2008; Tong et al., 2009; Sun et al., 2013; Zhu et al., 2015). Other BR regulators, such as *OsBU1* and *SG1*, are also involved in the regulation of plant height and grain shape (Tanaka et al., 2009; Nakagawa et al., 2011). BR plays an important role in the regulation of multiple phenotypes in rice, such as plant height, grain size, seed germination, etc, involved in complex molecular regulatory networks (Xiao et al., 2020; Niu et al., 2022; Xiong et al., 2022), which is worth of further study. In addition, multiple hormones regulate the plant height and grain size of rice, but how BR coordinates with other hormones requires further study.

In this study, we identified a gene, *HFR131*, that regulates plant height and grain length in rice. Further study of the *hfr131* mutant demonstrated that *hfr131* is a new allele of *OsBRI1*. The replacement of amino acids in the kinase domain of HFR131 caused dwarfing and reduced grain length in the near-isogenic line NIL-*hfr131*. The auxin response factor OsARF17 was found to regulate plant height and grain length through BR signaling by binding to the promoter of *HFR131* and thereby regulating its expression. In addition, the consociation of $OsAFR17^{Hap1}$/$HFR131^{Hap6}$ was confirmed to increase the grain length.

Materials and methods

Plant materials and growth conditions

Using wild rice *Oryza longistaminata* as the donor parent (♀) and *indica* restorer rice '187R', cultivated by Guangxi Academy of Agricultural Sciences, as the recipient parent (♂), the BC_4F_1 population was obtained through hybridization and backcrossing with 187R (♂) for four successive generations and the BC_4F_4 population was obtained by selfing the BC_4F_1 population for three successive generations. Then we identified the genotype of BC_4F_4 population with genetic markers and a set of introgression lines were constructed (Figure 1A). Among these lines, by phenotype identification, we identified an introgression line, IL-*hfr131*, that exhibited significantly reduced plant height, small grain size and erect leaf. The alleles in IL-*hfr131* and 187R were therefore designated *hfr131* and *HFR131*, respectively. Subsequently, we backcrossed IL-*hfr131* (♀) with its recurrent parent 187R (♂) for four successive generations and selfed for one generation, then identified the genotype of the offspring with genetic markers, obtaining a near-isogenic line, NIL-*hfr131*, with an introgression fragment on the long arm of chromosome 1 (Supplementary Figure 1). Next, NIL-*hfr131* was used to construct a F_2 population comprising approximately 4000 plants for fine-mapping and phenotypic analysis (Figure 1A). The plant height (n = 13), leaf angle (n = 10), internode length (n = 13) and panicle length (n = 13)

of 187R and NIL-*hfr131* were measured. For grain length, grain width and grain weight, each variable was measured on three groups with 50 seeds per group.

All rice materials in this study were planted in Shanghai (31°20′ 26″ N, 121°30′26″ E) in summer and in Hainan (18°18′52″ N, 109° 03′05″ E) in winter. Standard field management practices were applied for all lines.

Histological analysis

The second internode of rice stems was sampled and fixed with 50% FAA solution (50 mL anhydrous ethanol, 10 mL 37% formaldehyde, 5 mL glacial acetic acid, 35 mL ddH_2O). Histological sections of the samples were prepared by the Wuhan Misp Biotechnology Co., Ltd. (Wuhan, China). The internode cell length of the left, middle, and right side of the section were measured and three layers of cells were measured in each region.

BR treatment

Rice seeds were cultured in 0.5% agar semi-solid medium containing 0 or 1 μM/L 24-epibrassinolide (24-eBL) (Zhang et al., 2015), which is commonly used to test the responsiveness of plants to BR, for 1 week in a greenhouse at 26 °C under shading. The coleoptile length and root length were then measured (n = 10). The change in root length was calculated as a measure of the degree of inhibition of root elongation (n = 10) using the following formula: (root length with 1 μM/L 24-eBL treatment − root length without 24-eBL treatment)/root length without 24-eBL treatment.

Map-based cloning

Total DNA was extracted from leaf samples. The corresponding markers were amplified by PCR, and the PCR products were

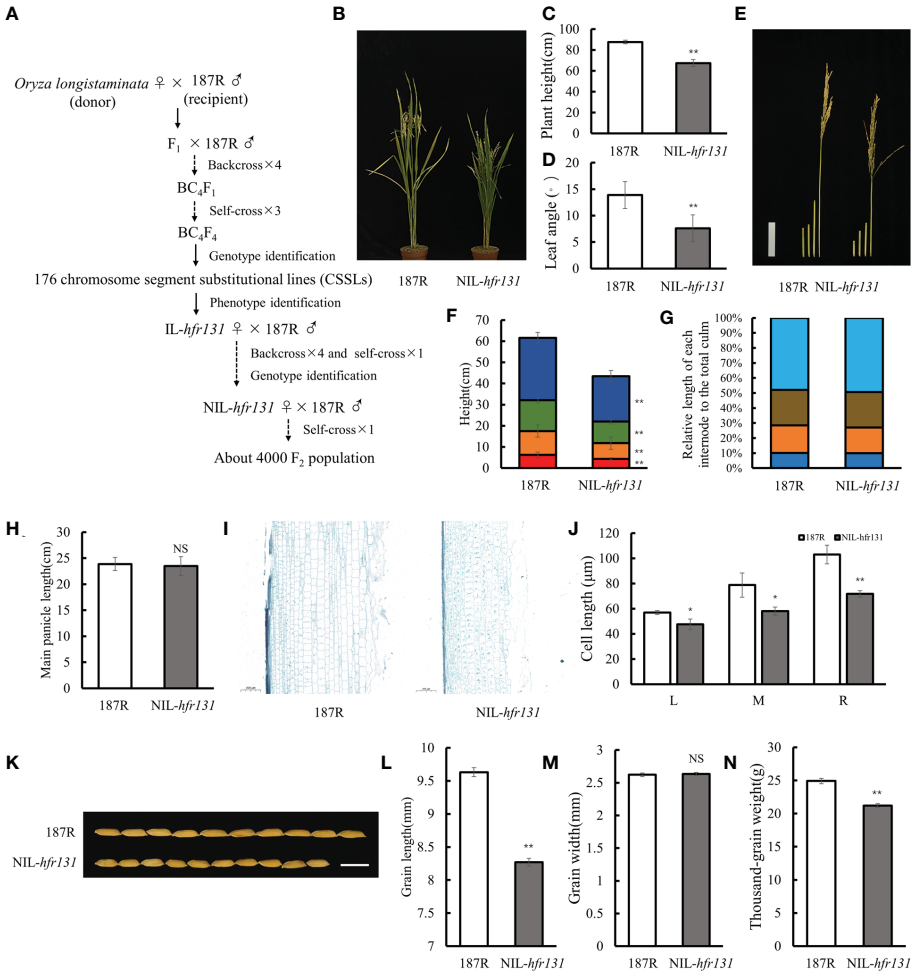


FIGURE 1
Phenotypes of rice '187R' and NIL-*hfr131*. **(A)** Construction of infiltrating lines and F_2 population. **(B–D)** Plant architecture, plant height, and the third leaf angle of 187R and NIL-*hfr131*. **(E–G)** Internode morphology (bar = 10 cm), internode length, and internode elongation pattern of 187R and NIL-*hfr131*. **(H)** Panicle length of 187R and NIL-*hfr131*. **(I, J)** Longitudinal sections of the stem and internode cell length (L, M, and R refer to the left, middle, and right side of the section, respectively) of 187R and NIL-*hfr131* (bar = 200 μm). **(K–N)** Grain morphology (bar = 1 cm), grain length, grain width, and thousand-grain weight of 187R and NIL-*hfr131*. * P < 0.05, ** P < 0.01, NS non-significant.

separated by polyacrylamide gel vertical electrophoresis. The primers used for fine mapping are listed in Supplementary Table 3.

RNA isolation and quantitative real-time PCR

Total RNA from rice samples was extracted using the FastPure® Plant Total RNA Isolation Kit (Vazyme, Nanjing, China). First-strand cDNA was generated by reverse transcription using the PrimeScript™ RT Reagent Kit (Takara, Otsu, Japan). Quantitative real-time PCR (qPCR) was conducted using TB Green Premix Ex Taq II (Tli RNaseH Plus) (Takara) and a CFX96 Real-Time PCR Detection system (Bio-Rad, Hercules, CA, USA). The reactions were conducted with 40 cycles of 95°C for 5 s and 60°C for 30 s. Melting curves were prepared and analyzed. *OsActin* or *OsUBQ5* was used as an internal reference gene. Three biological replicates were performed for each sample. The primers used are listed in Supplementary Table 3.

Dual-luciferase reporter assay

The promoter fragment of *HFR131* in the same region of 187R (3500 bp) and NIL-*hfr131* (3067 bp) was cloned and inserted into the *pGreenII 0800-LUC* vector using seamless cloning technology to construct the recombinant vectors *proHFR131-LUC* and *prohfr131-LUC*. Seeds of rice '9311' were germinated on wet napkins at 37 °C for 14 days. The stem of the seedlings was cut into thin slices and protoplasts were prepared by hydrolyzing the cell walls with cellulase and macerozyme. The recombinant vectors were transferred into the '9311' protoplasts using 40% polyethylene glycol solution and cultured at 28 °C for 12–16 h. The protoplast lysate was collected and the promoter activity of the *HFR131* promoter was tested using the Dual-Luciferase® Reporter Assay System (Promega, Madison, WI, USA) and a Synergy™ 2 Multi-Mode Reader (BioTek, Winooski, VT, USA). The activation efficiency of luciferase was expressed as the ratio of firefly luciferase (LUC) to *Renilla* luciferase (REN); the higher the ratio, the stronger the promoter activity. Three technical replicates were performed for each sample. The primers used are listed in Supplementary Table 3.

Transcriptome analysis

Total RNA was extracted from young panicles (n = 3) before heading date of 187R and NIL-*hfr131* using the FastPure® Plant Total RNA Isolation Kit (Vazyme). The extracted RNA was used as the template for construction of cDNA libraries, which were sequenced using a MGI MGISEQ-2000 platform by Sangon Biotech (Shanghai, China). The relative gene expression level was calculated and normalized to fragments per kilobase of transcript per million mapped reads. Differentially expressed genes (DEGs) were annotated and classified according to their putative or proven function. The transcriptome datasets were submitted to the GEO database as accession nos. GSE224118.

Construction of transformation lines

Overexpression lines of *HFR131* were based on the NIL-*hfr131* mutant background. For the coding sequences (CDSs) of *OsARF17* and the promoter sequences of *HFR131* in 'ZH11' and 187R were consistent, respectively, and ZH11 had high seed setting rate, which helped us investigate the phenotype of the grains well, we construct *OsARF17* knockout and overexpression lines based on ZH11 background. CRISPR/Cas9 knockout lines of *OsARF17* were generated by BIOGLE GeneTech (Jiangsu, China), and the target sequence was 5′- CCAATTATCCTAACTTGCCTCCACAAC TTATTTGCCAACTTCACAATGTGACG-3′. All knockout lines of the target sequence were cloned by PCR and sequenced. Regarding the overexpression lines, the CDSs of *OsARF17* (LOC_Os06g46410.1) and *HFR131* (LOC_Os01g52050.1) were cloned and inserted into the *PCHF-FLAG* vector to construct the overexpression vectors *CaMV35-OsARF17-FLAG* and *CaMV35S-HFR131-FLAG* using a seamless cloning technique. *CaMV35S-OsARF17-FLAG* was transformed by BIOGLE GeneTech and *CaMV35S-HFR131-FLAG* was transformed by BioRun (Wuhan, China). All transformed lines were verified by qPCR, and eight biological replicates were performed for *OsARF17* overexpression lines (T_2 generation) and three technical replicates were performed for *HFR131* overexpression lines (T_0 generation). The primers used are listed in Supplementary Table 3. The plant height, grain length and grain width of transformation lines were measured. Ten biological replicates were performed for T_3 generation of *OsARF17* knockout lines and T_2 generation of *OsARF17* overexpression lines. For the grain length and thousand-grain weight of T_0 generation of *HFR131* overexpression lines, each variable was measured on three groups with 20 seeds per group.

Yeast one-hybrid assay

Eight 150–300 bp DNA fragments containing an auxin response element (AuxRE) in the 3500 bp promoter region of *HFR131* were cloned and inserted into the *pAbAi* vector. The recombinant vectors were transferred into the yeast strain Y1H Gold, and cultured and screened on SD−Ura solid medium. The positive clones were detected by colony PCR using the primers Y1-test and pAbAi (Supplementary Table 3). The positive colonies were diluted and cultured in SD−Ura solid medium containing aureobasidin A (ABA^r) at different concentrations ranging from 0 to 1000 ng/ml. The culture with the optimal concentration of ABA^r was selected to eliminate self-activation. The CDS of *OsARF17* was cloned by PCR and inserted into the *pGADT7* vector to construct the *AD-ARF17* recombinant vector, which was transformed into yeast Y1H Gold containing one of the DNA fragments mentioned above. The transformed Y1H Gold cells were cultured on SD−Leu solid medium containing the selected ABA^r concentration. The empty *pGADT7* vector, *KAD*, was transformed as a negative control. The

Yeastmaker Yeast Transformation System 2 (Clontech) was used following the manufacturer's instructions. The primers used are listed in Supplementary Table 3.

ChIP-qPCR

Approximately 2.5 g rice seeds were weighed, soaked in water at 37 °C for 24 h, and then sown in nutrient soil 48 h after germination. The seeds were incubated at 26 °C for 14 days and then used for chromatin immunoprecipitation–quantitative real-time PCR (ChIP-qPCR). The experimental procedure followed the method of Wang et al. (Wang et al., 2018). The antibody used in this study was the FLAG antibody and the magnetic beads used were protein A/G magnetic beads. *OsActin* was used as an internal standard for data normalization. Three technical replicates were performed for each sample. The primers used are listed in Supplementary Table 3.

Haplotype analysis

A haplotype network for *HFR131* and *OsARF17* based on pairwise differences between two adjacent gene–CDS–haplotypes (gcHaps) was constructed using the pegas R package. The gcHaps data used for building the network were downloaded from the Rice Functional Genomics and Breeding (RFGB) database (https://www.rmbreeding.cn). The 3K Phenotype data were downloaded from RFGB (https://www.rmbreeding.cn/Phenotype). One-way analysis of variance (ANOVA) of the gcHap data was conducted for three agronomic traits (plant height, grain length, and thousand-grain weight), followed by Duncan's multiple-range test for multiple comparisons using the R package agricolae (v.1.3-5). Two-way ANOVA of the gcHaps data was conducted for analysis of grain length, followed by Duncan's multiple-range test for multiple comparisons using the R package agricolae (v.1.3-5). The probability level $P < 0.05$ was considered to be statistically significant. The multiple comparisons and the distribution of gcHaps among the five major populations, *Aus*, *Bas*, *XI*, *GJ* and *Adm*, were visualized using the R package ggplot2 (v.3.4.0) as boxplots and stacked barplots. The single-nucleotide polymorphism (SNP) dataset used for calculation of nucleotide diversity (π) and Tajima's *D* was downloaded from the Rice SNP-Seek Database (https://snp-seek.irri.org/_download.zul). Nucleotide diversity (π) at each window in the target region was estimated using VCFtools (v.0.1.15) with the "–window-pi" and "–window-pi-step" options, and the average value was used as its *p*-value. Tajima's *D* for the target region was calculated using VCFtools (v.0.1.15) with the "–TajimaD" option.

Results

The near-isogenic line NIL-*hfr131* exhibited reduced plant height and grain length

Compared with cultivated rice, wild rice has a larger gene pool, which allows adaptation to different environments and is a valuable resource for gene functional research in rice. In this study, we identified an introgression line, IL-*hfr131*, with significantly reduced plant height and grain size using wild rice *Oryza longistaminata* as the donor parent (♀) and *indica* restorer rice 187R as the recipient parent (♂) (Figure 1A), and further constructed a near-isogenic line, NIL-*hfr131*, containing the gene *hfr131* (Supplementary Figure 1). Subsequently, a F_2 population containing approximately 4000 plants was constructed using NIL-*hfr131* (♀) and 187R (♂) (Figure 1A).

Phenotypic investigation showed that the ratio of dwarf and small-grained plants to high and big-grained plants in the F_2 population was close to 1:3, indicating that *hfr131* was a recessive gene. Compared with 187R, NIL-*hfr131* showed a dwarf plant height accompanied by erect leaves (Figures 1B–D). To determine the direct cause of the plant height reduction, we measured the length of the first four internodes (counting the internode directly connected to the panicle as the first internode) and the panicle length of 187R and NIL-*hfr131*. The results showed that each internode of NIL-*hfr131* was significantly shortened to equal proportions (Figures 1E–G). However, there was no significant difference in panicle length between 187R and NIL-*hfr131* (Figure 1H), indicating that the reduced plant height of NIL-*hfr131* was caused by the equally proportional shortening of the internodes. The length of the internode is directly related to the growth and development of the internode cells. To further clarify the mechanism of internode shortening of NIL-*hfr131*, longitudinal sections of the internodes of NIL-*hfr131* and 187R were prepared. The results showed that the stem cell length of NIL-*hfr131* was significantly shortened both externally and internally (Figures 1I, J), indicating that the reduced internode length of NIL-*hfr131* was mainly caused by shortening of the cell length in the stem. Although the panicle length of NIL-*hfr131* was not changed significantly, the grain size was significantly decreased. Compared with 187R, the grain length and thousand-grain weight of NIL-*hfr131* were significantly decreased, but there was no significant difference in grain width (Figures 1K–N). Taken together, these results showed that the reduced plant height and grain size of NIL-*hfr131* was caused by shortening of the internodes and grain length, respectively.

NIL-*hfr131* is insensitive to brassinosteroid

Given that NIL-*hfr131* plants were reduced in height and had erect leaves, which is similar to the mutant phenotype of genes associated with the BR synthesis and signaling pathways (Hong et al., 2005; Nakamura et al., 2006; Tong et al., 2009), we speculated that *HFR131* may be involved in the BR synthesis or signaling pathways. To test this hypothesis, we conducted a BR treatment assay. The results showed that the coleoptile length of 187R was significantly increased after 24-eBL treatment, whereas NIL-*hfr131* showed no significant change (Figures 2A, B). Furthermore, after 24-eBL treatment, the root length of 187R was extremely reduced (Figure 2A). Although the root of NIL-*hfr131* also showed significant shortening, the inhibitory effect of 24-eBL on NIL-*hfr131* was weaker than on 187R (Figures 2A–C). In conclusion,

10.3389/fpls.2023.1152196

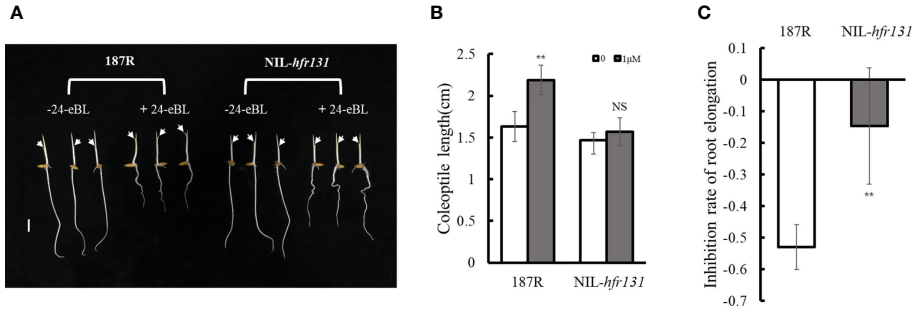


FIGURE 2
Responses of rice '187R' and NIL-*hfr131* to 24-epibrassinolide (24-eBL). **(A)** Effect of 24-eBL on coleoptile and root elongation in seedlings. The arrows indicate the top of the coleoptile. Bar = 1 cm. **(B)** Coleoptile length of 187R and NIL-*hfr131* seedlings in the presence or absence of 1 μM/L 24-eBL. **(C)** Inhibition of root elongation of 187R and NIL-*hfr131* in response to 1 μM/L 24-eBL. ** $P < 0.01$, NS non-significant.

NIL-*hfr131* was insensitive to BR, suggesting that *HFR131* was involved in the BR signaling pathway rather than the BR synthesis pathway.

hfr131 is a new allele of *OsBRI1*

To map the location of the *HFR131* locus, we constructed two pools comprising 36 dwarf small-grained plants (DS; recessive plants) and 36 tall large-grained plants (HB; dominant plants), selected from the F_2 population, and analyzed the two pools using simple sequence repeat (SSR) markers distributed on chromosome 1. The markers RM1297 and RM5931 showed polymorphism between the pools, and exhibited markedly different banding patterns between recessive individuals and dominant individuals (Supplementary Figure 2), indicating the two markers were closely linked to *HFR131*. Combining these two markers and neighboring SSR markers for further analysis, we preliminarily narrowed the effective locus to the interval between markers RM5931 and RM1152 (Figure 3A).

To fine map *HFR131*, we used RM5931 and RM1152 to screen 926 recessive plants selected from the F_2 population of almost 4000 plants, and 193 recombinants were identified. The SSR marker RM7514, located between RM5931 and RM1152, was used to analyze the 193 recombinants. The *HFR131* locus was located in the region between RM7514 and RM1152. In addition, Indel 64 was designed based on sequence differences between 187R and NIL-*hfr131*, and was found to co-segregate with NIL-*hfr131* (Figure 3B), indicating that *HFR131* was located at or adjacent to the site of Indel 64. From a BLAST search of the National Center for Biotechnology Information databases (https://www.ncbi.nlm.nih.gov/), we found that Indel 64 was located in the promoter region of *OsBRI1*, implying that *OsBRI1* may be the candidate gene for *HFR131*.

We then sequenced the promoter and coding regions of *OsBRI1* of 187R and NIL-*hfr131*. Two main mutant sites were found. One mutation was a 433 bp deletion in the promoter region, and the other was a SNP (G to A transition) in the coding region, converting threonine to isoleucine at position 988 in the kinase domain of OsBRI1 (Figure 3C). These results indicated that *HFR131* may be *OsBRI1*.

To investigate whether *OsBRI1* correspond to the phenotypic change in NIL-*hfr131*, we overexpressed the wild-type *OsBRI1*, cloned from 187R, in NIL-*hfr131* (Figure 3D). The results showed that, compared with NIL-*hfr131*, the plant height, leaf angle, grain length, and thousand-grain weight of the overexpression lines (*HFR131OE1* and *HFR131OE2*) were significantly increased (Figures 3E, F), indicating that *HFR131* is *OsBRI1* and that *hfr131* is a new allele of *OsBRI1*.

Amino acid non-synonymous substitution in the HFR131 kinase domain resulted in the reduction of plant height and grain length

Given that there were two mutation sites in *hfr131*, to determine which of the sites (the SNP in the coding region or the 433 bp deletion in the promoter region) was responsible for the height reduction and short grain length of NIL-*hfr131*, we first detected the expression levels of *HFR131* in the stem and leaf of 187R and NIL-*hfr131* by qPCR. Deletion of large fragments in the promoter region often leads to a significant reduction in the gene transcript level, resulting in abnormal gene function. Interestingly, compared with 187R, expression levels of *HFR131* in the stem and leaf of NIL-*hfr131* were not decreased, but in both were up-regulated significantly (Figure 4A). To further verify whether the 433 bp deletion in the *HFR131* promoter enhanced the promoter activity, we conducted a dual-luciferase reporter assay. We constructed two recombinant vectors, *proHFR131-LUC* and *prohfr131-LUC*, and transferred them into rice protoplasts. The empty vector was transferred as a negative control. The results showed that the luciferase activation efficiency of the *prohfr131-LUC* construct was significantly higher than that of the *proHFR131-LUC* construct (Figure 4B). This result suggested that the 433 bp deletion in the *HFR131* promoter region did enhance the promoter activity.

To determine which mutation site was the direct factor causing the phenotypic effects of NIL-*hfr131*, we considered the following three points. First, *hfr131* is a recessive gene. Generally, an increase in the expression of a recessive gene does not cause phenotypic effects. Second, we compared the amino acid sequence of HFR131 in NIL-

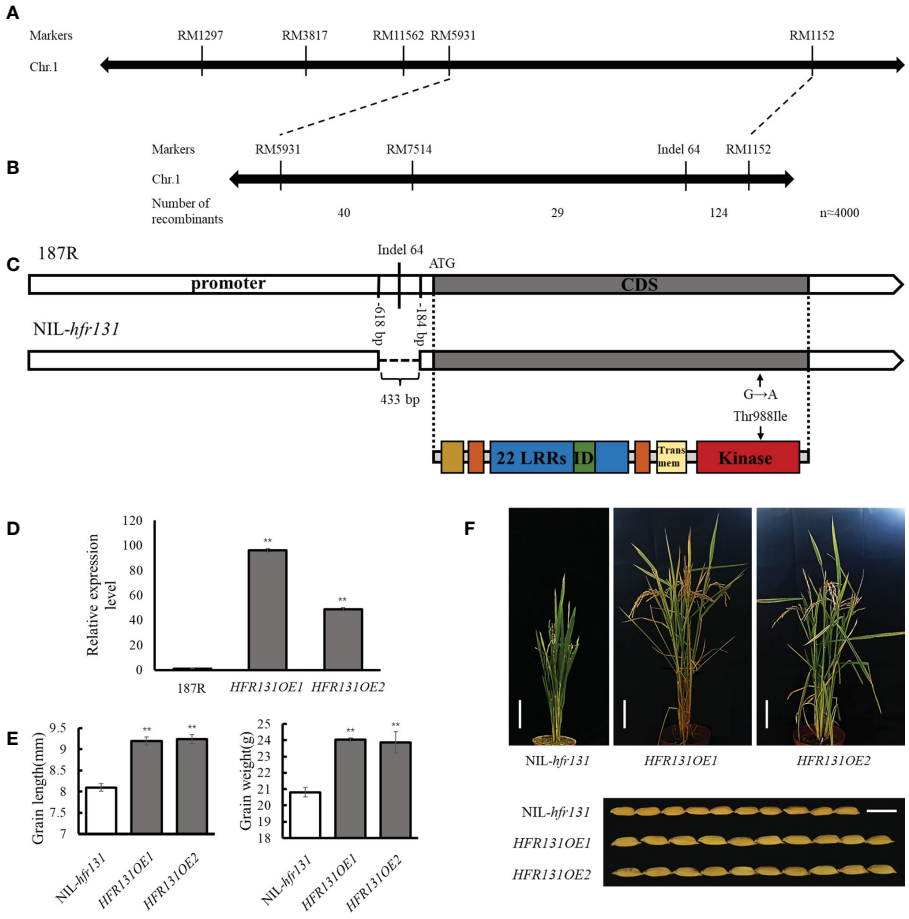


FIGURE 3
Identification of *HFR131*. **(A)** Linkage map of *HFR131*. **(B)** Fine mapping of *HFR131* using almost 4000 F_2 plants. The number of recombinants between adjacent markers is indicated under the linkage map. **(C)** Gene structure of *OsBRI1* in rice '187R' and NIL-*hfr131*, and the corresponding protein structure redrawn from Nakamura et al. (Nakamura et al., 2006). Positions of the 433 bp deletion are presented by base number before the ATG. **(D)** Transgene expression level of *OsBRI1*/*HFR131*. *OsActin* was used as an internal reference. **(E)** Grain length and thousand-grain weight of NIL-*hfr131* and *HFR131* overexpression lines (T_0 generation). **(F)** Plant architecture (Bar = 10 cm) and grain morphology (Bar = 1 cm) of NIL-*hfr131* and *HFR131* overexpression lines (T_0 generation). The plant height of NIL-*hfr131*, *HFR131OE1* and *HFR131OE2* is 50.9 cm, 68 cm and 66 cm, respectively. ** $P < 0.01$.

hfr131 with homologous proteins of *Arabidopsis*, rice, tomato, pea, and barley, and found that the threonine at position 988 of HFR131 is highly conserved (Figure 4C). The mutation site is located in the kinase domain of HFR131, a domain critical to BR signal transduction. However, the amino acid at position 988 of HFR131 was changed to isoleucine in NIL-*hfr131*, a residue with a quite different structure and physical properties to threonine, implying that this site was the target site causing the phenotypic change of NIL-*hfr131*. Third, the overexpression of *HFR131* in NIL-*hfr131* resulted in a significant increase in plant height and grain length (Figures 3E, F), indicating that the increased expression level caused by the deletion of 433 dp in the promoter region of *HFR131* was not the direct cause of the decrease in plant height and grain length of NIL-*hfr131*. In conclusion, the amino acid non-synonymous substitution in the kinase domain of HFR131 was the effective site that led to the decrease in plant height and grain length.

To determine how *hfr131* affects plant height and grain length, we selected young panicles before heading date of NIL-*hfr131* and 187R for transcriptome analysis and 2444 differentially expressed genes (DEGs) were detected. Among the DEGs, we found that the expression levels of *D2*, *D11*, and *BRD1*, which are associated with BR synthesis, were significantly increased (Supplementary Table 1), which was consistent with the hypothesis of Yamamuro et al. (2000) that plants could compensate for the defects caused by decreased BR sensitivity by increasing BR synthesis. We also observed that genes associated with plant height and grain size showed a significant change in expression level. The positive regulators of grain length *GS2* (Hu et al., 2015) and *OsLG3* (Yu et al., 2017) were significantly down-regulated. The negative regulators of plant height and grain length *OML4* (Lyu et al., 2020) and *OsREM4.1* (Gui et al., 2016) were significantly up-regulated. The expression level of *OsAP2-39* (Yaish et al., 2010) and *OsCYP96B4* (Ramamoorthy et al., 2011), elevating whose expression level would result in dwarf plant height, was significantly upregulated, whereas the expression of *LHD2* (Xiong et al., 2006), a positive regulator of plant height, was significantly decreased (Supplementary Table 1). These genes may be the response genes downstream of *HFR131* involved in regulating plant height and grain length.

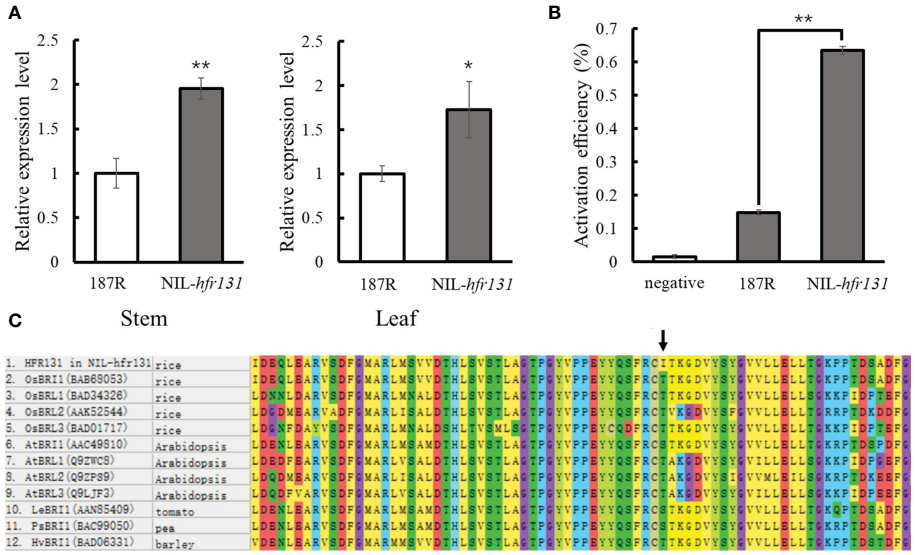


FIGURE 4
Analysis of the effective mutation site. **(A)** Expression level of *HFR131* in the stem and leaf of 187R and NIL-*hfr131*. **(B)** Detection of promoter activity. Activation efficiency was calculated as LUC/REN. **(C)** Alignment of a portion of the kinase domain of *Arabidopsis thaliana*, rice, tomato, pea, and barley BRI family members. The arrow indicates the mutation site. * $P < 0.05$, ** $P < 0.01$.

In addition, hormone-related genes were differentially expressed; for example, the gibberellin (GA)-related gene *OsGA2ox10* (Lo et al., 2008) was significantly up-regulated and the auxin-related genes *OsPIN1d*, *OsPIN8*, and *OsAUX2* (Wang et al., 2009; Yu et al., 2015) were significantly down-regulated (Supplementary Table 1), indicating that *HFR131* regulated plant height and grain length by coordinating with other hormones.

OsARF17 is involved in the regulation of plant height and grain length by the BR signaling pathway through regulating the expression of *HFR131*

OsARF17 plays an important role in the regulation of rice architecture, such as leaf angle and tiller angle (Chen et al., 2018; Li et al., 2020; Huang et al., 2021), suggesting that *OsARF17* may be an important target gene in rice breeding. However, whether *OsARF17* is also involved in the regulation of plant height and grain size, two important agronomic traits, has not been reported. It has been reported that *OsARF11* and *OsARF19* can regulate the expression of *OsBRI1* (Sakamoto et al., 2013; Zhang et al., 2015), so we speculated that *OsARF17* may regulate plant height and grain size through *HFR131*.

We constructed two *OsARF17* knockout (*KO1* and *KO2*) and overexpression lines (*OE1* and *OE2*), respectively (Figures 5A, B). Phenotypic observation showed that *OsARF17* knockout lines were significantly reduced in plant height compared with ZH11, but *OsARF17* overexpression lines did not show the opposite phenotype (Figures 5C, D), which may be related to the interaction of OsARF17 with other proteins in the ARF family, such as OsARF19, whose overexpression lines showed a reduced plant height (Zhang et al., 2015). Whatever, this result showed that *OsARF17* was necessary for the maintenance of plant height in rice. Compared with ZH11, the *OsARF17* knockout lines had significantly shorter grains and decreased thousand-grain weight, whereas *OsARF17* overexpression lines had significantly longer grains and increased thousand-grain weight (Figures 5E–G). These results suggested that *OsARF17* affected both plant height and grain length, which was similar to *HFR131*, indicating that *OsARF17* may regulate plant height and grain length through *HFR131*.

To verify whether *OsARF17* can regulate *HFR131*, we searched for TGTCT(A/C)C auxin response elements (AuxREs), which are bound by auxin response factors (ARFs) and confer auxin responsiveness (Wang et al., 2007; Shen et al., 2010), within 3500 bp region of *HFR131* promoter and 10 AuxREs were found (Figure 5H). We divided the 10 AuxREs into eight fragments with lengths of 150–300 bp (A1–A8) (Figure 5H) and verified by yeast one-hybrid assay and ChIP-qPCR. The yeast one-hybrid assay showed that yeast co-transformed with *ProHFR131-A6* and *AD-ARF17* grew strongly in the selective medium, whereas yeast co-transformed with *ProHFR131-A6* and *KAD* grew weakly (Figure 5I). The ChIP-qPCR analysis showed that the A6 fragment of the *HFR131* promoter was significantly enriched in the two overexpressed lines (Figure 5J). These results indicated that OsARF17 could bind to the A6 fragment of the *HFR131* promoter. A qPCR analysis revealed that the expression level of *HFR131* was significantly increased in *OsARF17* overexpression lines compared with that of the control ZH11 (Figure 5K). These results showed that OsARF17 could bind to the promoter of *HFR131* and positively regulated the expression of *HFR131*.

To further verify whether *OsARF17* was involved in the regulation of *HFR131*, we conducted a BR treatment assay. The results showed that, after treatment with 1 μM/L 24-eBL, root elongation was inhibited more slightly in *OsARF17* knockout lines than in the control ZH11. The length of the coleoptile of ZH11 was

10.3389/fpls.2023.1152196

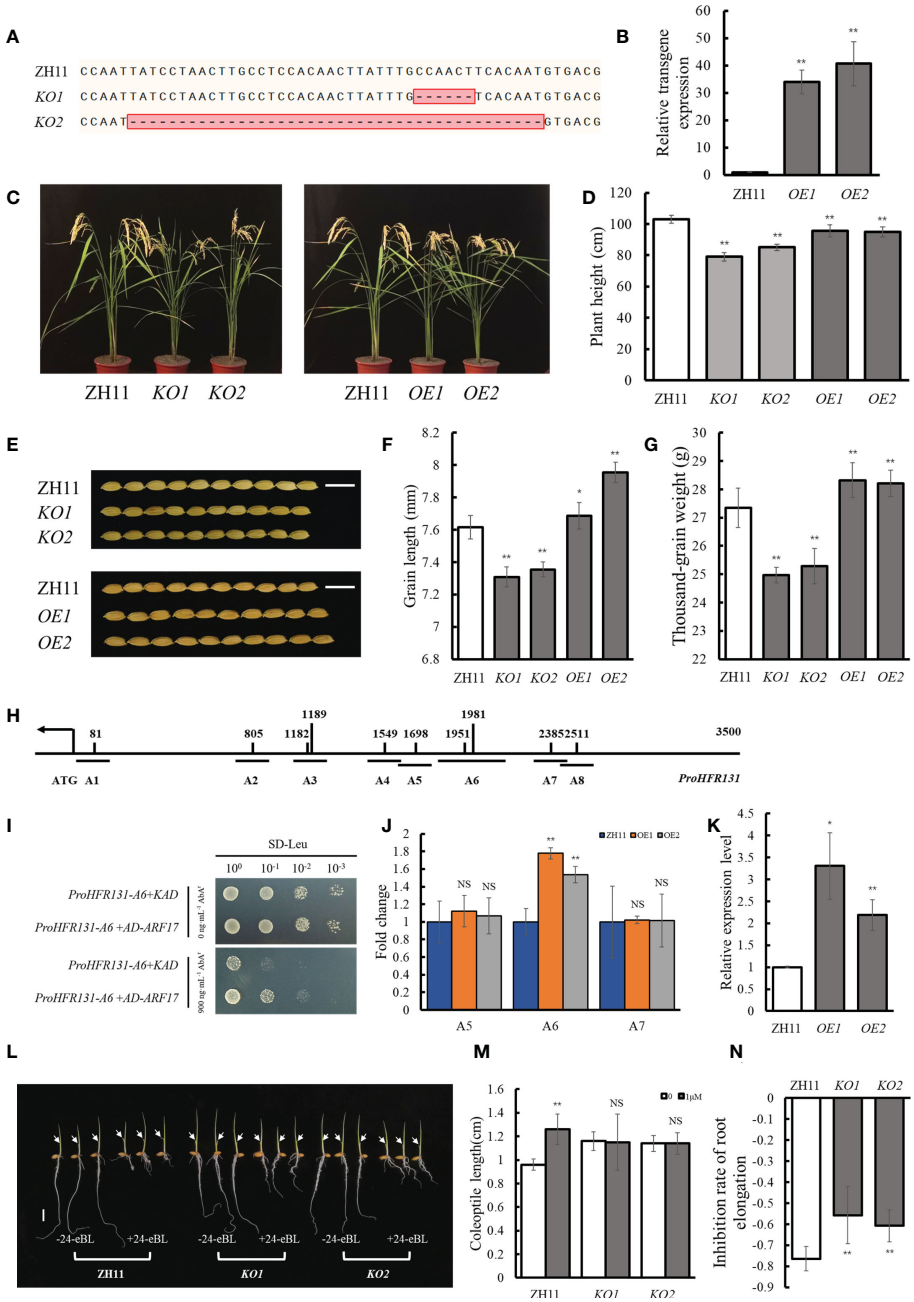


FIGURE 5
Relationship between *OsARF17* and *HFR131*. **(A)** Target site in the *OsARF17* sequence of rice 'ZH11' in the *OsARF17* knockout lines. **(B)** Transgene expression level of *OsARF17*. *OsActin* was used as an internal reference. **(C, D)** Plant architecture and plant height of *OsARF17* knockout lines and overexpression lines. **(E)** Grain morphology of *OsARF17* knockout lines and overexpression lines. Bar = 1 cm. **(F, G)** Grain length and thousand-grain weight of *OsARF17* knockout lines and overexpression lines. **(H)** Positions of 10 auxin response elements in the *HFR131* promoter. **(I)** Yeast one-hybrid assays to detect interaction between OsARF17 and the promoter of *HFR131*. **(J)** ChIP-qPCR analysis to detect interaction between OsARF17 and the promoter of *HFR131*. Fold change is relative to ZH11. *OsActin* was used as an internal reference. A5, A6, and A7 corresponding to the fragments indicated in Figure 5A. **(K)** Relative expression level of *HFR131* in *OsARF17* overexpression lines. *OsUBQ5* was used as an internal reference. **(L)** Effect of 24-eBL on coleoptile and root elongation in seedlings. The arrows indicate the top of the coleoptile. Bar = 1 cm. **(M)** Coleoptile length of ZH11 and *OsARF17* knockout-line seedlings in the presence or absence of 1 μM/L 24- epibrassinolide (24-eBL). **(N)** Inhibition of root elongation in ZH11 and *OsARF17* knockout lines under 1 μM/L 24-eBL treatment. * $P < 0.05$, ** $P < 0.01$, NS non-significant.

significantly increased, whereas no significant change was observed in the *OsARF17* knockout lines (Figures 5L–N), indicating that plant sensitivity to BR was decreased after *OsARF17* knockout. These results indicated that *OsARF17* was involved in BR response.

Based on these results, we concluded that OsARF17 could regulate plant height and grain length through regulating the expression of *HFR131*, that *OsARF17* may participate in BR response through *HFR131*, and that *OsARF17–HFR131*

10.3389/fpls.2023.1152196

interaction provides a bridge for communication between auxin and BR.

Evolutionary correlation analysis of *HFR131* and *OsARF17*

To further understand the relationship between *HFR131* and *OsARF17* in regulating plant height and grain length, we downloaded the gcHap data for *HFR131* and *OsARF17* from the RFGB and conducted a haplotype analysis. Eight major Haps (≥30 rice accessions) (Zhang et al., 2021) were resolved based on 20 SNPs distributed among the CDS region and the 3′ untranslated region of *HFR131*, and a haplotype network was constructed (Figures 6A, B). *Xian/Indica* (*XI*) accessions predominantly harbored Hap1, Hap4, Hap7, and Hap8, whereas *Geng/Japonica* (*GJ*) accessions predominantly carried Hap2, Hap3, and Hap6, indicating that there was a strong *XI*–*GJ* differentiation of *HFR131* Haps (Figures 6B, C). Nucleotide diversity (π) among different populations revealed that *GJ* harbored higher diversity than *XI*, not only in the gene region of *HFR131*, but also in the ~500 kb upstream/downstream regions of *HFR131* (Figure 6D). Furthermore, *HFR131* has not evolved neutrally; the gene was subject to balancing selection in *GJ*, and directional selection or population expansion in *XI* (Figure 6E), which indicated that *HFR131* may have contributed to *XI* domestication.

We then investigated the potential associations of major Haps with plant height, grain length, and thousand-grain weight by one-

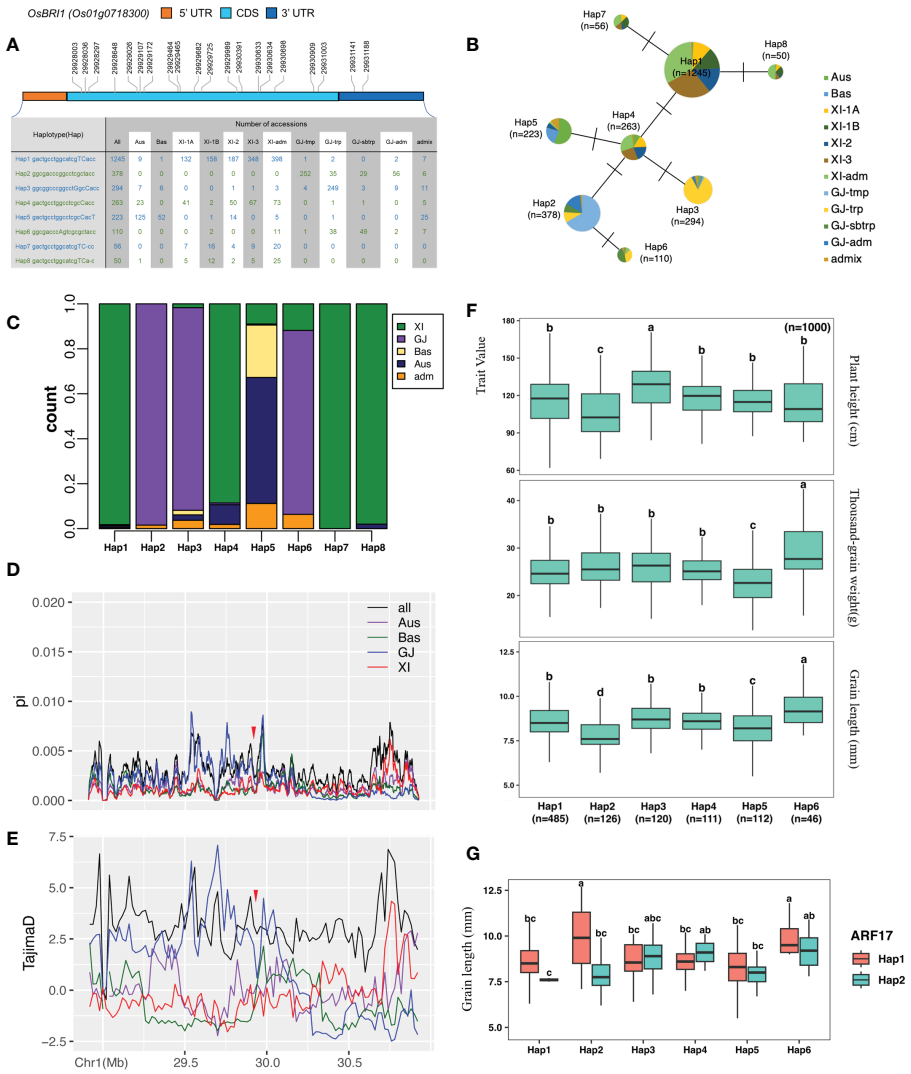


FIGURE 6
Haplotype analysis of *HFR131* and *OsARF17*. **(A)** Haplotypes of *HFR131*/*OsBRI1*. An uppercase letter in the haplotype sequence indicates a nonsynonymous mutation compared with Hap2. **(B)** Evolutionary network of *HFR131* haplotypes. **(C)** Frequency of *HFR131* haplotypes among subgroups of rice accessions. **(D, E)** Nucleotide diversity (π) and Tajima's *D* for ~1 Mb genomic region flanking *HFR131*. The red arrow indicates the position of *HFR131*. **(F)** Distribution of plant height, grain length, and thousand-grain weight among rice accessions with different *HFR131* haplotypes. **(G)** Grain length of rice accessions with different haplotypes of *HFR131* and *OsARF17* *Aus*, *Bas*, *XI*, and *GJ* refer to different subgroups of rice and *Adm* (*admix*) is a fully mixed group. *XI-1A*, *XI-1B*, *XI-2*, *XI-3*, and *XI-adm* belong to the *XI* group. *GJ-tmp*, *GJ-trp*, *GJ-sbtrp*, and *GJ-adm* belong to the *GJ* group. Different letters on the boxplots indicate statistically significant differences at $P < 0.05$ based on Duncan's multiple range test.

way ANOVA followed by Duncan's multiple-range tests. Hap3 was associated with the highest average plant height (127.64 ± 18.21 cm), whereas Hap2 was associated with the shortest average plant height (106.18 ± 21.3 cm). Both grain length and thousand-grain weight showed low and high values in Hap5 and Hap6, respectively, inferring that the thousand-grain weight of Hap5 and Hap6 was mainly determined by grain length. Hap7 and Hap8 were not included in the haplotype–trait analysis because the samples with phenotype data for the selected traits were too small (nine and four, respectively) (Figure 6F).

As a regulator of *HFR131*, we chose two major haplotypes (Hap1 and Hap2) of *OsARF17* (≥600 rice accessions) to further investigate its influence on *HFR131*. $OsARF17^{Hap1}$ and $OsARF17^{Hap2}$ were present in 2680 rice accessions and also showed an unequal distribution between *XI* and *GJ* (Supplementary Figure 3A). $OsARF17^{Hap1}$ was present mostly in *XI*, whereas $OsARF17^{Hap2}$ was predominantly in *GJ*, especially in *GJ-tmp* (*GJ* from East Asian temperate) and *GJ-trp* (GJ from Southeast Asian tropical) accessions (Supplementary Figures 3A, B). We performed two-way ANOVA followed by Duncan's multiple-range test to determine the consociation between *HFR131* and *OsARF17*, and we observed that a significant consociation only existed when considering grain length. Both $HFR131^{Hap2}$/$OsARF17^{Hap1}$ and $HFR131^{Hap6}$/$OsARF17^{Hap1}$ showed an "a" level in multiple comparisons with long grain length (Figure 6G). However, the $HFR131^{Hap2}$/$OsARF17^{Hap1}$ group consisted of only two samples and the grain length of these two samples was divergent (12.7 mm and 7.1 mm, respectively). The grain length of $HFR131^{Hap6}$ and $OsARF17^{Hap1}$ was 9.391 mm and 8.567 mm, respectively; however, the grain length of $HFR131^{Hap6}$/$OsARF17^{Hap1}$ increased to 9.9 mm (Figure 6G, Supplementary Figure 3C), revealing that the consociation of $OsAFR17^{Hap1}$/$HFR131^{Hap6}$ conferred a longer grain length, and thus improved rice yield.

Discussion

The mutation of conserved amino acid in the kinase domain defects the function of *OsBRI1*

In *Arabidopsis thaliana*, BRI1 is a receptor of BR, containing a leucine-rich repeat (LRR) domain, a transmembrane domain, and a kinase domain, of which the kinase domain is crucial for the transduction of BR signals (Li and Chory, 1997). BR binds to the extracellular LRR domain of BRI1, and induces association and transphosphorylation between the kinase domains of BRI1 and the co-receptor kinase BAK1. BRI1 activates BAK1 by phosphorylating the activation loop of BAK1, and BAK1 phosphorylates the juxtamembrane and kinase domains of BRI1 to enhance BRI1 signaling in turn. The BR signal is then transmitted downstream (Wang et al., 2012). It has been reported that OsBRI1 and OsBAK1 have a similar structure and function to BRI1 and BAK1, respectively (Yamamuro et al., 2000; Song et al., 2008; Li et al., 2009). In rice, 11 OsBRI1-associated mutants (*d61-1* to *d61-10* and *Fn189*) have been reported, among which the *Fn189*, *d61-1*, *d61-4*, and *d61-10* mutants harbor mutations in the kinase domain of OsBRI1 (Yamamuro et al., 2000; Nakamura et al., 2006; Zhao et al., 2013). In the present study, we identified a new allele of *OsBRI1*, *hfr131*, in which a new mutation site, T988I, in the kinase domain of OsBRI1 was detected, causing significant reductions in plant height and grain length. We compared the protein sequences of the BR receptor and its homologous genes in *Arabidopsis thaliana*, rice, tomato, pea, and barley, and determined that the mutation site is located in a highly conserved region (Figure 4C), indicating that the site is critical for the function of the OsBRI1 kinase domain and therefore for the function of the gene. Serine/threonine and tyrosine residues are the common phosphorylation sites in BRI1 and BAK1 (Wang et al., 2012). However, T988 in the kinase domain of OsBRI1 was changed to I, an amino acid that cannot be phosphorylated, in NIL-*hfr131*, indicating that this change may affect the phosphorylation of OsBAK1 to OsBRI1, thus hindering transduction of the BR signal.

The plant height of rice mainly depends on the elongation of internodes. Dwarf mutants can be categorized into six types according to the elongation pattern of the upper four or five internodes, namely N-type (wild type), dn-type, dm-type, d6-type, nl-type, and sh-type (Yamamuro et al., 2000). Mutants defective in BR synthesis or signaling are often of the dm-type or d6-type. The *d2*, *d11*, and *d61-1* mutants exhibit the dm-type internode elongation pattern (Supplementary Table 2), showing a specific shortening of the second internode. The *brd2*, *d61-2*, and *Fn189* mutants show the d6-type internode elongation pattern (Supplementary Table 2), in which all the internodes are extremely shortened except for the first internode. However, brd1 failed to elongate any internodes (Supplementary Table 2). The T988I conversion in the kinase domain of *OsBRI1*/*HFR131* caused reduction in plant height with equally proportional shortening of all internodes (Figure 1G), which represented the dn-type internode elongation pattern and had not been reported previously in BR synthesis- or *OsBRI1*-defective mutants (Supplementary Table 2).

hfr131 confers the dn-type internode elongation pattern by coordinating different hormones

The formation of different internode elongation patterns may be associated with differences in the main regulatory hormones at different internodes. GA is considered to play an important role in the elongation of the first internode in rice. The first internode of the GA synthesis-defective *d35* mutant is shortened to a significantly greater degree than the other internodes (Itoh et al., 2004). *HOX12* and *EUI* influence elongation of the first internode by regulating the contents of active GAs (Luo et al., 2006; Zhu et al., 2006; Gao et al., 2016). All phenotypic effects of BR result from BR synthesis and perception. BR-synthesis and *OsBRI1*-related mutants will assist in providing an improved understanding of the role of BR in rice internode regulation. All BR-synthesis and *OsBRI1*-related mutants reported to date show severe shortening or non-elongation of the second internode, with dm- or d6-type elongation patterns

(Supplementary Table 2), indicating that elongation of the second internode may be mainly determined by BR. For the lower internodes, auxin may be the primary functional hormone. Loss-of-function of BR2, which functions in indole-3-acetic acid export (Knöller et al., 2010), disrupts polar auxin transport in the maize stalk, resulting in a reduced plant height with compact lower-stalk internodes (Multani et al., 2003). These findings confirmed that the distribution of auxin plays a crucial role in regulating the lower internodes. In the present study, NIL-*hfr131*, harboring a new allele of *OsBRI1*, *hfr131*, showed a dn-type internode elongation pattern rather than the d6- or dm-type, indicating that this allele can coordinate different hormones. It was further confirmed by transcriptome analysis that *OsGA2ox10*, which participates in the GA catabolic pathways (Lo et al., 2008), was significantly up-regulated and that *OsPIN1d*, *OsPIN8*, and *OsAUX2*, which are auxin transporters (Wang et al., 2009; Yu et al., 2015), were significantly down-regulated (Supplementary Table 1), implying that there may be coordination between different hormones.

We noticed that the location of *HFR131*/*OsBRI1* was closed to the famous "Green Revolution" gene *SD1* (Monna et al., 2002; Sasaki et al., 2002; Su et al., 2021), both of which were located on the long arm of chromosome 1, and the distance of them was nearly 8500 kb (Supplementary Figure 1). Either *SD1*-mutant or NIL-*hfr131* showed a semi-dwarf plant height, but the expression pattern of them were different. *SD1* was widely expressed in different organs especially in leaf blades (Monna et al., 2002; Sasaki et al., 2002; Su et al., 2021), however, *OsBRI1* was strongly expressed in shoot apex but hardly in leaf blades (Yamamuro et al., 2000), indicating that they regulated plant height through different pathway, also implying that *hfr131* may be another useful dwarfing gene resource after *sd1*.

OsARF17-HFR131 links brassinosteroid and auxin signaling during regulation of plant height and grain length

Auxin and BR coordinate to regulate many aspects of plant growth and development, such as root and bud development (Mazzoni-Putman et al., 2021), and several studies have shown that there is crosstalk between BR and auxin; for example, exogenous auxin improves the sensitivity of rice to BR (Fujioka et al., 1998), and BR affects the transport of auxin (Symons et al., 2007). With extension of the research scope, ARF family proteins were found to be important factors communicating between the BR and auxin signaling pathways. In *Arabidopsis thaliana*, *ARF2*, *ARF7*, and *ARF19* are involved in BR and auxin communication through different mechanisms (Vert et al., 2008; Sun et al., 2010; Cho et al., 2014). In rice, OsARF11 and OsARF19 can bind to the *OsBRI1* promoter and promote the expression of *OsBRI1*, thus enhancing the rice response to BR and regulating leaf angle (Sakamoto et al., 2013; Zhang et al., 2015). However, how the BR and auxin communication affects plant height and grain size in rice is poorly understood. In the present study, we confirmed that OsARF17 may be involved in the response to BR through *HFR131*. OsARF17 can bind to the promoter region of *HFR131* and positively regulate the expression of *HFR131* (Figures 5I–K). *OsARF17* knockout lines showed decreased sensitivity to BR (Figures 5L–N). *OsARF17* and *HFR131* defective and overexpression lines showed a similar phenotype. Both *OsARF17* knockout lines and NIL-*hfr131* showed a reduced plant height and grain length, whereas both overexpression lines of *OsARF17* and *HFR131* exhibited an increased grain length, indicating that BR and auxin regulate plant height and grain length synergistically through *OsARF17–HFR131* interaction. In addition, $HFR131^{Hap6}$/$OsARF17^{Hap1}$ in combination conferred a longer grain length than both $HFR131^{Hap6}$ and $OsARF17^{Hap1}$ individually. This may be the result of efficient coordination of the BR and auxin signaling pathways between the two haplotypes, and highlights the importance of breaking through the limitation of traditional single-gene breeding to increase rice yield.

In summary, this study identified a new allele of *OsBRI1*, *hfr131*, which caused a significant decrease in plant height and grain length of rice. It was further found that *OsARF17* could regulate the plant height and grain length of rice, and affect the sensitivity of rice to BR, by regulating the expression of *HFR131*. In addition, the $OsAFR17^{Hap1}$/$HFR131^{Hap6}$ consociation was confirmed to contribute to longer grain length. These findings are of great importance for improvement of rice yield by molecular breeding in the future.

Data availability statement

The data presented in the study are deposited in the GEO repository, accession number GSE224118

Author contributions

XL, YW, DL, and LC contributed to conception and design of the study. DL, LC, ML, FM, and PY performed the experiments. DL, LC, XZ, JH, FN, SH, JC, XY, and JY analyzed and interpreted the data. DL, LC, and ML wrote the manuscript and XL and YW revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was sponsored by National Natural Science Foundation of China (32172043, 31971918 and 31900393), Shanghai Rising-Star Program (20QB1402100), Natural Science Foundation of Shanghai (19ZR1436600) and Innovation Program of Shanghai Municipal Education Commission (2023ZKZD05).

Acknowledgments

We thank Jinying Gou (Fudan University) for providing *pAbAi* vector and Yan Yan and Yue Hu (Fudan University) for the guidance of yeast one-hybrid assay.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1152196/full#supplementary-material

References

Bai, M.-Y., Zhang, L.-Y., Gampala, S. S., Zhu, S.-W., Song, W.-Y., Chong, K., et al. (2007). Functions of OsBZR1 and 14-3-3 proteins in brassinosteroid signaling in rice. *Proc. Natl. Acad. Sci.* 104 (34), 13839–13844. doi: 10.1073/pnas.0706386104

Chen, S.-H., Zhou, L.-J., Xu, P., and Xue, H.-W. (2018). SPOC domain-containing protein leaf inclination3 interacts with LIP1 to regulate rice leaf inclination through auxin signaling. *PloS Genet.* 14 (11), e1007829. doi: 10.1371/journal

Cho, H., Ryu, H., Rho, S., Hill, K., Smith, S., Audenaert, D., et al. (2014). A secreted peptide acts on BIN2-mediated phosphorylation of ARFs to potentiate auxin response during lateral root development. *Nat. Cell Biol.* 16 (1), 66–76. doi: 10.1038/ncb2893

Clouse, S. D. (2011). Brassinosteroids. *Arabidopsis Book* 9, e0151. doi: 10.1199/tab.0151

Economic and Social Affairs (2022). *World population prospects 2022: Summary of results* (NY, USA: United Nations New York).

Fujioka, S., Noguchi, T., Takatsuto, S., and Yoshida, S. (1998). Activity of brassinosteroids in the dwarf rice lamina inclination bioassay. *Phytochemistry* 49 (7), 1841–1848. doi: 10.1016/S0031-9422(98)00412-9

Gao, S., Fang, J., Xu, F., Wang, W., and Chu, C. (2016). Rice HOX12 regulates panicle exsertion by directly modulating the expression of *ELONGATED UPPERMOST INTERNODE1*. *Plant Cell* 28 (3), 680–695. doi: 10.1105/tpc.15.01021

Gui, J., Zheng, S., Liu, C., Shen, J., Li, J., and Li, L. (2016). OsREM4.1 interacts with OsSERK1 to coordinate the interlinking between abscisic acid and brassinosteroid signaling in rice. *Dev. Cell* 38 (2), 201–213. doi: 10.1016/j.devcel.2016.06.011

Hargrove, T. R., and Cabanilla, V. L. (1979). The impact of semidwarf varieties on Asian rice-breeding programs. *BioScience* 29 (12), 731–735. doi: 10.2307/1307667

Hong, Z., Ueguchi-Tanaka, M., Fujioka, S., Takatsuto, S., Yoshida, S., Hasegawa, Y., et al. (2005). The rice *brassinosteroid-deficient dwarf2* mutant, defective in the rice homolog of arabidopsis DIMINUTO/DWARF1, is rescued by the endogenously accumulated alternative bioactive brassinosteroid, dolichosterone. *Plant Cell* 17 (8), 2243–2254. doi: 10.1105/tpc.105.030973

Hong, Z., Ueguchi-Tanaka, M., Umemura, K., Uozu, S., Fujioka, S., Takatsuto, S., et al. (2003). A rice brassinosteroid-deficient mutant, *ebisu dwarf* (*d2*), is caused by a loss of function of a new member of cytochrome P450. *Plant Cell* 15 (12), 2900–2910. doi: 10.1105/tpc.014712

Hu, J., Wang, Y., Fang, Y., Zeng, L., Xu, J., Yu, H., et al. (2015). A rare allele of *GS2* enhances grain size and grain yield in rice. *Mol. Plant* 8 (10), 1455–1465. doi: 10.1016/j.molp.2015.07.002

Huang, Y., Dong, H., Mou, C., Wang, P., Hao, Q., Zhang, M., et al. (2022). Ribonuclease h-like gene *SMALL GRAIN2* regulates grain size in rice through brassinosteroid signaling pathway. *J. Integr. Plant Biol.* 64 (10), 1883–1900. doi: 10.1111/jipb.13333

Huang, G., Hu, H., van de Meene, A., Zhang, J., Dong, L., Zheng, S., et al. (2021). AUXIN RESPONSE FACTORS 6 and 17 control the flag leaf angle in rice by regulating secondary cell wall biosynthesis of lamina joints. *Plant Cell* 33 (9), 3120–3133. doi: 10.1093/plcell/koab175

Itoh, H., Tatsumi, T., Sakamoto, T., Otomo, K., Toyomasu, T., Kitano, H., et al. (2004). A rice semi-dwarf gene, *Tan-ginbozu* (*D35*), encodes the gibberellin biosynthesis enzyme, ent-kaurene oxidase. *Plant Mol. Biol.* 54 (4), 533–547. doi: 10.1023/b:plan.0000038261.21060.47

Khush, G. S. (1999). Green revolution: preparing for the 21st century. *Genome* 42 (4), 646–655. doi: 10.1139/g99-044

Knöller, A. S., Blakeslee, J. J., Richards, E. L., Peer, W. A., and Murphy, A. S. (2010). Brachytic2/*Zm*ABCB1 functions in IAA export from intercalary meristems. *J. Exp. Bot.* 61 (13), 3689–3696. doi: 10.1093/jxb/erq180

Li, J., and Chory, J. (1997). A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. *Cell* 90 (5), 929–938. doi: 10.1016/S0092-8674(00)80357-8

Li, Y., Li, J., Chen, Z., Wei, Y., Qi, Y., and Wu, C. (2020). OsmiR167a-targeted auxin response factors modulate tiller angle *via* fine-tuning auxin distribution in rice. *Plant Biotechnol. J.* 18 (10), 2015–2026. doi: 10.1111/pbi.13360

Li, D., Wang, L., Wang, M., Xu, Y.-Y., Luo, W., Liu, Y.-J., et al. (2009). Engineering *OsBAK1* gene as a molecular tool to improve rice architecture for high yield. *Plant Biotechnol. J.* 7 (8), 791–806. doi: 10.1111/j.1467-7652.2009.00444.x

Liu, F., Wang, P., Zhang, X., Li, X., Yan, X., Fu, D., et al. (2018). The genetic and molecular basis of crop height based on a rice model. *Planta* 247 (1), 1–26. doi: 10.1007/s00425-017-2798-1

Liu, D., Yu, Z., Zhang, G., Yin, W., Li, L., Niu, M., et al. (2021). Diversification of plant agronomic traits by genome editing of brassinosteroid signaling family genes in rice. *Plant Physiol.* 187 (4), 2563–2576. doi: 10.1093/plphys/kiab394

Lo, S.-F., Yang, S.-Y., Chen, K.-T., Hsing, Y.-I., Zeevaart, J. A. D., Chen, L.-J., et al. (2008). A novel class of gibberellin 2-oxidases control semidwarfism, tillering, and root development in rice. *Plant Cell* 20 (10), 2603–2618. doi: 10.1105/tpc.108.060913

Luo, A., Qian, Q., Yin, H., Liu, X., Yin, C., Lan, Y., et al. (2006). *EUI1*, encoding a putative cytochrome P450 monooxygenase, regulates internode elongation by modulating gibberellin responses in rice. *Plant Cell Physiol.* 47 (2), 181–191. doi: 10.1093/pcp/pci233

Lyu, J., Wang, D., Duan, P., Liu, Y., Huang, K., Zeng, D., et al. (2020). Control of grain size and weight by the GSK2-LARGE1/OML4 pathway in rice. *Plant Cell* 32 (6), 1905–1918. doi: 10.1105/tpc.19.00468

Mazzoni-Putman, S. M., Brumos, J., Zhao, C. S., Alonso, J. M., and Stepanova, A. N. (2021). Auxin interactions with other hormones in plant development. *Cold Spring Harbor Perspect. Biol.* 13 (10), a039990. doi: 10.1101/cshperspect.a039990

McClung, C. R. (2014). Making hunger yield. *Science* 344 (6185), 699–700. doi: 10.1126/science.1254135

Monna, L., Kitazawa, N., Yoshino, R., Suzuki, J., Masuda, H., Maehara, Y., et al. (2002). Positional cloning of rice semidwarfing gene, *sd-1*: rice "green revolution gene" encodes a mutant enzyme involved in gibberellin synthesis. *DNA Res.* 9 (1), 11–17. doi: 10.1093/dnares/9.1.11

Mori, M., Nomura, T., Ooka, H., Ishizaka, M., Yokota, T., Sugimoto, K., et al. (2002). Isolation and characterization of a rice dwarf mutant with a defect in brassinosteroid biosynthesis. *Plant Physiol.* 130 (3), 1152–1161. doi: 10.1104/pp.007179

Morinaka, Y., Sakamoto, T., Inukai, Y., Agetsuma, M., Kitano, H., Ashikari, M., et al. (2006). Morphological alteration caused by brassinosteroid insensitivity increases the biomass and grain production of rice. *Plant Physiol.* 141 (3), 924–931. doi: 10.1104/pp.106.077081

Multani, D. S., Briggs, S. P., Chamberlin, M. A., Blakeslee, J. J., Murphy, A. S., and Johal, G. S. (2003). Loss of an MDR transporter in compact stalks of maize *br2* and sorghum *dw3* mutants. *Science* 302 (5642), 81–84. doi: 10.1126/science.1086072

Nakagawa, H., Tanaka, A., Tanabata, T., Ohtake, M., Fujioka, S., Nakamura, H., et al. (2011). *SHORT GRAIN1* decreases organ elongation and brassinosteroid response in rice. *Plant Physiol.* 158 (3), 1208–1219. doi: 10.1104/pp.111.187567

Nakamura, A., Fujioka, S., Sunohara, H., Kamiya, N., Hong, Z., Inukai, Y., et al. (2006). The role of *OsBRI1* and its homologous genes, *OsBRL1* and *OsBRL3*, in rice. *Plant Physiol.* 140 (2), 580–590. doi: 10.1104/pp.105.072330

Niu, M., Wang, H., Yin, W., Meng, W., Xiao, Y., Liu, D., et al. (2022). Rice DWARF AND LOW-TILLERING and the homeodomain protein OSH15 interact to regulate internode elongation *via* orchestrating brassinosteroid signaling and metabolism. *Plant Cell* 34 (10), 3754–3772. doi: 10.1093/plcell/koac196

Qiao, J., Jiang, H., Lin, Y., Shang, L., Wang, M., Li, D., et al. (2021). A novel *miR167a-OsARF6-OsAUX3* module regulates grain length and weight in rice. *Mol. Plant* 14 (10), 1683–1698. doi: 10.1016/j.molp.2021.06.023

Ramamoorthy, R., Jiang, S.-Y., and Ramachandran, S. (2011). *Oryza sativa* cytochrome P450 family member OsCYP96B4 reduces plant height in a transcript dosage dependent manner. *PloS One* 6 (11), e28069. doi: 10.1371/journal.pone.0028069

Sakamoto, T., Morinaka, Y., Inukai, Y., Kitano, H., and Fujioka, S. (2013). Auxin signal transcription factor regulates expression of the brassinosteroid receptor gene in rice. *Plant Cell* 73 (4), 676–688. doi: 10.1111/tpj.12071

Sasaki, A., Ashikari, M., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Swapan, D., et al. (2002). A mutant gibberellin-synthesis gene in rice. *Nature* 416 (6882), 701–702. doi: 10.1038/416701a

Shen, C., Wang, S., Bai, Y., Wu, Y., Zhang, S., Chen, M., et al. (2010). Functional analysis of the structural domain of ARF proteins in rice (*Oryza sativa* l.). *J. Exp. Bot.* 61 (14), 3971–3981. doi: 10.1093/jxb/erq208

Song, D., Li, G., Song, F., and Zheng, Z. (2008). Molecular characterization and expression analysis of *OsBISERK1*, a gene encoding a leucine-rich repeat receptor-like kinase, during disease resistance responses in rice. *Mol. Biol. Rep.* 35 (2), 275–283. doi: 10.1007/s11033-007-9080-8

Su, S., Hong, J., Chen, X., Zhang, C., Chen, M., Luo, Z., et al. (2021). Gibberellins orchestrate panicle architecture mediated by DELLA-KNOX signalling in rice. *Plant Biotechnol. J.* 19 (11), 2304–2318. doi: 10.1111/pbi.13661

Sun, Y., Fan, X.-Y., Cao, D.-M., Tang, W., He, K., Zhu, J.-Y., et al. (2010). Integration of brassinosteroid signal transduction with the transcription network for plant growth regulation in *Arabidopsis*. *Dev. Cell* 19 (5), 765–777. doi: 10.1016/j.devcel.2010.10.010

Sun, L., Li, X., Fu, Y., Zhu, Z., Tan, L., Liu, F., et al. (2013). *GS6*, a member of the GRAS gene family, negatively regulates grain size in rice. *J. Integr. Plant Biol.* 55 (10), 938–949. doi: 10.1111/jipb.12062

Symons, G. M., Ross, J. J., Jager, C. E., and Reid, J. B. (2007). Brassinosteroid transport. *J. Exp. Bot.* 59 (1), 17–24. doi: 10.1093/jxb/erm098

Tanabe, S., Ashikari, M., Fujioka, S., Takatsuto, S., Yoshida, S., Yano, M., et al. (2005). A novel cytochrome P450 is implicated in brassinosteroid biosynthesis *via* the characterization of a rice dwarf mutant, *dwarf11*, with reduced seed length. *Plant Cell* 17 (3), 776–790. doi: 10.1105/tpc.104.024950

Tanaka, A., Nakagawa, H., Tomita, C., Shimatani, Z., Ohtake, M., Nomura, T., et al. (2009). *BRASSINOSTEROID UPREGULATED1*, encoding a helix-Loop-Helix protein, is a novel gene involved in brassinosteroid signaling and controls bending of the lamina joint in rice. *Plant Physiol.* 151 (2), 669–680. doi: 10.1104/pp.109.140806

Tong, H., Jin, Y., Liu, W., Li, F., Fang, J., Yin, Y., et al. (2009). DWARF AND LOW-TILLERING, a new member of the GRAS family, plays positive roles in brassinosteroid signaling in rice. *Plant J.* 58 (5), 803–816. doi: 10.1111/j.1365-313X.2009.03825.x

Tong, H., Xiao, Y., Liu, D., Gao, S., Liu, L., Yin, Y., et al. (2014). Brassinosteroid regulates cell elongation by modulating gibberellin metabolism in rice. *Plant Cell* 26 (11), 4376–4393. doi: 10.1105/tpc.114.132092

Vert, G., Walcher, C. L., Chory, J., and Nemhauser, J. L. (2008). Integration of auxin and brassinosteroid pathways by auxin response factor 2. *Proc. Natl. Acad. Sci.* 105 (28), 9829–9834. doi: 10.1073/pnas.0803996105

Wang, Z.-Y., Bai, M.-Y., Oh, E., and Zhu, J.-Y. (2012). Brassinosteroid signaling network and regulation of photomorphogenesis. *Annu. Rev. Genet.* 46 (1), 701–724. doi: 10.1146/annurev-genet-102209-163450

Wang, J.-R., Hu, H., Wang, G.-H., Li, J., Chen, J.-Y., and Wu, P. (2009). Expression of *PIN* genes in rice (*Oryza sativa* l.): Tissue specificity and regulation by hormones. *Mol. Plant* 2 (4), 823–831. doi: 10.1093/mp/ssp023

Wang, Y., Luo, X., Sun, F., Hu, J., Zha, X., Su, W., et al. (2018). Overexpressing lncRNA LAIR increases grain yield and regulates neighbouring gene cluster expression in rice. *Nat. Commun.* 9 (1), 3516. doi: 10.1038/s41467-018-05829-7

Wang, D., Pei, K., Fu, Y., Sun, Z., Li, S., Liu, H., et al. (2007). Genome-wide analysis of the *auxin response factors* (ARF) gene family in rice (*Oryza sativa*). *Gene* 394 (1-2), 13–24. doi: 10.1016/j.gene.2007.01.006

Wang, L., Xu, Y., Zhang, C., Ma, Q., Joo, S.-H., Kim, S.-K., et al. (2008). OsLIC, a novel CCCH-type zinc finger protein with transcription activation, mediates rice architecture *via* brassinosteroids signaling. *PloS One* 3 (10), e3521. doi: 10.1371/journal.pone.0003521

Xiao, Y., Zhang, G., Liu, D., Niu, M., Tong, H., and Chu, C. (2020). GSK2 stabilizes OFP3 to suppress brassinosteroid responses in rice. *Plant J.* 102 (6), 1187–1201. doi: 10.1111/tpj.14692

Xing, Y., and Zhang, Q. (2010). Genetic and molecular bases of rice yield. *Annu. Rev. Plant Biol.* 61 (1), 421–442. doi: 10.1146/annurev-arplant-042809-112209

Xiong, G. S., Hu, X. M., Jiao, Y. Q., Yu, Y. C., Chu, C. C., Li, J. Y., et al. (2006). *LEAFY HEAD2*, which encodes a putative RNA-binding protein, regulates shoot development of rice. *Cell Res.* 16 (3), 267–276. doi: 10.1038/sj.cr.7310034

Xiong, M., Yu, J., Wang, J., Gao, Q., Huang, L., Chen, C., et al. (2022). Brassinosteroids regulate rice seed germination through the BZR1-*RAmy3D* transcriptional module. *Plant Physiol.* 189 (1), 402–418. doi: 10.1093/plphys/kiac043

Yaish, M. W., El-kereamy, A., Zhu, T., Beatty, P. H., Good, A. G., Bi, Y.-M., et al. (2010). The APETALA-2-Like transcription factor OsAP2-39 controls key interactions between abscisic acid and gibberellin in rice. *PloS Genet.* 6 (9), e1001098. doi: 10.1371/journal.pgen.1001098

Yamamuro, C., Ihara, Y., Wu, X., Noguchi, T., Fujioka, S., Takatsuto, S., et al. (2000). Loss of function of a rice *brassinosteroid insensitive1* homolog prevents internode elongation and bending of the lamina joint. *Plant Cell* 12 (9), 1591–1605. doi: 10.1105/tpc.12.9.1591

Yoo, M.-J., Albert, V. A., Soltis, P. S., and Soltis, D. E. (2006). Phylogenetic diversification of *glycogen synthase kinase 3/SHAGGY-like kinase* genes in plants. *BMC Plant Biol.* 6 (1), 3. doi: 10.1186/1471-2229-6-3

Yu, C., Sun, C., Shen, C., Wang, S., Liu, F., Liu, Y., et al. (2015). The auxin transporter, OsAUX1, is involved in primary root and root hair elongation and in cd stress responses in rice (*Oryza sativa* l.). *Plant J.* 83 (5), 818–830. doi: 10.1111/tpj.12929

Yu, J., Xiong, H., Zhu, X., Zhang, H., Li, H., Miao, J., et al. (2017). *OsLG3* contributing to rice grain length and yield was mined by ho-LAMap. *BMC Biol.* 15 (1), 28. doi: 10.1186/s12915-017-0365-7

Zhang, F., Wang, C., Li, M., Cui, Y., Shi, Y., Wu, Z., et al. (2021). The landscape of gene-CDS-haplotype diversity in rice: Properties, population organization, footprints of domestication and breeding, and implications for genetic improvement.. *Mol Plant.* 14 (5), 787–804. doi: 10.1016/j.molp.2021.02.003

Zhang, S., Wang, S., Xu, Y., Yu, C., Shen, C., Qian, Q., et al. (2015). The auxin response factor, OsARF19, controls rice leaf angles through positively regulating *OsGH3-5* and *OsBRI1*. *Plant Cell Environ.* 38 (4), 638–654. doi: 10.1111/pce.12397

Zhao, J., Wu, C., Yuan, S., Yin, L., Sun, W., Zhao, Q., et al. (2013). Kinase activity of OsBRI1 is essential for brassinosteroids to regulate rice growth and development. *Plant Sci.* 199, 199–200. doi: 10.1016/j.plantsci.2012.10.011

Zhu, X., Liang, W., Cui, X., Chen, M., Yin, C., Luo, Z., et al. (2015). Brassinosteroids promote development of rice pollen grains and seeds by triggering expression of carbon starved anther, a MYB domain protein. *Plant J.* 82 (4), 570–581. doi: 10.1111/tpj.12820

Zhu, Y., Nomura, T., Xu, Y., Zhang, Y., Peng, Y., Mao, B., et al. (2006). *ELONGATED UPPERMOST INTERNODE* encodes a cytochrome P450 monooxygenase that epoxidizes gibberellins in a novel deactivation reaction in rice. *Plant Cell* 18 (2), 442–456. doi: 10.1105/tpc.105.038455

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 10 May 2023
DOI 10.3389/fpls.2023.1184416

OPEN ACCESS

EDITED BY
Xueyong Li,
Institute of Crop Sciences (CAAS), China

REVIEWED BY
Rao YuChun,
Zhejiang Normal University, China
Zheng Chongke,
Shandong Academy of Agricultural Sciences, China

*CORRESPONDENCE
Detang Zou
zoudtneau@126.com
Hongliang Zheng
zhenghongliang008@126.com

†These authors have contributed equally to this work

RECEIVED 11 March 2023
ACCEPTED 11 April 2023
PUBLISHED 10 May 2023

CITATION
Xu S, Cui J, Cao H, Liang S, Ma T, Liu H, Wang J, Yang L, Xin W, Jia Y, Zou D and Zheng H (2023) Identification of candidate genes for salinity tolerance in *Japonica* rice at the seedling stage based on genome-wide association study and linkage mapping.
Front. Plant Sci. 14:1184416.
doi: 10.3389/fpls.2023.1184416

Identification of candidate genes for salinity tolerance in *Japonica* rice at the seedling stage based on genome-wide association study and linkage mapping

Shanbin Xu†, Jingnan Cui†, Hu Cao†, Shaoming Liang, Tianze Ma, Hualong Liu, Jingguo Wang, Luomiao Yang, Wei Xin, Yan Jia, Detang Zou* and Hongliang Zheng*

Key Laboratory of Germplasm Enhancement, Physiology and Ecology of Food Crops in Cold Region, Ministry of Education, Northeast Agricultural University, Harbin, China

Background: Salinity tolerance plays a vital role in rice cultivation because the strength of salinity tolerance at the seedling stage directly affects seedling survival and final crop yield in saline soils. Here, we combined a genome-wide association study (GWAS) and linkage mapping to analyze the candidate intervals for salinity tolerance in Japonica rice at the seedling stage.

Results: We used the Na+ concentration in shoots (SNC), K+ concentration in shoots (SKC), Na+/K+ ratio in shoots (SNK), and seedling survival rate (SSR) as indices to assess the salinity tolerance at the seedling stage in rice. The GWAS identified the lead SNP (Chr12_20864157), associated with an SNK, which the linkage mapping detected as being in qSK12. A 195-kb region on chromosome 12 was selected based on the overlapping regions in the GWAS and the linkage mapping. Based on haplotype analysis, qRT-PCR, and sequence analysis, we obtained LOC_Os12g34450 as a candidate gene.

Conclusion: Based on these results, LOC_Os12g34450 was identified as a candidate gene contributing to salinity tolerance in Japonica rice. This study provides valuable guidance for plant breeders to improve the response of Japonica rice to salt stress.

KEYWORDS

Japonica rice, GWAS, linkage mapping, salinity tolerance, seedling stage

Introduction

Soil salinity is an important limiting factor affecting high and stable resistance in rice and expansion areas for the economically important crop (Ouhibi et al., 2014). Salt stress inhibits plant protein synthesis, reduces photosynthetic efficiency, causes ion imbalances and high osmotic stress, and leads to plant wilting and apoptosis (Ma et al., 2017). This stress response significantly reduces crop yield and has become an important environmental factor affecting crop growth and development. But at present, due to unreasonable irrigation methods and excessive fertilization, the salinization of soil is becoming more and more serious (Qadir et al., 2014). Rice can effectively help people by ensuring food security and advancing sustainable agriculture development. Growing rice on saline land can improve the utilization of saline soils and the soil conditions of the land. Rice is a moderately salt-sensitive crop, and its seedling salinity tolerance is a key factor determining its final yield. Therefore, its fine positioning can not only reveal its molecular mechanism, but also provide a theoretical basis for improving salt-tolerant varieties of rice.

Multiple rice genes control the complex quantitative trait known as salinity tolerance (Liang et al., 2015), and scientists have made a series of important research advances in cloning salt stress genes (Lin et al., 2004; Bimpong et al., 2014; Yu et al., 2017; Li et al., 2022). For example, *SKC1*, rice's first salinity tolerance QTL, encodes a transporter protein of the HKT family. Previous studies have shown that *SKC1* protein is a sodium (Na^+)-selective transporter protein that can effectively regulate the aboveground sodium and potassium (Na^+/K^+) balance and improve salinity tolerance (Ren et al., 2005). Meanwhile, Huang et al. (2009) reported a rice drought and salinity tolerance gene, *DST*, which negatively regulates salinity tolerance. The functional deletion of *DST* directly down-regulates the expression of genes related to hydrogen peroxide metabolism, reduces water evaporation under drought stress and Na^+ entry into the plant, and ultimately improves the salinity tolerance of rice. Li et al. (2014) identified a lectin receptor-like kinase *SIT1* in rice, which mediates the salt-sensitive response in rice, and its rapid increase in response to high salt stimulation activates the downstream effectors *MAPK3* and *MAPK6*, which in turn increases ethylene content by activating ethylene synthase and causing an increase in reactive oxygen species, resulting in reduced survival under salt stress. Zhou et al. (2018) identified a receptor-like cytoplasmic kinase gene, *STRK1*, significantly improving salinity tolerance in rice. The *STRK1* protein undergoes autophosphorylation upon salt stress and significantly increases its activity by phosphorylating and activating CatC, thereby degrading the large amount of hydrogen peroxide produced by salt stress.

Cloning salinity tolerance genes has produced important breakthroughs in the molecular mechanism elucidation of important complex traits in rice, which have become a reference for studying the genetic mechanisms behind complex traits in rice. This work has significant application value and lays the foundation for identifying and cloning salinity tolerance genes in crops. Combining GWAS and QTL mapping for gene mining can significantly improve the efficiency of QTL identification, and the results can be verified against each other, resulting in stable and reliable QTLs. For example, Zhang et al. (2019) identified a candidate gene, *GmCDF1*, which is closely linked to soybean salinity tolerance on chromosome 8 through linkage and association analysis. Wu et al. (2016) combined GWAS and linkage analysis to identify 125 QTLs regulating maize male inflorescence size. Tang et al. (2021) used linkage localization and GWAS analysis to locate one QTL region controlling rice grain length, further identifying *OsGASR7* as a functional gene within this QTL interval. Therefore, combining the two methods is important for mining candidate genes for target traits.

This study explored the genetic mechanism of salinity tolerance as assessed by GWAS and linkage mapping using Na^+ concentration in shoots (SNC), K^+ concentration in shoots (SKC), Na^+/K^+ ratio in shoots (SNK), and seedling survival rate (SSR) at the rice seedling salt treatment. We identified Chr12_20864157 and *qSK12* on chromosome 12 by GWAS and linkage mapping, identifying an overlapping region of the 195-kb as a candidate region. Using haplotype analysis, qRT-PCR, and sequence analysis, *LOC_Os12g34450* was considered the most likely functional gene associated with salinity tolerance. The results procured by this study provide novel insights for improving salinity tolerance in *Japonica* rice.

Materials and methods

Plant materials

It consists of 295 *Japonica* rice varieties widely grown at home and abroad, of which domestic materials are mainly from Heilongjiang, Jilin, Liaoning, and Ningxia provinces, and foreign varieties were mainly from Korea, Russia, and Japan. The 195 RILs that comprised the linkage mapping population were a cross between the salt-sensitive Kongyu131 and the salt-tolerant Xiaobaijingzi. The salinity tolerance phenotypes of KY131 and XBJZ were shown in Figure 1. All rice varieties have been studied previously (Li et al., 2019; Li et al., 2020; Duan et al., 2022).

Salinity tolerance evaluation at the seedling stage

The experiment was divided into two groups, named group T_1 and group T_2, and set up for three repetitions. In group T_1, twenty-four uniformly germinated rice seeds of each variety were sown. One seed was placed in each hole and hydroponically grown with Yoshida nutrient solution. The germinated seeds were transferred to an artificial climatic chamber and incubated at 25 and 23°C with 14h light and 10h dark cycles, respectively. When seedlings grew to two leaves and one heart, salt stress treatment was performed with a pre-treatment sodium chloride (NaCl) concentration of 50 mmol/L.

After 3 days of pre-treatment, formal treatment was performed with a NaCl concentration of 120 mmol/L for 7 days. The shoot of each sample was dried at 120°C for 30 min and at 80°C to a constant weight. Weigh 0.1 grams (g) of the dry samples, add 5 ml of 1 mol/L

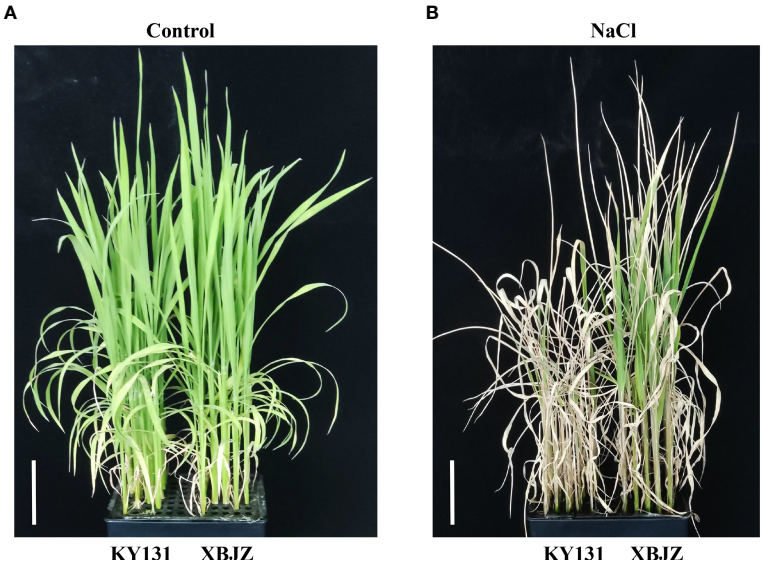


FIGURE 1
Phenotypes of KY131 and XBJZ seedlings under contral and salt stress. **(A)** Represent under control conditions. Bar = 5 cm. **(B)** Represent under salt stress conditions. Bar = 5 cm.

HCl to the tube, and put it into a water bath for 6 h at 70°C in a constant temperature water bath shaker. The SNC and SKC of the samples were measured using a flame photometer (Sherwood 410, Cambridge, UK) and the SNK was calculated. In group T_2, 100 uniformly germinated seeds of each variety were grown with Yoshida nutrient solution, with the same salt treatment as the T_1 group. The culture medium was replaced every day, and replaced with the same medium as the control group after 7 days. After 10 days, the survival rate statistics counted the number of plants with new leaf production.

GWAS for salinity tolerance

In total, 788,396 single nucleotide polymorphisms (SNPs) were used for genotyping 295 *Japonica* rice accessions for GWAS. The threshold for identifying significantly associated SNPs was set at $-\log_{10}(P) > 5.26$, according to a pre-laboratory study by Li et al. (2020). The Manhattan map was created using the R package 'qqman'. Redundant SNPs with the smallest P values were filtered within a minimal distance interval and the LDBlockShow software was used to calculate the pairwise R^2 value between any two SNPs in the interval of leading SNPs ± 2 Mb. In the interval of 1.5-2.0 Mb of leading SNPs, the average of the top 10% R^2 values was recorded, plus 0.2 to define the LD attenuation interval of leading SNPs interval (Dong et al., 2021).

QTL mapping for salinity tolerance

The genetic linkage map constructed using 195 RILs contained 527 bin markers (Figure S1). QTL localization was performed using the composite interval mapping method with QTL ICIMapping 4.2 software, and the threshold value was set to LOD>2.5, according to a pre-laboratory study by Li et al. (2020).

Haplotype analysis of candidate gene

Non-synonymous mutant SNPs in the exonic regions of all genes in the candidate interval and SNPs in the promoter region (1.5 kb before ATG) were extracted from the RiceSNPSeekDatabase website (https://snp-seek.irri.org/_snp.zul) and haplotype analysis was performed using DnaSP software. Also, materials with different haplotypes needed to be greater than or equal to 10.

Identification of candidate genes by gene expression and sequence analysis

The expression levels of the four genes of KY131 and XBJZ were verified by qRT-PCR analysis under salinated and normal conditions. qRT-PCR analysis was performed using Roche LightCycler96. All primer sequences were shown in Table S3. The CDS and promoter regions of KY131 and XBJZ candidate genes were cloned using PCR. Sequences comparison were performed using DNAMAN.

Results

Phenotypic variation

In this study, we analyzed the phenotypes of 295 *Japonica* rice accessions and RIL lines at the seedling stage under salinity stress and evaluated four salinity tolerance indices: SNC, SKC, SNK, and

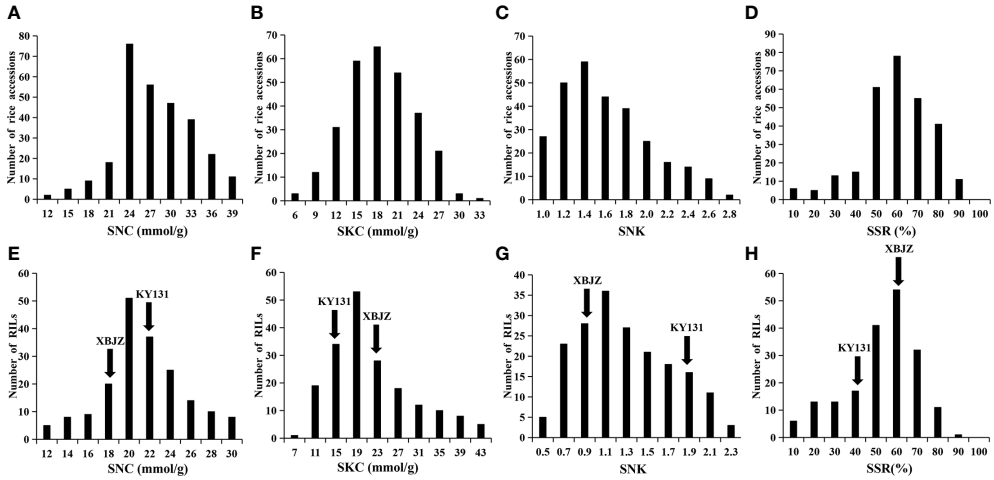


FIGURE 2
Phenotypic variation in the SNC, SKC, SNK, and SSR in 295 *Japonica* rice accessions and RIL lines. **(A–D)** Represent the SNC, SKC, SNK, and SSR of 295 rice accessions. **(E–H)** Represent the SNC, SKC, SNK, and SSR of RIL lines.

SSR. Under the salinity stress treatment, the SNC, SKC, SNK, and SSR among the 295 rice accessions varied in range from 10.14–38.83 mmol/g, 5.21–32.82 mmol/g, 0.79–2.78, and 9.33–88.66%, respectively, and the coefficient of variation was 20.13%, 26.73%, 28.61%, and 28.12%, respectively (Table S1). The variation for SNC, SKC, SNK, and SSR in the RIL lines ranged from 8.12–29.56 mmol/g, 6.98–39.56 mmol/g, 0.46–2.19, and 8.66–89.33%, respectively, and the coefficient of variation was 20.48%, 40.58%, 37.01%, and 32.99%, respectively (Table S1). Phenotypic data indicate that this study's natural and RIL populations have abundant phenotypic variation in rice seedling salinity tolerance. Meanwhile, the values of these four traits differed significantly between the two parents (Figures 2E–H; Table S1), the phenotype values of the four traits indicating that XBJZ was more salinity-tolerant than KY131. The SNC, SKC, SNK, and SSR phenotypic values in the 295 rice accessions and RIL lines were normally distributed, demonstrating that these indices are quantitative traits under the control of numerous factors (Figure 2).

GWAS for salinity tolerance-related traits in a natural population

The 788,396 SNPs obtained from previous studies were used for GWAS analysis (Li et al., 2019; Zheng et al., 2022). Manhattan and quantile–quantile plots were shown in Figure 3. Fourteen lead SNPs significantly associated with SNC, SNK, and SSR were provided in Table 1. Three QTLs associated with SNC were detected and located on chromosomes 1, 4, and 12, with R^2 values ranging from 9.20–10.36%. No QTL associated with SKC was detected, and eight QTLs associated with SNK were detected and located on chromosomes 1, 4, 6, 9, 10, and 12, with R^2 values ranging from 9.05–10.30%. Meanwhile, three QTLs associated with SSR were detected and located on chromosomes 4, 8, and 10, with R^2 values ranging from 9.76–12.02% (Table 1).

Linkage mapping for salinity tolerance at the seedling stage

A total of five QTLs associated with SNC, SKC, SNK, and SSR were identified on chromosomes 1, 4, and 12 using linkage mapping (Table 2; Figure S1), with LOD values from 2.51–8.22 and proportions of phenotypic variation ranged from 5.49–18.27%. In addition, *qSKC12* and *qSNK12* were considered to be the same QTL because of the same interval, named *qSK12* (chromosome 12) (Figure 4B), located in the physical region between markers C12_20029364 and C12_20873254 and explaining 11.54–18.27% of the phenotypic variation. The GWAS identified the lead SNPs, Chr12_20864157, associated with an SNK, which the linkage mapping detected as being in *qSK12*. The LD block region on chromosome 12 was predicted to be 20.718–20.905 Mb (186 kb) (Figure 4A). A 195-kb overlap region was filtered out based on the GWAS and the linkage mapping (Figure 4C).

Haplotype analysis of candidate genes

According to the *Phytozome* database, the 195-kb candidate interval did not include known salinity tolerance genes from previous studies. The 195-kb region on chromosome 12 contained 35 genes (Figure 4D; Table S2). We performed haplotype analysis of 35 genes and found that 4 genes (*LOC_Os12g34320*, *LOC_Os12g34330*, *LOC_Os12g34450*, and *LOC_Os12g34460*) within the overlapping interval were associated with significantly different haplotypes of SNKs. *LOC_Os12g34320* and *LOC_Os12g34450* were classified into two haplotypes by non-synonymous mutant SNPs in the exon region. One SNP located in the 5′ untranslated region of *LOC_Os12g34330* formed two haplotypes, while no non-synonymous SNP was found in the exon region. The haplotype analysis of *LOC_Os12g34460* indicated that 2 SNPs were identified in the promoter region and

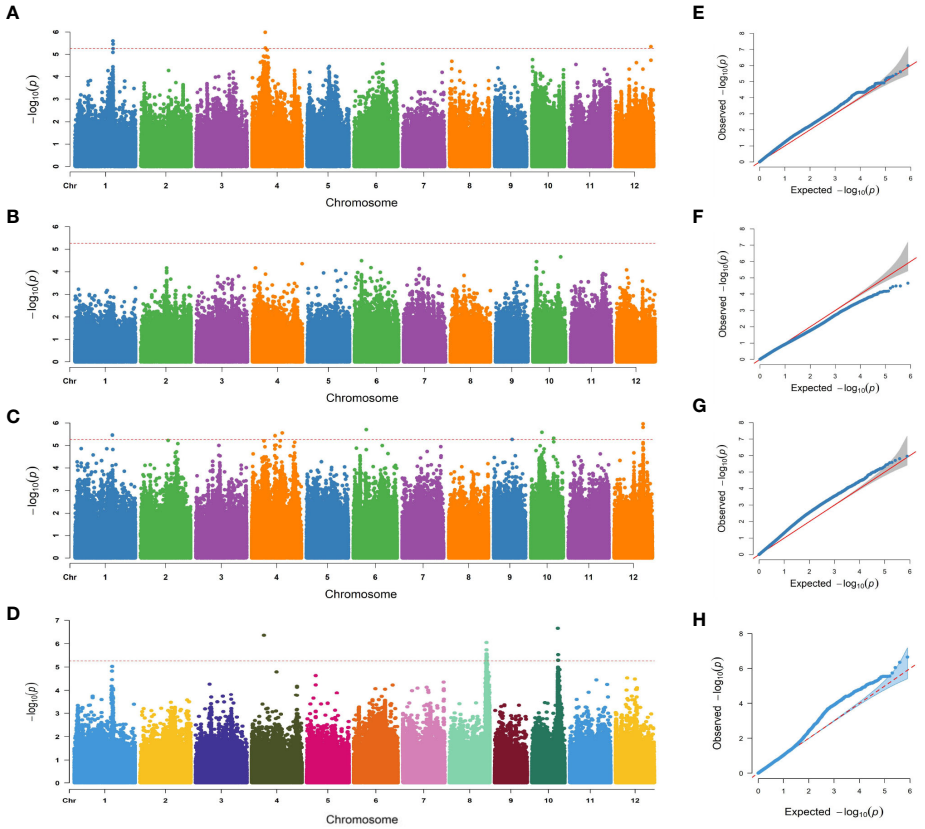


FIGURE 3
Manhattan plots and quantile-quantile (Q-Q) plots of GWAS for the SNC, SKC, SNK, and SSR. **(A–D)** Manhattan plot for the SNC, SKC, SNK, and SSR. **(E–H)** Q-Q plot for the SNC, SKC, SNK, and SSR.

TABLE 1 Lead SNPs for SNC, SNK, and SSR identified by GWAS.

Traits	Lead SNP	Chromosome	Position	P value	R^2(%)	Known QTLs	Known genes
SNC	Chr1:27088613	1	27088613	2.48E-06	9.66	*qNAUP-1a* (Sabouri and Sabouri, 2008)	
	Chr4:9305690	4	9305690	1.02E-06	10.36	*qPn4b* (Tong et al., 2006)	
	Chr12:25487805	12	25487805	4.49E-06	9.20		
SNK	Chr1:26805021	1	26805021	3.44E-06	9.36	*qNAUP-1a* (Sabouri and Sabouri, 2008)	*OSBZ8* (Mukherjee et al., 2006)
	Chr4:16679366	4	16679366	3.70E-06	9.46	*qPn4b* (Tong et al., 2006)	
	Chr4:21840798	4	21840798	2.77E-06	9.57	*qDTF4.1s* (Mohammadi et al., 2013)	*OsCLC-1* (Nakamura et al., 2006)
	Chr6:9215769	6	9215769	1.99E-06	9.83		
	Chr9:13269677	9	13269677	5.31E-06	9.05		
	Chr10:7916859	10	7916859	2.61E-06	9.78	*SalTol10-1* (Islam et al., 2011)	
	Chr10:16340314	10	16340314	4.70E-06	9.32	*SalTol*10-1 (Islam et al., 2011)	

(Continued)

TABLE 1 Continued

Traits	Lead SNP	Chromosome	Position	P value	R^2(%)	Known QTLs	Known genes
	Chr12:20864157	12	20864157	1.10E-06	10.30	*QSst12* (Cheng et al., 2012)	
SSR	Chr4:8601162	4	8601162	4.40E-07	9.76	*qPn4b* (Tong et al., 2006)	
	Chr8:26210079	8	26210079	8.94E-07	10.68	*qGY8.1s* (Mohammadi et al., 2013)	
	Chr10:18354286	10	18354286	2.20E-07	12.02	*qSKC-10b* (Ghomi et al., 2013)	

R^2 (%): Phenotypic variance explained.

TABLE 2 QTLs for SNC, SKC, SNK, and SSR identified by linkage mapping analysis in 195 RILs.

Traits	QTLs	Left Marker	Right Marker	Chr.	LOD	R^2(%)	Additive effect	Known QTLs	Known genes
SNC	*qSNC12*	C12_23512814	C12_24090250	12	2.51	6.17	1.03	*qSH12.1* (Wang et al., 2012)	*OsMYB91* (Zhu et al., 2015)
SKC	*qSKC12*	C12_20029364	C12_20873254	12	8.22	18.27	-3.50	*QSst12* (Cheng et al., 2012)	
SNK	*qSNK12*	C12_20029364	C12_20873254	12	4.99	11.54	-0.15	*QSst12* (Cheng et al., 2012)	
SSR	*qSSR1*	C1_33960158	C1_34422650	1	3.18	5.49	-0.05	*qPH1.1s* (Mohammadi et al., 2013)	
	qSSR4	C4_22587609	C4_25541936	4	4.28	8.88	0.06	*qDTF4.1s* (Mohammadi et al., 2013)	*OsNAC2* (Shen et al., 2017)

R^2 (%): Phenotypic variance explained.

4 non-synonymous SNPs were detected in the exon region (Figures 5A–D).

Identification of candidate genes by gene expression and sequence analysis

The two parents, KY131 and XBJZ, were treated with 120 mM NaCl for 0, 1, 3, 6, 12, and 24 h, respectively. The expression patterns of these four genes were assessed by qRT-PCR analysis. The mean results in triplicate were shown in Figure 6. Among these four genes, salinity stress did not increase the expression levels of *LOC_Os12g34320*, *LOC_Os12g34330*, or *LOC_Os12g34460* (Figures 6A, B, D). *LOC_Os12g34450* was significantly induced by salinity stress, and the expression levels were significantly different between KY131 and XBJZ, which showed opposite expression patterns. The expression level of *LOC_Os12g34450* was more than 18-fold higher in KY131 than in XBJZ under 6 h of salinity stress (Figure 6C). Meanwhile, we performed qRT-PCR analysis of other functionally annotated genes within the 195-kb region, and the results of the four genes expressed between the two parents are shown in Figure S2, which were not differentially expressed between the two parents. After observing these results, we completed further sequencing of the promoter regions and genes of *LOC_O12g34450* in the two parents KY131 and XBJZ. Compared with the sequence of KY131, *LOC_Os12g34450* of XBJZ had a 2-bp deletion (A and C bases) and 2 SNPs (T→G, G→A) in the first exon of the CDS region, 1 SNP (T→C) in the third exon, and multiple SNPs in the promoter region containing 1 cis-element (TGA-element) related to auxin response (AACGAC→GACGAT) (Figure S3). Therefore, *LOC_Os12g34450* was considered a functional gene associated with salinity tolerance. *LOC_Os12g34450* encodes an auxin-binding protein 4 precursor gene, which has not been reported to affect salinity tolerance in rice in previous studies. In subsequent studies, transgenic plants will be examined for the biological function of their associated genomic variant's responses to salinity tolerance in *Japonica* rice.

Discussion

Rice seedlings have poor salinity tolerance and are susceptible to salt stress, which seriously affects their planting and yield. Therefore, cultivating salinity tolerance in seedlings has great practical significance for agricultural production (Nam et al., 2015). This study selected SNC, SKC, SNK, and SSR as indicators to assess salinity tolerance in rice seedlings, which have also been used in previous studies (Rahman et al., 2016). Under salinity stress conditions, the Na^+ content in rice seedling shoots and roots increases, and the K^+ content decreases. This chemical change leads to a higher Na^+/K^+ ratio and disrupts the ionic balance by decreasing Mg, Zn, and Mn content (Tuncturk et al., 2008), also leading to osmotic stress and growth inhibition (Munns, 2011).

Salinity tolerance is a complex trait, and identifying such QTLs helps obtain relative salt tolerance genes for molecular-assisted

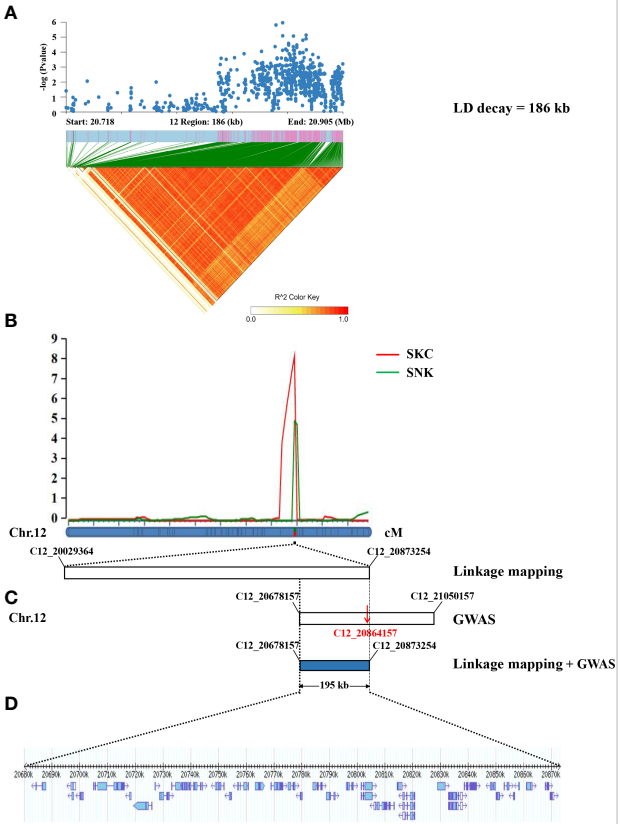


FIGURE 4
Identification of candidate genes by GWAS and linkage mapping. **(A)** The local Manhattan plot (top) and LD heatmap (bottom) surround the lead SNP. **(B)** Salinity tolerance-related QTLs were identified in 195 RILs and mapped to the interval between markers C12_20029364 and C12_20873254 by linkage mapping. **(C)** The physical location of the lead SNP (C12_20864157) on chromosome 12 was detected by the GWAS (LD decay = 186 kb). **(D)** The 195-kb region contained 35 genes.

breeding (Ganie et al., 2019). Eleven and three QTLs were identified by GWAS and linkage mapping, respectively, and these were close to or overlapped with the loci of some known genes and QTLs compared with the results of previous studies. For example, *OsCLC-1*, which encodes a voltage-gated chloride channel protein, could avoid ion toxicity by transporting chloride ions across the vesicle membrane to the vesicles (Nakamura et al., 2006); the lead SNP Chr4_21840798 identified by GWAS was approximately 45 kb closer to *OsCLC-1*. *OsMYB9*, encoding an R2R3-type MYB transcription factor that functions in the ABA-mediated signaling pathway, thereby improving the salinity tolerance of rice (Zhu et al., 2015), was within the *qSNC12* identified by the linkage mapping. Additionally, Shen et al. (2017) found that *OsNAC2*, was able to regulate abiotic stress response, *OsNAC2* overexpression plants showed reduced tolerance under salt stress, and *OsNAC2* was within the *qSSR4* identified by the linkage mapping.

Sabouri and Sabouri (2008) detected a salinity tolerance-associated QTL (*qNAUP-1a*) on chromosome 1, and the lead SNPs (Chr1_27088613 and Chr1_26805021), detected by GWAS, were located within *qNAUP-1a*. Islam et al. (2011) detected a salinity tolerance-related QTL (*SalTol10-1*) on chromosome 10, and the lead SNPs Chr10_7916859 and Chr10_16340314 detected by GWAS were both located within *SalTol10-1*. In addition, linkage mapping revealed that the QTL *qSNC12* was within the salinity tolerance interval *qSH12.1* (Wang et al., 2012), which supported our findings. In our study, GWAS and linkage mapping identified another salinity tolerance QTL, *QSst12*, containing Chr12_20864157 and *qSK12*.

The *LOC_Os12g34450* was considered the most likely functional gene associated with salinity tolerance. The expression of *LOC_Os12g34450* was significantly up-regulated in XBJZ and down-regulated in KY131 after salinity stress. Meanwhile, we found that *LOC_Os12g34450* of XBJZ contained multiple SNPs in the promoter region, including a cis-element (TGA-element) related to auxin response (AACGAC→GACGAT), three non-synonymous

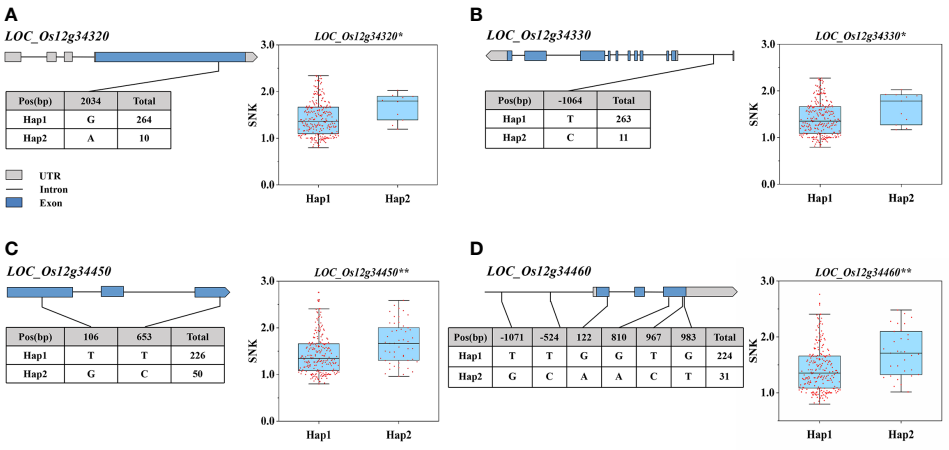


FIGURE 5
Haplotype analysis of *LOC_Os12g34320*, *LOC_Os12g34330*, *LOC_Os12g34450*, *LOC_Os12g34460*. **(A–D)** Represent the gene structure and haplotype analysis of *LOC_Os12g34320*, *LOC_Os12g34330*, *LOC_Os12g34450*, *LOC_Os12g34460*. (The * and ** suggest significance of ANOVA at $P < 0.05$ and $P < 0.01$, respectively).

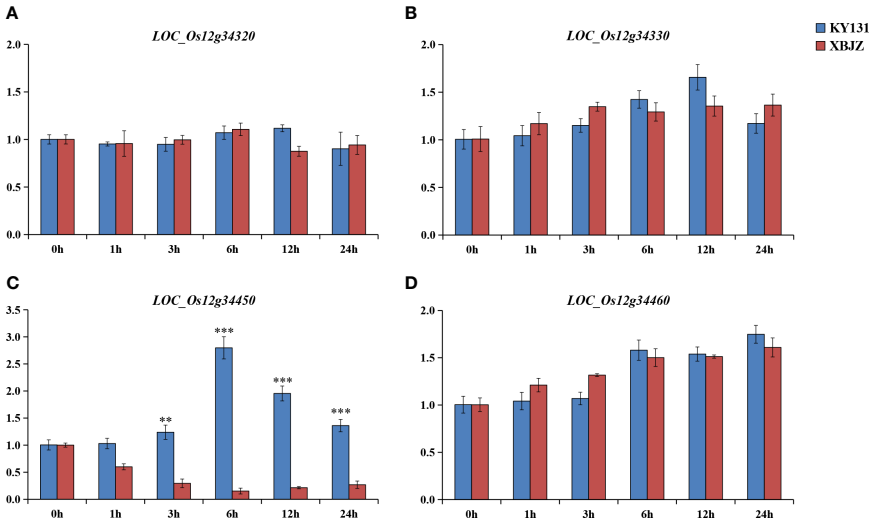


FIGURE 6
Expression patterns of the four genes under normal growth conditions and salinity stress. **(A–D)** represent the gene expression of LOC_Os12g34320, LOC_Os12g34330, LOC_Os12g34450, LOC_Os12g3446 under normal growth conditions and salinity stress. (**P < 0.01, ***P < 0.001, Students't-test).

mutation SNPs (T→G, G→A, T→C) in the CDS region, and a 2-bp deletion (A and C bases) in the first exon compared with the sequence of KY131. These SNPs and deletions result in a frameshift that leads to a premature stop codon, which leads to premature termination of the mRNA of the *LOC_Os12g34450*. *LOC_Os12g34450* is predicted to encode an auxin-binding protein, and accumulating evidence in recent years supports an important role for auxin in abiotic stress responses in plants.

In *Arabidopsis*, the *NTM2* plays an important role in regulating plant seed germination under high-salt stress (Jung and Park, 2011). Overexpression of *OsmiR393* in rice resulted in the down-regulation of two auxin receptor gene homologs (*OsTIR1* and *OsAFB2*), which reduced salinity tolerance in rice (Xia et al., 2012). Deng et al. (2022) found that *RST1* encodes the growth factor response factor *OsARF18* and that *RST1* loss of function leads to up-regulation of *OsAS1* expression, which improves nitrogen utilization, reduces Na^+/K^+ ratios, and decreases NH_4^+ overaccumulation by promoting asparagine synthesis, thereby improving salt tolerance and yield of plants.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

SX, DZ, and HZ designed the study and provided experimental materials. SX, JC, HC, SL, and TM performed the experiments. SX, JC, and HZ analyzed the results and wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the National Natural Science Foundation of China (grant No. U20A2025, 31872884), and the "Academic Backbone" Project of Northeast Agricultural University (grant No. 20XG24).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1184416/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Genetic linkage map and QTL mapping results.

SUPPLEMENTARY FIGURE 2
Expression patterns of the other four genes under normal growth conditions and salinity stress (** P < 0.01, *** P < 0.001, Students't-test).

SUPPLEMENTARY FIGURE 3
The gene structure and sequence difference analysis of *LOC_Os12g34450*. **(A)** The gene structure of *LOC_Os12g34450*. **(B)** The sequence difference analysis of *LOC_Os12g34450* in Xiaobaijingzi, Kongyu131, and Nipponbare.

References

Bimpong, I. K., Manneh, B., Diop, B., Ghislain, K., Sow, A., Amoah, N. K. A., et al. (2014). New quantitative trait loci for enhancing adaptation to salinity in rice from hasawi, a Saudi landrace into three African cultivars at the reproductive stage. *Euphytica* 200 (1), 45–60. doi: 10.1007/s10681-014-1134-0

Cheng, L. R., Wang, Y., Meng, L. J., Hu, X., Cui, Y. R., Sun, Y., et al. (2012). Identification of salt-tolerant QTLs with strong genetic background effect using two sets of reciprocal introgression lines in rice. *Genome* 55 (1), 45–55. doi: 10.1139/g11-075

Deng, P., Jing, W., Cao, C., Sun, M., Chi, W., Zhao, S., et al. (2022). Transcriptional repressor RST1 controls salt tolerance and grain yield in rice by regulating gene expression of asparagine synthetase. *Proc. Natl. Acad. Sci. U. S. A.* 119 (50), e2210338119. doi: 10.1073/pnas.2210338119

Dong, S. S., He, W. M., Ji, J. J., Zhang, C., and Yang, T. L. (2021). LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Brief Bioinform.* 22 (4), bbaa227. doi: 10.1093/bib/bbaa227

Duan, Y. X., Zheng, H. L., Wen, H. R., Qu, D., Cui, J. N., Li, C., et al. (2022). Identification of candidate genes for salt tolerance at the germination stage in japonica rice by genome-wide association analysis. *Agriculture-Basel* 12 (10), 15. doi: 10.3390/agriculture12101588

Ganie, S. A., Molla, K. A., Henry, R. J., Bhat, K. V., and Mondal, T. K. (2019). Advances in understanding salt tolerance in rice. *Theor. Appl. Genet.* 132 (4), 851–870. doi: 10.1007/s00122-019-03301-8

Ghomi, K., Rabiei, B., Sabouri, H., and Sabouri, A. (2013). Mapping QTLs for traits related to salinity tolerance at seedling stage of rice (Oryza sativa l.): an agrigenomics study of an Iranian rice population. *OMICS* 17 (5), 242–251. doi: 10.1089/omi.2012.0097

Huang, X. Y., Chao, D. Y., Gao, J. P., Zhu, M. Z., Shi, M., and Lin, H. X. (2009). A previously unknown zinc finger protein, DST, regulates drought and salt tolerance in rice via stomatal aperture control. *Genes Dev.* 23 (15), 1805–1817. doi: 10.1101/gad.1812409

Islam, M. R., Salambr, M., Hassanbr, L., Collardbr, B., and Gregorio, B. (2011). QTL mapping for salinity tolerance at seedling stage in rice. *Journal of the Science of Food and Agriculture* emir. *J. Food Agric.* 23 (2), 137–146(2). doi: 10.9755/ejfa.v23i2.6348

Jung, J.-H., and Park, C.-M. (2011). Auxin modulation of salt stress signaling in arabidopsis seed germination. *Plant Signal Behav.* 6 (8), 1198–1200. doi: 10.4161/psb.6.8.15792

Li, C. H., Wang, G., Zhao, J. L., Zhang, L. Q., Ai, L. F., Han, Y. F., et al. (2014). The receptor-like kinase SIT1 mediates salt sensitivity by activating MAPK3/6 and regulating ethylene homeostasis in rice. *Plant Cell* 26 (6), 2538–2553. doi: 10.1105/tpc.114.125187

Li, N., Zheng, H. L., Cui, J. N., Wang, J. G., Liu, H. L., Sun, J., et al. (2019). Genome-wide association study and candidate gene analysis of alkalinity tolerance in japonica rice germplasm at the seedling stage. *Rice* 12, 12. doi: 10.1186/s12284-019-0285-y

Li, X. W., Zheng, H. L., Wu, W. S., Liu, H. L., Wang, J. G., Jia, Y., et al. (2020). QTL mapping and candidate gene analysis for alkali tolerance in japonica rice at the bud stage based on linkage mapping and genome-wide association study. *Rice* 13 (1), 11. doi: 10.1186/s12284-020-00412-5

Li, Y. X., Zhou, J. H., Li, Z., Qiao, J. Z., Quan, R. D., Wang, J., et al. (2022). SALT AND ABA RESPONSE ERF1 improves seed germination and salt tolerance by repressing ABA signaling in rice. *Plant Physiol.* 189 (2), 1110–1127. doi: 10.1093/plphys/kiac125

Liang, J. L., Qu, Y. P., Yang, C. G., Ma, X. D., Cao, G. L., Zhao, Z. W., et al. (2015). Identification of QTLs associated with salt or alkaline tolerance at the seedling stage in rice under salt or alkaline stress. *Euphytica* 201 (3), 441–452. doi: 10.1007/s10681-014-1236-8

Lin, H. X., Zhu, M. Z., Yano, M., Gao, J. P., Liang, Z. W., Su, W. A., et al. (2004). QTLs for na^+ and k^+ uptake of the shoots and roots controlling rice salt tolerance. *Theor. Appl. Genet.* 108 (2), 253–260. doi: 10.1007/s00122-003-1421-y

Ma, C. Q., Wang, Y. G., Gu, D., Nan, J. D., Chen, S. X., and Li, H. Y. (2017). Overexpression of s-Adenosyl-L-Methionine synthetase 2 from sugar beet M14 increased arabidopsis tolerance to salt and oxidative stress. *Int. J. Mol. Sci.* 18 (4), 16. doi: 10.3390/ijms18040847

Mohammadi, R., Mendioro, M. S., Diaz, G. Q., Gregorio, G. B., and Singh, R. K. (2013). Mapping quantitative trait loci associated with yield and yield components under reproductive stage salinity stress in rice (Oryza sativa l.). *J. Genet.* 92 (3), 433–443. doi: 10.1007/s12041-013-0285-4

Mukherjee, K., Choudhury, A. R., Gupta, B., Gupta, S., and Sengupta, D. N. (2006). An ABRE-binding factor, OSBZ8, is highly expressed in salt tolerant cultivars than in salt sensitive cultivars of indica rice. *BMC Plant Biol.* 6, 14. doi: 10.1186/1471-2229-6-18

Munns, R. (2011). "Plant adaptations to salt and water stress: differences and commonalities," in *Plant responses to drought and salinity stress: developments in a post-genomic era*. Ed. I. Turkan (London: Academic Press Ltd-Elsevier Science Ltd), 1–32.

Nakamura, A., Fukuda, A., Sakai, S., and Tanaka, Y. (2006). Molecular cloning, functional expression and subcellular localization of two putative vacuolar voltage-gated chloride channels in rice (Oryza sativa l.). *Plant Cell Physiol.* 47 (1), 32–42. doi: 10.1093/pcp/pci220

Nam, M. H., Bang, E., Kwon, T. Y., Kim, Y., Kim, E. H., Cho, K., et al. (2015). Metabolite profiling of diverse rice germplasm and identification of conserved metabolic markers of rice roots in response to long-term mild salinity stress. *Int. J. Mol. Sci.* 16 (9), 21959–21974. doi: 10.3390/ijms160921959

Ouhibi, C., Attia, H., Rebah, F., Msilini, N., Chebbi, M., Aarrouf, J., et al. (2014). Salt stress mitigation by seed priming with UV-c in lettuce plants: growth, antioxidant activity and phenolic compounds. *Plant Physiol. Biochem.* 83, 126–133. doi: 10.1016/j.plaphy.2014.07.019

Qadir, M., Quillerou, E., Nangia, V., Murtaza, G., Singh, M., Thomas, R. J., et al. (2014). Economics of salt-induced land degradation and restoration. *Natural Resour. Forum* 38 (4), 282–295. doi: 10.1111/1477-8947.12054

Rahman, A., Nahar, K., Hasanuzzaman, M., and Fujita, M. (2016). Calcium supplementation improves Na^+/K^+ ratio, antioxidant defense and glyoxalase systems in salt-stressed rice seedlings. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00609

Ren, Z. H., Gao, J. P., Li, L. G., Cai, X. L., Huang, W., Chao, D. Y., et al. (2005). A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat. Genet.* 37 (10), 1141–1146. doi: 10.1038/ng1643

Sabouri, H., and Sabouri, A. (2008). New evidence of QTLs attributed to salinity tolerance in rice. *Afr. J. Biotechnol.* 7 (24), 4376–4383. doi: 10.5897/AJB08.667

Shen, J. B., Lv, B., Luo, L. Q., He, J. M., Mao, C. J., Xi, D. D., et al. (2017). The NAC-type transcription factor OsNAC2 regulates ABA-dependent genes and abiotic stress tolerance in rice (vol 7, 40641, 2017). *Sci. Rep.* 7, 2. doi: 10.1038/srep46890

Tang, Z. B., Gao, X. Y., Zhan, X. Y., Fang, N. Y., Wang, R. Q., Zhan, C. F., et al. (2021). Natural variation in OsGASR7 regulates grain length in rice. *Plant Biotechnol. J.* 19 (1), 14–16. doi: 10.1111/pbi.13436

Tong, H. H., Mei, H. W., Yu, X. Q., Xu, X. Y., Li, M. S., Zhang, S. Q., et al. (2006). Identification of related QTLs at late developmental stage in rice (Oryza sativa l.) under two nitrogen levels. *Yi Chuan Xue Bao.* 33 (5), 458–467. doi: 10.1016/s0379-4172(06)60073-5

Tuncturk, M., Tuncturk, R., and Yasar, F. (2008). Changes in micronutrients, dry weight and plant growth of soybean (Glycine max l. Merrill) cultivars under salt stress. *Afr. J. Biotechnol.* 7 (11), 1650–1654. doi: 10.5897/AJB08.248

Wang, Z. F., Cheng, J. P., Chen, Z. W., Huang, J., Bao, Y. M., Wang, J. F., et al. (2012). Identification of QTLs with main, epistatic and QTL x environment interaction effects for salt tolerance in rice seedlings under different salinity conditions. *Theor. Appl. Genet.* 125 (4), 807–815. doi: 10.1007/s00122-012-1873-z

Wu, X., Li, Y. X., Shi, Y. S., Song, Y. C., Zhang, D. F., Li, C. H., et al. (2016). Joint-linkage mapping and GWAS reveal extensive genetic loci that regulate male inflorescence size in maize. *Plant Biotechnol. J.* 14 (7), 1551–1562. doi: 10.1111/pbi.12519

Xia, K. F., Wang, R., Ou, X. J., Fang, Z. M., Tian, C. G., Duan, J., et al. (2012). OsTIR1 and OsAFB2 downregulation via OsmiR393 overexpression leads to more tillers, early flowering and less tolerance to salt and drought in rice. *PLoS One* 7 (1), 364–373. doi: 10.1371/journal.pone.0030039

Yu, J., Zao, W. G., He, Q., Kim, T. S., and Park, Y. J. (2017). Genome-wide association study and gene set analysis for understanding candidate genes involved in salt tolerance at the rice seedling stage. *Mol. Genet. Genomics* 292 (6), 1391–1403. doi: 10.1007/s00438-017-1354-9

Zhang, W., Liao, X. L., Cui, Y. M., Ma, W. Y., Zhang, X. N., Du, H. Y., et al. (2019). A cation diffusion facilitator, GmCDF1, negatively regulates salt tolerance in soybean. *PloS Genet.* 15 (1), 27. doi: 10.1371/journal.pgen.1007798

Zheng, H. L., Sun, S. C., Bai, L. M., Jiang, S. K., Ding, G. H., Wang, T. T., et al. (2022). Identification of candidate genes for panicle length in oryza sativa l. ssp. japonica via genome-wide association study and linkage mapping. *Euphytica* 218 (2), 12. doi: 10.1007/s10681-022-02972-7

Zhou, Y. B., Liu, C., Tang, D. Y., Yan, L., Wang, D., Yang, Y. Z., et al. (2018). The receptor-like cytoplasmic kinase STRK1 phosphorylates and activates CatC, thereby regulating H_2O_2 homeostasis and improving salt tolerance in rice. *Plant Cell* 30 (5), 1100–1118. doi: 10.1105/tpc.17.01000

Zhu, N., Cheng, S. F., Liu, X. Y., Du, H., Dai, M. Q., Zhou, D. X., et al. (2015). The R2R3-type MYB gene OsMYB91 has a function in coordinating plant growth and salt stress tolerance in rice. *Plant Sci.* 236, 146–156. doi: 10.1016/j.plantsci.2015.03.023

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 19 June 2023
DOI 10.3389/fpls.2023.1206165

OPEN ACCESS

EDITED BY
Xueyong Li,
Chinese Academy of Agricultural Sciences, China

REVIEWED BY
Yongzhong Xing,
Huazhong Agricultural University, China
Jun Fang,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE
Yuchun Rao
ryc@zjnu.cn
Yijian Mao
yijian891001@163.com

RECEIVED 15 April 2023
ACCEPTED 31 May 2023
PUBLISHED 19 June 2023

CITATION
Zhong Q, Jia Q, Yin W, Wang Y, Rao Y and Mao Y (2023) Advances in cloning functional genes for rice yield traits and molecular design breeding in China. *Front. Plant Sci.* 14:1206165. doi: 10.3389/fpls.2023.1206165

Advances in cloning functional genes for rice yield traits and molecular design breeding in China

Qianqian Zhong[1], Qiwei Jia[1], Wenjing Yin[1], Yuexing Wang[2], Yuchun Rao[1]* and Yijian Mao[2]*

[1]College of Life Sciences, Zhejiang Normal University, Jinhua, Zhejiang, China, [2]State Key Laboratory of Rice Biology and Breeding, China National Rice Research Institute, Hangzhou, China

Rice, a major food crop in China, contributes significantly to international food stability. Advances in rice genome sequencing, bioinformatics, and transgenic techniques have catalyzed Chinese researchers' discovery of novel genes that control rice yield. These breakthroughs in research also encompass the analysis of genetic regulatory networks and the establishment of a new framework for molecular design breeding, leading to numerous transformative findings in this field. In this review, some breakthroughs in rice yield traits and a series of achievements in molecular design breeding in China in recent years are presented; the identification and cloning of functional genes related to yield traits and the development of molecular markers of rice functional genes are summarized, with the intention of playing a reference role in the following molecular design breeding work and how to further improve rice yield.

KEYWORDS

rice, genes, yield traits, molecular markers, molecular design breeding

Introduction

With the development of urbanization and population growth, people have new demands for material living standards. As one of the main food crops for human survival, rice has new requirements in production, quality, and plant breeding (Zeng et al., 2017a). In modern breeding history, the effective utilization of dwarfing and heterosis resulted in two dramatic leaps in crop yield per unit area. Since the beginning of the 21st century, China's rice production has faced serious challenges. As the world population's need for higher grain yield and superior quality escalates, building upon our current successes to further enhance rice yield becomes essential. This endeavor necessitates a comprehensive analysis of the intricate regulatory networks governing genes associated with major yield traits in rice. Utilization of key yield-related genes, or quantitative trait loci (QTLs), coupled with the discovery of additional genes with the potential to increase rice yield, is crucial. By doing so, we can devise innovative

breeding strategies that circumvent traditional breeding limitations, propelling our efforts toward meeting the escalating demand. The application of molecular marker-assisted selection in rice breeding has been conducted for many years in China, leading to the construction of multiple chromosome segment substitution lines (CSSLs) and the identification of many QTLs. With the completion of rice genome sequencing and the mapping and isolation of important yield genes, significant progress has also been made in the functional analysis of these genes, providing the foundation for the application of molecular design in rice breeding (Zhou Y. et al., 2017).

In the practice of genetic improvement of crop yield-related traits, the breakthrough in yield potential is to make new progress in plant type while ensuring a high grain weight. Specifically, crop yield can be increased by improving panicle morphology, tiller traits, leaf shape, and angle, in addition to increasing the field planting density of plants to improve the light energy utilization rate, thus changing the number of grains per panicle and 1000-grain weight to increase the crop yield. Therefore, the development of plant architecture and grain formation are key traits for current and future high-yield breeding, and the analysis of genetic regulatory networks provides an important basis for improving these traits. In the case of rice, the analysis of rice plant architecture development (e.g., tiller number, plant height, tiller angle, leaf shape, stem and leaf angle, and panicle type) and grain formation (e.g., flower/panicle completion and grain development), which affect rice yield and the molecular genetics, physiological, and biochemical regulatory networks of important biological processes closely related to them, proposes the molecular design and breeding theory of major rice yield traits, which can provide support for the cultivation of high-yield varieties of rice in China. Chinese scientists have made a series of achievements in the cultivation of high-yield and high-quality varieties of rice. In particular, the research achievements of Jiayang Li, Bin Han, and Qian Qian's team, "Molecular mechanism and variety design of the formation of high-yield and high-quality traits of rice" won the first prize of the State Natural Science Award in 2017 - another breakthrough after the "green revolution" and "hybrid rice" in China, which is of great significance to the promotion of the development of molecular design breeding in China (Chen F. et al., 2018).

The current study classifies and summarizes the gene cloning and molecular design breeding of various important yield traits of rice in China in recent years and provides a prospective investigation on the future direction of molecular breeding of rice. Given the numerous accomplishments of Chinese researchers in rice genomics, there may have been unintentional oversights in compiling this review. This review of genes and gene regulatory networks is summarized as an example.

Research advances in rice yield traits

The complex quantitative traits that determine rice yield are governed by numerous genes and highly sensitive to environmental variables. The key determinants of rice yield encompass plant morphology, panicle structure, and grain type. These yield characteristics mutually supplement each other to collectively determine the overall yield. Research teams in China have discovered and cloned a series of genes associated with these yield traits and shed light on the gene regulatory networks that influence these traits (Figure 1), substantially advancing the progress of rice research in the country.

Cloning and genetic network analysis of key rice plant type genes

Rice plant type traits mainly include tiller number, tiller angle, and angle between stem and leaf, which are important agronomic

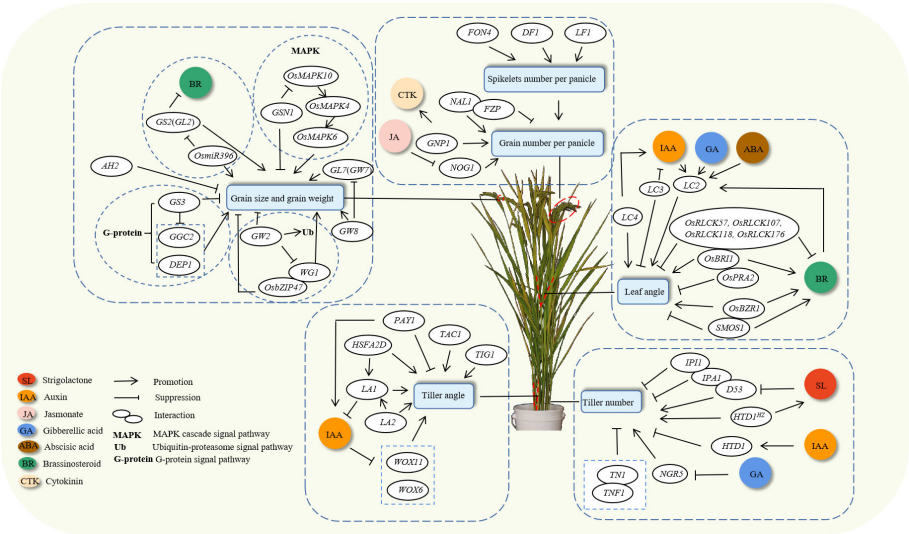


FIGURE 1
Gene regulatory network for yield traits.

traits that determine rice yield. Chinese breeders have carried out extensive research on the idealization of rice plant type, cloned a series of key genes that regulate plant type (Table 1), constructed relevant genetic regulatory networks, and cultivated many super rice varieties with ideal plant type characteristics.

Regulatory network for rice tiller numbers

The quantity of tillers in rice plants is a primary factor influencing yield. Moreover, it is a crucial agronomic attribute that determines tolerance to density and resistance to lodging. It has been demonstrated that the gene *Ideal Plant Architecture1* (*IPA1*), one of the major rice genes responsible for ideal plant architecture, encodes a transcription factor called OsSPL14 (SOUAMOSA PROMOTER BINDING PROTEIN-LIKE 14), which is involved in the regulation of multiple growth and developmental processes (Jiao et al., 2010). In-depth research on the interaction protein system of the IPA1 protein revealed that it associates with IPA1 INTERACTING PROTEIN1 (IPI1), a nuclear-localized RING-finger E3 ligase that can ubiquitinate the IPA1 protein. Haiyang Wang, affiliated with the Chinese Academy of

TABLE 1 Representative functional genes for the characteristics of rice plants.

Trait	Gene	Coding product	Regulatory phenotype	Reference
Plant type	*IPA1*	transcription factor OsSPL14	ideotype	(Jiao et al., 2010)
	IPI1	RING-finger E3 ligase	tiller number and panicle size	(Wang and Wang, 2017)
	NGR5	AP2-domain transcription factor	tiller number	(Wu et al., 2020)
	TN1	functional protein with a BAH domain and an RNA recognition domain	tiller number	(Zhang et al., 2023)
	D53	a substrate of the SCFD3 ubiquitination complex	tiller number	(Jiang et al., 2013; Zhou et al., 2013)
	HTD1	carotenoid cleavage dioxygenase 7	tiller number	(Zou et al., 2006)
	$HTD1^{HZ}$	carotenoid cleavage dioxygenase 7	tiller number	(Wang et al., 2020)
	SD1	gibberellin 20-oxidase	plant height	(Wang et al., 2020)
	TAC1	specific proteins of graminaceous plants	tiller angle	(Yu et al., 2007)
	TIG1	TCP transcription factor	tiller angle	(Zhang et al., 2019)
	LA2	a novel chloroplastic protein	tiller angle	(Huang et al., 2021)
	LA1	a novel and unique protein	tiller angle	(Li et al., 2007)
	HSFA2D	heat shock transcription factor	tiller angle	(Zhang et al., 2018)
	WOX6	WUSCHEL-related homeobox domain-containing protein	tiller angle	
	WOX11	WUSCHEL-related homeobox domain-containing protein	tiller angle	
	PAY1	containing a peptidase S64 domain	tiller number and tiller angle	(Zhao et al., 2015)
	OsBZR1	BR-signaling factor	leaf angle	(Qiao et al., 2017)
	SMOS1	AP2-domain transcription factor	leaf angle	
	OsPRA2	small GTP-binding protein	leaf angle	(Zhang et al., 2016)
	OsBRI1	BR receptor kinase	leaf angle	
	OsRLCK107	receptor-like cytoplasmic kinase	leaf angle	(Zhou et al., 2016)
	OsRLCK57	receptor-like cytoplasmic kinase	leaf angle	
	OsRLCK118	receptor-like cytoplasmic kinase	leaf angle	
	OsRLCK176	receptor-like cytoplasmic kinase	leaf angle	
	LC3	a SPOC domain-containing transcription suppressor	leaf angle	(Chen SH. et al., 2018)
	LC4	F-box protein	leaf angle	(Qu et al., 2019)
	LC2	a vernalization-insensitive 3-like protein	leaf angle	(Zhao et al., 2010)

Agricultural Sciences, expressed his thoughts on this gene in the journal Molecular Plant. He indicated, "The loss of *IPI1* function can increase the number of tillers and the size of panicles, which in turn can increase yield. This gene is expected to be applied in variety modification to increase yield by increasing tiller number" (Wang and Wang, 2017). In an article recently published in Cell Research, Reynante et al. argued that *IPI1* is a candidate gene with yield-increasing potential that can improve panicle number without sacrificing the number of tillers (Ordonio and Matsuoka, 2017). Further biochemical analysis has revealed that the ubiquitination of IPA1 protein by IPI1 protein is tissue-specific, allowing fine regulation of IPA1 protein levels in different tissues. The natural allelic combination of different *IPI1* genes results in different levels of *IPA1* expression, which affect the number of tillers in a dose-dependent manner and allow for achieving a reasonably balanced number of tillers and the highest yield potential (Wang et al., 2017). Thus, rice breeders can create plant architecture characteristics that meet specific farming environments and genetic backgrounds to improve the controllability of breeding. In addition, Xiangdong Fu's group screened an EMS mutant library with a background of the semidwarf cultivar 9311 and identified a nitrogen-insensitive gene, *Nitrogen-mediated Tiller Growth Response5* (*NGR5*). This led to significant advances in the understanding of the gibberellin (GA) response to nitrogen in rice. Research has found that *NGR5* positively regulates rice growth and development in response to nitrogen (agronomic traits such as plant height, tillering, and grains per panicle) and can negatively regulate the expression of tillering inhibitory genes, thereby promoting tillering in rice. Further research has found that the NGR5 protein can interact with GIBBERELLIN INSENSITIVE DWARF1 (GID1), which inhibits the growth of rice tillering by promoting the degradation of the NGR5 protein (Wu et al., 2020). This study elucidated the key molecular mechanism of the intersection of the GA signaling pathway and nitrogen allocation and utilization. This finding provides a new target for advancing the high yield and efficiency goals of the "Green Revolution". Professor Zichao Li from the College of Agriculture of China Agricultural University recently cloned a new rice tillering regulatory gene, *Tiller Number1* (*TN1*), which encodes a functional protein containing the BAH and RNA recognition domains. It negatively regulates the establishment of rice tillering numbers. At the same time, the team also found that the natural variation in the TN1 protein would affect its interaction with TN1 INTERACTION FACTOR1 (TIF1). When the *TIF1* gene was knocked out, the number of plant tillers increased. Therefore, *TN1* and *TNF1* have similar regulatory effects on the number of tillers in rice (Zhang et al., 2023). This research provides excellent gene resources and genetic information for designing the ideal type of rice plant. As a new type of hormone, strigolactone (SL) plays a key role in regulating the number of tillers in rice. Jiayang Li's research group and Jianmin Wan's research group published articles on the regulation of SL in rice tiller number in the journal Nature, revealing that *Dwarf 53* (*D53*) is a negative regulatory factor in the SL signaling pathway and that SL can induce *D53* degradation through the proteasome pathway, suppress lateral bud growth, and regulate the number of tillers in rice (Jiang et al., 2013; Zhou et al., 2013). In addition, Song et al. found that *D53* interacts with the ideal plant type gene *IPA1*, jointly regulating the number of rice tillers (Song et al., 2017). Recently, there has been a new development in the study of SL regulation in the number of rice tillers. In the research by Zou et al., the *TilleringI and Dwarf 1* (*HTD1*) gene, which encodes CAROTENOID CLEAVAGE DIOXYGENASE 7 (CCD7), negatively regulates the number of rice tillers, and IAA can upregulate the transcription of this gene to control them (Zou et al., 2006). Further research by Wang et al. has demonstrated a partial loss-of-function allele of the SL biosynthesis gene, $HTD1^{HZ}$, which can increase the number and yield of rice tillers. At present, this gene and the semidwarf gene *Semidwarf 1* (*SD1*) are jointly selected to be widely used, which has to some extent contributed to the green revolution in rice (Wang et al., 2020).

Regulatory network for rice tiller angles

An appropriate tiller angle is critical for the growth of rice and the formation of farmland yield. Recently, several genes controlling rice tiller angle have been identified, and their function in the network regulating rice plant architecture has been analyzed to some extent. *Tiller Angle Control 1 (TAC1)* is an important QTL previously reported for rice tillering angle regulation (Yu et al., 2007). The laboratory of Professor Chuanqing Sun at the China Agricultural University isolated a gene, *TAC1*, corresponding to the main QTL that controls rice tillering angle through positional cloning. The analysis of the plant architecture of *TAC1*-overexpressing and RNAi transgenic lines indicated that the upregulation of *TAC1* resulted in an increased tiller angle, while its downregulation decreased the tillering angle. These findings document that the level of *TAC1* mRNA determines the size of the tillering angle (Jiang et al., 2012). Another research team identified a Tiller Inclined Growth1 (*TIG1*) gene that affects tillering angle in wild rice and positively regulates cell elongation and tillering angle size. It encodes the transcription factors TEOSINTE BRANCHED1, CYCLOIDEA AND PCF (TCP) and is expressed on the paraxial side of the tillering base. After the artificial selection of the indica rice variety *tig1*, the natural variation of its promoter caused a significant decrease in the expression of the *TIG1* gene, resulting in a decrease in cell length and tillering angle (Zhang et al., 2019). This achieved the transition of wild rice from inclined growth to vertical tillering growth and provided strong support for the design of the ideal rice plant type.

Recently, a research group used map-based cloning to identify a gene, known as *LAZY2* (*LA2*), that controls tiller angle. This gene encodes a novel chloroplast protein that specifically governs starch in gravity-sensitive cells. By interacting with the starch biosynthetic enzyme, *Oryza sativa* plastidic phosphoglucomutase (OspPGM), *LA2* regulates the tillering angle, indicating the importance of gravity response in shaping the tiller angle (Huang et al., 2021). Additionally, a group of researchers obtained dynamic transcriptome data by combining RNA-seq technology to study the dynamic development process of rice stem gravity response and excavated important nodal genes that regulate the formation of rice tiller angles, such as *Heat Stress Transcription Factor2D* (*HSFA2D*),

LAZY1 (*LA1*), *Wuschel Related HomeoboX6* (*WOX6*), and *Wuschel Related HomeoboX11* (*WOX11*), and constructed a molecular network for the dynamic regulation of rice tillering angle at the whole genome level. Among them, *LA1* is a previously identified classical tiller angle-regulating gene. Previous studies have confirmed that *LA1* controls the gravitropism of the aboveground portion of the rice shoot and ultimately controls the size of the tillering angle by regulating the polar transport of auxin (IAA) (Li et al., 2007). Genetic evidence confirmed that *LA1* and SL may co-participate in regulating aboveground gravitropism and tillering angle in the auxin transport pathway (Sang et al., 2014). *HSFA2D*, *WOX6*, and *WOX11* are new regulatory genes implicated in the regulatory pathway mediated by *LA1*. *HSFA2D* is located upstream of the *LA1*-mediated auxin pathway and is a positive regulator of this pathway, implicating that *HSFA2D* can cause an unbalanced distribution of auxin by inducing the expression of *LA1*. Auxin promotes the asymmetric expression of the *WOX6* and *WOX11* genes, producing changes in the tiller angle (Zhang et al., 2018). Other genetic analyses show that the *LA2* gene also causes the change in tillering angle by acting upstream of *LA1* to mediate the lateral transport of auxin (Huang et al., 2021). In addition, Zhao et al. have made important progress in the research direction of polar transport of auxin to control the tillering angle of rice using map-based cloning technology. A novel gene, *Plant Architecture and Yield1* (*PAY1*), which controls plant architecture and yield, was cloned, and it was found that its mutation alters the absorption and polar transport capacity of auxin, and then changes the polar distribution of auxin *in vivo*, which is an important reason for the change in tillering angle of *PAY1*. *PAY1* is pleiotropic, and its overexpression enables smaller tillering angles, a lower number of tillers, increased plant height, thicker stalks, increased numbers of panicles, and higher yield (Zhao et al., 2015). This provides new ideas for research in this field by exploring an effective way to mine regulatory genes and regulatory pathways of the rice tiller angle.

Regulatory network for stem-leaf angle in rice

Leaf angle affects the light-receiving posture of plants, and the appropriate angle is beneficial to increase the light-receiving area of the leaves and improve the light energy utilization rate. Presently, the genes regulating leaf angle are mostly related to the synthesis and signaling pathways of brassinolide (BR) and IAA. *Oryza sativa Brassinazole-resistant1* (*OsBZR1*), which encodes the BR-signaling factor, interacts with *Small Organ Size1* (*SMOS1*) to enhance its transcriptional activity. *SMOS1* encodes a transcription factor with an APETALA2 (AP2) DNA-binding domain, reduces the angle between rice stem and leaf, and plays an important role in BR signaling in rice (Qiao et al., 2017). *Oryza sativa Brassinosteroid Insensitive1* (*OsBRI1*) is the orthologue of *Arabidopsis thaliana Brassinosteroid Insensitive1* (*AtBRI1*) (Yamamuro et al., 2000). Several genes have recently been reported to interact with it, including *Oryza sativa Pea Pra2* (*OsPRA2*), *Oryza sativa receptor-like cytoplasmic kinase107* (*OsRLCK107*), *Oryza sativa receptor-like cytoplasmic kinase57* (*OsRLCK57*), *Oryza sativa receptor-like cytoplasmic kinase118* (*OsRLCK118*), and *Oryza sativa receptor-like cytoplasmic kinase176* (*OsRLCK176*). *OsPRA2* encodes a small GTP-binding protein, and its overexpression is characterized by a smaller leaf inclination and reduced sensitivity to exogenous brassinolide (Zhang et al., 2016). *OsRLCK57*, *OsRLCK107*, *OsRLCK118*, *OsRLCK176*, and other *OsRLCKs* encode receptor-like cytoplasmic kinases that interact with the BR receptor *OsBRI1* to suppress BR signaling and affect leaf inclination (Zhou et al., 2016). Chen et al. showed that the *Leaf Inclination 3* (*LC3*) gene, which controls the angle between the stem and leaf, encodes a transcriptional suppressor containing a SPOC domain and plays a role in suppressing IAA homeostasis and signal transduction. In addition to its involvement in regulating the stem-leaf angle mechanism, it also ensures the normal development of lamina joints (Chen SH. et al., 2018). Qu et al. found that microRNAs also play a certain role in regulating the angle between stems and leaves. This study shows that miR394 regulates the angle between stem and leaf by targeting the gene *Leaf Inclination 4* (*LC4*), and IAA will affect the transcription of miR394 and *LC4* (Qu et al., 2019). Other plant hormones such as GA, jasmonic acid (JA), abscisic acid (ABA), and ethylene (ETH) also participate in the regulation of the angle between rice stems and leaves (Li QF. et al., 2019), and all of them achieve the regulatory function by affecting the synthesis and signal transmission of BR (Gan et al., 2015). Zhao et al. obtained a *Leaf Inclination 2* (*LC2*) gene that positively regulates the angle between stem and leaf through positional cloning and encodes a vernalization-insensitive 3-like protein. This gene is mainly expressed at the lamina joint, and its expression is influenced by GA, ABA, IAA, and BR (Zhao et al., 2010).

Cloning and genetic network analysis of key genes for rice panicle type

Panicle type is a key factor in rice yield. Genetic analysis shows that it is controlled by multiple genes, mainly by influencing the number of grains per panicle and panicle morphology (the number of spikelets) to affect rice yield. In recent years, scientists have made great progress in exploring panicle type genes (Table 2) and have analyzed the regulatory network of rice panicle type to some extent.

Regulatory network for rice grain number per panicle

The number of grains per panicle is important in determining rice yield. Wu et al. reported that the *Grain Number per Panicle1* (*GNP1*) gene, which encodes GA 20-oxidase, can increase grain number and yield by upregulating cytokinin (CTK) activity in rice panicle meristems (Wu et al., 2016). Since then, researchers have identified a gene, *Number Of Grains1* (*NOG1*), that affects the number of grains per panicle and rice yield. This gene is located on the long arm of chromosome 1 and encodes enoyl-CoA hydratase of the fatty acid β-oxidation pathway. A 12-bp transcription factor binding site in the *NOG1* promoter region exhibits copy number

TABLE 2 The representative functional genes for rice panicle type traits.

Trait	Gene	Coding product	Regulatory phenotype	Reference
Panicle type	*GNP1*	GA 20-oxidase	grain number per panicle	(Wu et al., 2016)
	NOG1	enoyl-CoA hydratase	grain number per panicle	(Huo et al., 2017)
	FZP	ERF transcription factor	grain number per panicle and 1000 grain weight	(Bai et al., 2017)
	NAL1	trypsin-like serine protease	grain number per panicle	(Huang et al., 2018)
	LF1	HD-ZIP III transcription factor	spikelet number per panicle	(Zhang et al., 2017)
	DF1	lipase	spikelet number per panicle	(Ren et al., 2017)
	FON4	small secreted protein containing the CLE domain	spikelet number per panicle	(Ren et al., 2019b)

variation, which can increase the gene expression level, decrease the concentration of fatty acids and JA in the plant, and increase the number of panicle grains and yield. Introducing *NOG1* into the *NOG1*-deficient cultivar "Zhonghua 17" increases the yield by 25.8%, and overexpression of *NOG1* in the high-yielding cultivar "Teqing" carrying this gene increases the yield by nearly 20% without affecting plant height, heading date, panicle number, grain weight, and other traits. The cloning of *NOG1* provides an important new gene-based strategy for breeding high-yielding rice varieties and offers new insights into the molecular mechanisms underlying the regulation of traits that determine rice yield (Huo et al., 2017). Recently, it has been documented that rice yield is controlled by modulating the expression level of *Frizzy Panicle* (*FZP*). The manipulation of the *FZP* gene in rice leads to a surge in the number of grains per panicle, but with smaller grain sizes, whereas its overexpression triggers the growth of larger grains but decreases the number of grains per panicle (Bai et al., 2017; Ren DY. et al., 2018). The expression level of *FZP* determines the number of grains per panicle and the 1000-grain weight. Fine-tuning the balance between these two parameters to maximize yield by modulating *FZP* expression warrants further exploration. Further functional analysis showed that the trypsin-like serine protease encoded by *Narrow Leaf 1* (*NAL1*) interacts with the FZP protein, promoting its degradation. Reducing the expression of *FZP* or elevating the expression of *NAL1* in the cultivar Zhonghua 17 increases the number of secondary branches and grains per panicle, leading to a higher rice grain yield (Huang et al., 2018).

Regulatory network for rice spikelet number per panicle

Another critical factor influencing the formation of grains per panicle that is often overlooked is the number of spikelets in a panicle. A standard rice panicle comprises two pairs of glumes and one fertile floret, which directly contributes to grain yield. Zhang et al. reported a rice mutant, *lateral florets1* (*lf1*). The spikelet of this mutant produces normal terminal florets, while the lateral floret meristem at the protective glume develops into relatively normal lateral florets. The *LF1* gene is the first gene identified to regulate the growth of lateral florets in rice and encodes the HD-ZIPIII protein; its mutation induces ectopic expression of the meristem maintenance gene (Zhang et al., 2017). This study provides strong evidence for the "three flowered spikelets" hypothesis in rice. Ren et al. identified a rice mutant called *double floret1* (*df1*) that develops an intact floret and produces normal seeds within a pair of nursing spikelets and the normal terminal florets. The *DF1* gene encodes a lipase that regulates the determinacy of the spikelet meristem, and its mutation changes spikelet certainty to uncertainty or prolongs the spikelet determination phase, resulting in the formation of multiple grains (Ren et al., 2017). These findings provide an important basis for research aimed at increasing the number of grains per panicle in rice. In a recent study, Ren et al. discovered an important gene *Floral Organ Number4* (*FON4*), which is involved in inflorescence and floral organ morphogenesis in rice. They identified a novel allelic mutant, *fon4-7*, which forms a lateral floret in addition to normal terminal florets, but different *fon4* allelic mutants have different numbers of florets (Ren et al., 2019b). This study shows that mutations in *FON4* provide the potential to form multiflowered spikelets, which in turn form multiple seeds.

Cloning and genetic network analysis of key genes for rice grain types

Rice grain type is an important trait that determines grain weight and affects rice yield and quality. To date, researchers have cloned a series of genes related to grain type (Table 3), revealing that these genes regulate multiple signaling pathways and have made some progress in balancing rice yield and quality.

Regulatory network between rice grain type and signaling pathways

The latest reports show that the grain type of rice is regulated by signaling pathways such as BR, G-protein, and ubiquitin-proteasome (Li N. et al., 2019; Ren et al., 2023). Hu et al. successfully isolated and cloned an important gene, *Grain Size2* (*GS2*), from a local Zhejiang rice variety, "Baodali", which can significantly improve rice yield. The GS2 protein is localized in the nucleus and regulates the length and width of grains by promoting cell division and expansion (Hu et al., 2015). GS2 directly interacts with the negative regulator of the BR signal transduction pathway,

TABLE 3 The representative functional genes for rice grain type traits.

Trait	Gene	Coding product	Regulatory phenotype	Reference
Grain type	*GS2/GL2*	growth-regulating factor	grain size and length	(Che et al., 2015; Hu et al., 2015)
	OsmiR396	microRNA	grain size	(Duan et al., 2015)
	GS3	heterotrimeric G-protein γ subunit	grain length	(Sun et al., 2018)
	GGC2	heterotrimeric G-protein γ subunit	grain length	
	DEP1	heterotrimeric G-protein γ subunit	grain length	
	GW2	RING-type E3 ubiquitin ligase	grain width and weight	(Song et al., 2007)
	WG1	glutathione	grain size	(Hao et al., 2021)
	OsbZIP47	bZIP transcription factor	grain size	
	GSN1	mitogen-activated protein kinase phosphatase	grain size	(Guo et al., 2018)
	GW7/GL7	*Arabidopsis thaliana* LONGIFOLIA homologous proteins	grain length and width	(Wang SK. et al., 2015; Wang YX. et al., 2015)
	GW8	squamosa promoter binding protein-like	grain length and width	(Wang et al., 2012)
	AH2	MYB family transcription factor	grain size	(Ren et al., 2019a)

and its transcriptional activation activity is inhibited. In addition, it was found that *Grain Length2* (*GL2*) and *GS2* were located at the same locus. The *GL2* gene increased the grain size, and significantly increased the number of tillers and grains per ear, thus increasing the yield per plant (Che et al., 2015). Additionally, the gene *Oryza sativa MicroRNA396* (*OsmiR396*) can cleave *GS2*, but the rare AA allele variation of *GS2* will affect the binding site of *OsmiR396*, resulting in its inability to cleave *GS2*, thus increasing the expression of *GS2*. This indicates that the AA allelic variation of *GS2* can increase grain size (Duan et al., 2015). These studies point to a new strategy that can be used to generate new high-yielding rice cultivars. Gγ subunits in rice, including G PROTEIN GAMMA SUBUNIT3 (GS3), G PROTEIN GAMMA SUBUNIT2 (GGC2), and DENSE PANICLE 1 (DEP1), are involved in the regulation of grain type. Among them, *GS3* combines Gβ protein by competing with *GGC2* and *DEP1* genes to reduce grain length, while *GGC2* and *DEP1* genes are associated with Gβ protein interaction to increase grain length (Sun et al., 2018). In previous studies, the dominant gene at the *DEP1* gene locus was caused by acquired mutations, which can promote cell division, increase the number of branches and grains per panicle, and thus promote an increase in rice yield. This gene has been widely used in rice yield-increasing varieties in China (Huang et al., 2009). Song et al. discovered a gene called *Grain Width2* (*GW2*) that influences rice grain width and weight. This gene encodes an E3 ubiquitin ligase and is involved in the regulation of the ubiquitin-proteasome signaling pathway. Without functional *GW2*, the substrate that should undergo degradation is not specifically identified, leading to an increased cell number and a quicker filling rate. Consequently, this broadens the grain shape and enhances the grain weight of rice, ultimately boosting the overall yield of the rice crop (Song et al., 2007). Recent studies have found that the *GW2-Wide Grain1* (*WG1*)-*Oryza sativa Basic Leucine Zipper47* (*OsbZIP47*) pathway regulates rice seed size. *WG1* encodes glutaredoxin, which controls seed size by promoting cell proliferation, while *OsbZIP47* limits seed growth by reducing cell proliferation. Further research and analysis showed that *GW2* can ubiquitinate *WG1* and degrade it. This study provides a new target for understanding the mechanism of rice grain development and has an important reference value for improving rice yield (Hao et al., 2021). Additionally, some laboratories have shown that the MAPK cascade signaling pathway is an essential regulator of seed size and has potential value for improving crop yield. Xu et al. identified a *grain size and number1 (gsn1)* mutant, *GSN1*, which encodes the mitogen-activated protein kinase phosphatase *Oryza sativa* MAPK PHOSPHATASE1 (OsMKP1). When OsMKP1 loss-of-function results in larger but fewer grains, its overexpression produces smaller but more grains. The additional analysis documented that OsMKP1 interacts with the *Oryza sativa* MAPK PHOSPHATASE6 (OsMAPK6), leading to its dephosphorylation and inactivation. Further research found that OsMKP1 can also inhibit *Oryza sativa* MAPK PHOSPHATASE10 (OsMKKK10) and *Oryza sativa* MAPK PHOSPHATASE4 (OsMKK4*)* (Guo et al., 2018; Xu et al., 2018a; Xu et al., 2018b). Thus, these studies revealed that *GSN1* determines grain size by inhibiting the MAPK signaling pathway and establishing the *GSN1*-MAPK module; it laid the foundation for revealing the mechanism of rice grain size regulation.

Regulatory network for rice grain type and quality

The yield and quality of rice cannot be achieved simultaneously, so balancing the relationship between the two to obtain high-yielding, high-quality varieties is a big challenge. *Grain Length7* (*GL7*) is the major rice gene regulating grain length and width. The *GL7* gene, which is responsible for controlling the shape and quality of rice grains, has been successfully isolated. By modifying the pattern of cell division, *GL7* participates in regulating the longitudinal elongation of cells. Further meticulous genetic analysis demonstrated that a DNA fragment containing 17.1 kb of the *GL7* locus in American long-grain rice varieties underwent

tandem duplication, which increased the expression of the *GL7* gene but also caused the down-regulation of the expression of its adjacent negative regulator. The above two conditions are jointly regulated to cause an increase in grain length, a decrease in chalk content, and chalkiness. These changes improve the appearance and taste of rice (Wang SK. et al., 2015; Wang YX. et al., 2015). It has been documented that an allelic variation of the *Grain Width8* (*GW8*) gene in high-quality Pakistani Basmati rice, which changes the grain morphology to be more elongated, significantly improves rice quality but causes a 14% loss in rice yield (Wang et al., 2012). *Grain Width7* (*GW7*) and *GL7* are located at the same genetic locus. Further studies revealed that *GW8* binds directly to the promoter of the *GW7* gene and inhibits its expression. The *GW7* gene undergoes allelic variation and becomes a semi-dominant allele $GW7^{TEA}$, which can make its gene expression level not regulated by *GW8*. Therefore, controlling the *GW8-GW7* module can significantly improve rice quality by ensuring that the yield does not decrease. The results also showed that applying optimal allelic variants of the *GW7* and *GS3* genes to indica rice, which produces a high yield in China, could significantly improve rice quality and yield (Wang SK. et al., 2015). The above study identified new genes with important implications for molecular module design and breeding of high-yielding and high-quality rice. It also provides a new clue to unraveling the molecular nature of synergistic genetic modification of rice quality and yield. In a news report in *Nature* on the results of this study, Professor Susan McCouch of Cornell University was quoted as saying that no one had taken into account both high yield and high quality in rice breeding, "which will be **[SIC]** a huge impact on research results. " Academician Qian Qian's team at the China Rice Research Institute also recently published a paper to discuss precise and efficient strategies for improving rice grain size and rice quality, with the aim of obtaining a germplasm resource bank of excellent seed diameter and grain quality genes and achieving both high yield and high quality of rice (Ren et al., 2023). Glumelles and lemmas are unique organs of graminoids that also determine grain size and quality. Recent studies have shown that *Abnormal Hull2* (*AH2*) encodes an MYB domain protein and functions during glume and grain development. The grain of *AH2* mutants is smaller, and its quality is changed due to decreased amylose content, gel consistency, and increased protein content. At the same time, a part of the glumes loses its outer silicified cells, resulting in a conversion from the outer rough epidermis to the inner smooth epidermal cells (Ren et al., 2019a). The analysis of the *AH2* function helps to understand the potential mechanisms regulating rice grain shape and is important for improving rice yield and quality.

Research advances in molecular design breeding

Striking a balance between yield and quality in rice - much like the proverbial challenge of having both fish and bear's paws - is an intricate task but critical for future global food security. The field of rice breeding is constantly evolving, fueled by rapid advances in molecular biology, genomics, and bioinformatics. The techniques and methods used in rice breeding have undergone continuous refinement and innovation, ranging from traditional crossbreeding and mutation breeding to more modern approaches such as molecular marker-assisted breeding molecular design breeding, and gene editing breeding. These technologies can combine different rice yield traits to produce high-yielding, high-quality varieties. At the same time, they can also combine various excellent characteristics, such as disease resistance, drought resistance, lodging resistance, and high-quality grain, leading to the cultivation of superior rice varieties and promoting sustainable agriculture. Many functional genes have been identified, and scientists have employed some of them to develop corresponding molecular markers (Table 4; Figure 2). This has resulted in the cultivation of numerous superior rice varieties, significantly boosting the efficiency of rice breeding and enabling the breeding of rice varieties that incorporate numerous advantages and benefits.

Advances in key technologies for molecular design breeding

CRISPR/Cas9 technology for gene editing and breeding

Rice genome editing technology refers to the precise modification of genes in rice to improve its characteristics (Voytas, 2013), such as increasing plant resistance to disease, stress, and yield, making it possible to cultivate excellent rice varieties. CRISPR/Cas9 genome editing technology has been widely used for rice fertility and to improve resistance, quality, and agronomic traits. It has developed into a mainstream gene editing system, which is important for future rice breeding. In recent years, many scientists have used CRISPR/Cas9 technology to perform genome editing and have obtained many high-quality rice germplasm resources. Song et al. used this technology to edit the ideal plant type gene, *IPA1*, and obtained a material, IPA1-Pro10, which can increase both the number of tillers and the number of grains per panicle, thus greatly promoting rice yield (Song et al., 2022). In terms of rice quality improvement, Huang et al. used this technology to edit the promoter of the *Waxy* (*Wx*) gene, which controls the amylose content (AC) in rice grains, and create novel *Wx* alleles with fine AC levels, improving eating and cooking quality (ECQ) (Huang et al., 2020). This technology balances beneficial complementary traits and greatly facilitates the process of cultivating high-quality germplasm.

Apomixis

The utilization of heterosis has markedly improved rice yield, but it is hampered by trait separation in hybrid offspring, and heterosis cannot be maintained. Based on this, the possibility has been advanced to transform the process of sexual reproduction of rice to apomictic reproduction so that hybrid offspring would have the same genotype as the parents and heterosis would be preserved. Apomixis

10.3389/fpls.2023.1206165

TABLE 4 Functional gene molecular markers developed in China.

Donor parent	Recipient parent	Gene	Reference
Yuefeng B	320B without fragrance	*fgr*	(Li JH. et al., 2006)
Chuanxiang 29B	lemont, R2, Mei B, P18, G46B	*fgr*	(Sun, 2007)
Daohuaxiang 2	Kongyu 131	*fgr*	(Li et al., 2020)
Xiangjing 9407	IRBB60	*Xa21, fgr*	(Sang et al., 2009)
BL122, XH74	R290	*Pi-1, fgr*	(Wang et al., 2010)
CBB23	Chuanhui 907, Chuanhanghui 908, Chuanhui 991, Chuanhui 992	*xa23*	(Wang et al., 2021)
Mowanggu, Akihikari	IR64, Teqing	*Wx*	(Mao et al., 2017)
K81	Chuanxiang 29B, IR58025B, BoIIIB	*Wx*	(Li et al., 2009)
HC086	R838, R898, R476, R6547	*Wx*	(Hou et al., 2009)
Nanjing 9108	Shen 01B	*Wx-mq*	(Niu et al., 2019)
Kantou 194	Wuyujing 7	$Stv\text{-}b^{i}$	(Yao et al., 2013)
Kantou 194	Wuyujing 7	*Wx-mq*, $Stv\text{-}b^{i}$	(Yao et al., 2010)
Jiangtangdao 1	Miyang 23	*sbe3-rs*	(Yang et al., 2012)
Jiangtangdao 1	Hudao 89	*sbe3-rs*	(Bai et al., 2023)
Sanlicun, IAC1246, JAPPENI TUNGKUNGO, MIGA	SAGC-4	*GS3*	(Li et al., 2016)
htd1-2, Taibei 309	htd1-2, 9311	*HTD1-2*	(Jiang et al., 2009)
TN1, TN8, TN19, TN21, TN28, rice-1, rice-2, E6, ipa1-2010	Changhui 891, Changhui 121, Changhui T025, Changhui T470 and so on	*IPA1*	(Ye, 2013)
Zhendao 88	Wulingjing 1	$TAC1^{TQ}$, $qsB\text{-}9^{TQ}$	(Chen et al., 2014)
Fukei 138	LTH, Nipponbare, Kenjiandao 3, Kenjiandao 6, Kendao 8, Kendao 12, Kendao 16, Suijing 3, Longjing 10, Longhua00-835, Longjing 13, Longjing 34	*Pi35*	(Ma et al., 2015)
Kongyu 131	Zhennuo 19	*Pigm*	(Li et al., 2022)
RBR1-2	Guangzhan 63S	*Pi-1, Pi-2*	(Zhou GL. et al., 2017)
Katy	RU9101001	*Pi-ta*	(Wang et al., 2005)
Digu, BL-1, Pi-4	G46B, Jiangnanxiangnuo	$Pi\text{-}d(t)^{1}$, *Pi-b*, $Pi\text{-}ta^{2}$	(Chen et al., 2004)
Wuyunjing 8	Yandao 10	*Pi-ta, Pi-b, Pi-9*	(Liu et al., 2021)
Nanjing 5055	Wujing 15	*Pi-ta, Pi-b, Wx-mq*	(Yao et al., 2017)

(Continued)

TABLE 4 Continued

Donor parent	Recipient parent	Gene	Reference
Wuyunjing 8	Zhendao 42	*Pi-ta, Pi-b, Stv-b*^{i}	(Wang J. et al., 2011)
H401	Lianjing 05-45	*Pi-ta, Pi-2*	(Xing et al., 2023)
Nipponbare, 9311, Kantou 194, Ning 9108, Nanjing 11, Bijing 43.Bijing 44, Bijing 45	02428, Peiai 64S	S_5^n	(Tian et al., 2014)
O. rufipogon Griff.	Zhenshan 97	*TIG1*	(Li et al., 2023)
Kasalath, Shuhui 498	Nipponbare, Suxiu 867	*Hd6*	(Peng et al., 2022)
B5	Huhan 1B, Xiushui 123, Huhan 7B, Huanghuazhan, Zhongzu 14, Huhan 15, Nipponbare	*Bph14*	(Liu et al., 2018)
Rathu Heenati	9311	*Bph3*	(Deng et al., 2018)
B5	9311, 1826	*Bph14, Bph15*	(Li JB. et al., 2006)
B5	C815S	*Bph14, Bph15*	(Ren XM. et al., 2018)

is a plant that produces a clone of itself through asexual reproduction by seed, and this strategy could be revolutionary for agriculture (Liu et al., 2023). BABY BOOM (BBM), a member of the AP2 family of transcription factors specifically expressed in spermatids, continues to be expressed in zygotes after fertilization. Thus, this transcription factor may be necessary during the initiation of embryonic development (Horstman et al., 2014). Sundaresan's research group demonstrated that *BBM1* may play a crucial role in zygotic embryos and that ectopic expression of *BBM1* allows female cells to bypass fertilization and form normal embryos through parthenogenesis. By editing three *MiMe* genes and ectopically expressing the *BBM1* gene, the replacement of meiosis by mitosis was achieved, establishing an apomictic reproductive system in rice and allowing the asexual reproduction of rice seeds (Khanday et al., 2018). During the same period, the study performed by Kejian Wang's team at the China National Rice Research Institute showed that the heterozygosity of F_1 hybrid rice could be fixed. These researchers employed the Chunyou 84 hybrid rice cultivar as a parent to edit three genes using CRISPR/Cas9 technology, *PAIR1*, *OsREC8*, and *OsOSD1*, which are involved in the meiotic phase of rice, and the *MATRILINEAL* (*MTL*) gene, which is homologous to the maize haploid induction gene. The resulting tetraploid mutant rice replaced meiosis with a division process similar to mitosis, generating progeny with a genotype identical to that of the parents and achieving the fixation of heterozygous genotypes (Wang et al., 2019). Although both methods realize apomictic reproduction in rice, they affect rice

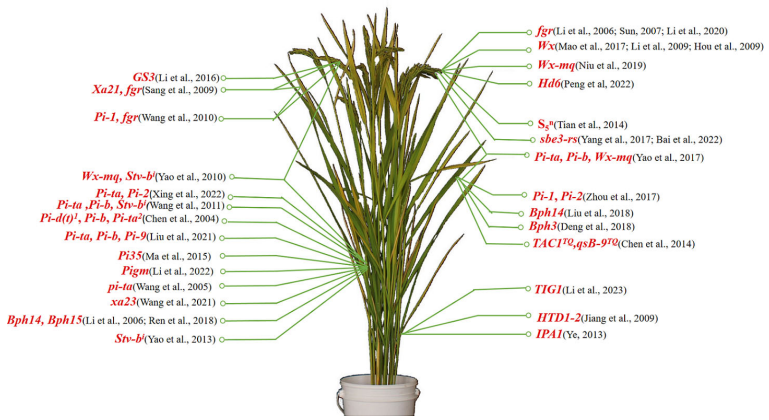


FIGURE 2
Distribution map of functional gene molecular markers developed in China.

fertility to a certain degree. Therefore, future research should focus on the improvement of apomictic reproduction to realize its application in rice farming.

Molecular design breeding for rice variety improvement

Molecular design breeding is a technology that uses gene editing, synthetic biology, and other modalities to directly manipulate DNA sequences. This approach expedites the breeding process and enhances its efficiency (Wang JK. et al., 2011). The rational aggregation of high-quality genes through molecular design methods is conducive to the breeding of superior new rice varieties (Figure 3). This technology has allowed researchers to consolidate numerous genes for breeding, marking a significant milestone in the history of rice breeding. Utilizing the established gene regulatory network for yield traits to conduct molecular design breeding holds great promise for enhancing crop yield and quality. High-quality genes from Japonica Nipponbare and Indica 9311 were used as desirable target gene donors to design 28 genes affecting rice yield, appearance quality, culinary quality, and ecological adaptability. With its inferior culinary quality, the ultra-high-yield cultivar "Teqing" served as a recipient of these desirable genes. After extensive periods of hybridization, backcrossing, and polymer selection, combined with directional selection using molecular markers, several exceptional offspring materials were produced. These offspring materials fully retained the genetic background and high-yielding characteristics of Teqing, while significantly improving the visual quality, culinary quality, taste, and flavor of the rice. The quality of the corresponding hybrid rice was also markedly enhanced (Zeng et al., 2017a). These findings embody an excellent transition from traditional crop breeding to efficient, precise, and directed molecular design breeding, and they are likely to inspire similar future efforts.

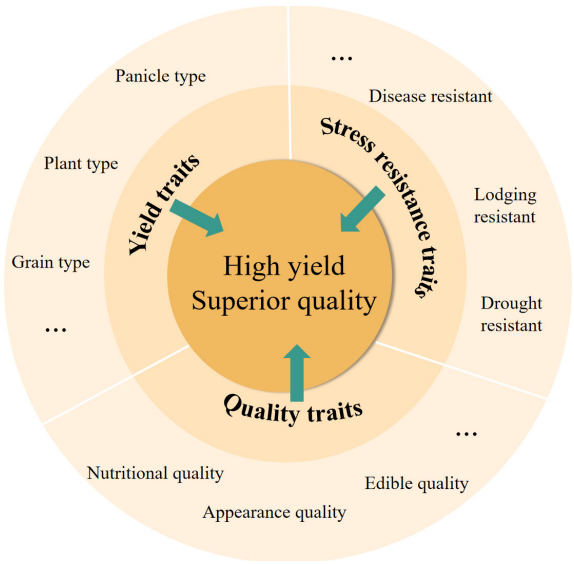


FIGURE 3
Reasonable molecular design breeding strategy.

A significant research effort led to the construction of a pan-genomic dataset of the *O.sativa-O.rufipogon* complex, which was generated by deep sequencing and reassembly of 66 different cultivars. The genome comparison identified 23 million sequence variants in the rice genome. These sequence variants, including many known quantitative traits, will help identify novel causal variants that constitute complex traits. Most importantly, this pan-genomic dataset was used to analyze the entire set of coding genes, revealing many variants in different rice cultivars (Zhao et al., 2018). This pan-genomic resource will promote evolutionary and functional studies of plant architecture and grain development in rice, and provide a reference for rice design and breeding. In addition, the polygenic (*GHD10-GS2-DEP1-IPA1*) pyramiding effect was studied, and a pyramiding breeding model for high-yielding genes was designed using the cloned series of yield-affecting genes (Zeng et al., 2017b). The model elaborates the molecular breeding theory based on the utilization of heterosis between the indica and japonica subspecies and provides a theoretical basis for the future design of super hybrid rice, offering a feasible strategy to achieve the third leap in rice production yield.

The "Zhongke 804" and "Zhongkefa" series are landmark new rice cultivars developed by Jiayang Li, a member of the Chinese Academy of Sciences and a researcher at the Institute of Genetics and Developmental Biology of the Chinese Academy of Sciences. These cultivars were developed based on the theoretical foundation and design concept of "Molecular Mechanism and Cultivar Design of Rice High-yield and High-quality Traits," and represent important achievements made with the support of the Class A Strategic Pilot Science and Technology Special Project "Innovation Systems of Molecular Module Design Breeding" of the Chinese Academy of Sciences. Compared with the main high-quality rice cultivar, "Daohuaxiang, " farmed in Northeast China, the new japonica rice cultivars "Zhongke 804" and "Zhongkefa" series have the characteristics of high quality, high yield, blast resistance, lodging resistance, and better taste (Zeng et al., 2017a). Given the characteristics and production needs of the northeastern region of rice farming, the research group of Jiayang Li, in cooperation with the teams of Qian at the China National Rice Research Institute and Guomin Zhang at the Northern Japonica Rice Molecular Breeding Joint Research Center, have developed new varieties, "Zhongkefa 5" and "Zhongkefa 6", which are characterized by high yield, high quality, and strong stress resistance (GUO et al., 2019). In addition, Jiayang Li's research group also developed a new conventional early maturing soft japonica rice variety, "Zhongkefa 928," based on excellent quality-related genes and blast resistance genes through in-depth analysis and molecular design, using new molecular design breeding techniques (Yu et al., 2018). As a typical example of the transition from scientific theory to production practice, these varieties will provide a higher level of sophistication to the field of variety design and breeding research and greatly promote the transition from traditional crop breeding to efficient breeding based on a precise and targeted molecular design.

Due to the pleiotropy between genes, potential genetic instability, and complex regulatory networks among genes, few gene resources are available for improving yield traits in rice varieties. When using molecular design techniques to aggregate

excellent genes, it is also difficult to balance the expression of each gene. Therefore, further exploring the genes of excellent yield traits and revealing the regulatory networks and genetic mechanisms between each gene is significant for cultivating high-yielding and high-quality varieties and establishing excellent germplasm.

Prospects

In recent years, the challenge of ensuring national food security through the cultivation of high-yielding, high-quality rice varieties has been exacerbated by population growth and shrinking agricultural land. Chinese scientists have accomplished numerous significant breakthroughs in rice yield trait research. They have applied advanced breeding techniques like molecular markers and molecular design, cloned an array of rice yield trait genes, and amassed a wealth of excellent genes to breed several high-quality rice varieties. For instance, they have used CRISPR/Cas9 and apomixis technology for germplasm improvement and used molecular design breeding methods to cultivate superior rice varieties, such as the Zhongkefa series.

Based on the existing functional genes and regulatory networks, further mining of related genes and analysis of gene regulatory networks, developing and using molecular markers with application value, the establishment of a high-throughput and low-cost genotype detection system through multiple gene pyramiding, gene editing, and apomixis, the creation of excellent new breeding materials, and cultivating new rice varieties with high yield and quality will be an inevitable trend in future rice research.

Author contributions

YR, YM and YW conceived and designed the article framework. QZ wrote the manuscript. QZ, YR, QJ, WY and YW revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the Zhejiang Provincial Natural Science Foundation of China (Grant No. LZ23C130003), the National Key R&D Program of China (Grant No: 2021YFA1300703), Hainan Yazhou Bay Seed Lab (Grant No. B21HJ0219), National Natural Science Foundation of China (Grant No. 31971921).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Bai, X. F., Huang, Y., Hu, Y., Liu, H. Y., Zhang, B., Smaczniak, C., et al. (2017). Duplication of an upstream silencer of *FZP* increases grain yield in rice. *Nat. Plants.* 3, 885–893. doi: 10.1038/s41477-017-0042-4

Bai, J. J., Tang, J. H., Pu, Z. Z., Wan, C. Z., Gong, C. C., and Yang, R. F. (2023). Development and utilization of KASP molecular marker for high resistant starch gene, *sbe3-rs*, in rice (*Oryza sativa* l.). *Mol. Plant Breed.*

Che, R. H., Tong, H. N., Shi, B. H., Liu, Y. Q., Fang, S. R., Liu, D. P., et al. (2015). Control of grain size and rice yield by *GL2*-mediated brassinosteroid responses. *Nat. Plants.* 2, 15195. doi: 10.1038/nplants.2015.195

Chen, X. W., Li, S. G., Ma, Y. Q., Li, H. Y., Zhou, K. D., and Zhu, L. H. (2004). Marker-assisted selection and pyramiding for three blast resistance genes, *Pi-d(t)*^{1}, *Pi-b*, *Pi-ta2*, in rice. *Chin. J. Biotechnol.* 5, 708–714. doi: 10.13345/j.cjb.2004.05.015

Chen, F., Qian, Q., Wang, T., Dong, A. W., Qi, X. Q., Zuo, J. R., et al. (2018). Research advances in plant science in China in 2017. *Bull. Bot.* 53, 391–440. doi: 10.11983/CBB18177

Chen, S. H., Zhou, L. J., Xu, P., and Xue, H. W. (2018). SPOC domain-containing protein leaf inclination3 interacts with LIP1 to regulate rice leaf inclination through auxin signaling. *PloS Genet.* 14, e1007829. doi: 10.1371/journal.pgen.1007829

Chen, Z. X., Zuo, S. M., Zhang, Y. F., Chen, H. Q., Feng, M. H., Jiang, W., et al. (2014). Breeding potential and interaction effects of *qSB-9*^{TQ} and *TAC1*^{TQ} in rice breeding against sheath blight disease. *Chin. J. Rice Sci.* 28, 479–486. doi: 10.3969/j.issn.1001-7216.2014.05.005

Deng, Z., Shi, Y. Y., Zhao, X. H., Qin, P., Liu, K. Y., Wang, K., et al. (2018). Development and application of gene-specific molecular markers for rice brown planthopper gene *Bph3*. *Mol. Plant Breed.* 16, 3563–3568. doi: 10.13271/j.mpb.016.003563

Duan, P. G., Ni, S., Wang, J. M., Zhang, B. L., Xu, R., Wang, Y. X., et al. (2015). Regulation of *OsGRF4* by OsmiR396 controls grain size and yield in rice. *Nat. Plants.* 2, 15203. doi: 10.1038/nplants.2015.203

Gan, L., Wu, H., Wu, D. P., Zhang, Z. F., Guo, Z. F., Yang, N., et al. (2015). Methyl jasmonate inhibits lamina joint inclination by repressing brassinosteroid biosynthesis and signaling in rice. *Plant Sci.* 241, 238–245. doi: 10.1016/j.plantsci.2015.10.012

Guo, T., Chen, K., Dong, N. Q., Shi, C. L., Ye, W. W., Gao, J. P., et al. (2018). *GRAIN SIZE AND NUMBER1* negatively regulates the *OsMKKK10-OsMKK4-OsMPK6* cascade to coordinate the trade-off between grain number per panicle and grain size in rice. *Plant Cell.* 30, 871–888. doi: 10.1105/tpc.17.00959

GUO, T., Yu, H., Qiu, J., Li, J. Y., Han, B., and Lin, H. X. (2019). Advances in rice genetics and breeding by molecular design in China. *Sci. Sin. Vitae.* 49, 185–1212. doi: 10.1360/SSV-2019-0209

Hao, J. Q., Wang, D. K., Wu, Y. B., Huang, K., Duan, P. G., Li, N., et al. (2021). The *GW2-WG1*-OsbZIP47 pathway controls grain size and weight in rice. *Mol. Plant* 14, 1266–1280. doi: 10.1016/j.molp.2021.04.011

Horstman, A., Willemsen, V., Boutilier, K., and Heidstra, R. (2014). AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. *Trends Plant Sci.* 19, 146–157. doi: 10.1016/j.tplants.2013.10.010

Hou, L. H., Xia, M. Y., Qi, H. X., Yuan, Q. H., and Yin, D. S. (2009). Developing restorer lines with intermediate amylose content by molecular marker-assisted selection in rice. *Chin. Agric. Sci. Bull.* 25, 32–36. doi: 10.11924/j.issn.1000-6850.2009-0266

Hu, J., Wang, Y. X., Fang, Y. X., Zeng, L. J., Xu, J., Yu, H. P., et al. (2015). A rare allele of *GS2* enhances grain size and grain yield in rice. *Mol. Plant* 8, 1455–1465. doi: 10.1016/j.molp.2015.07.002

Huang, L. C., Li, Q. F., Zhang, C. Q., Chu, R., Gu, Z. W., Tan, H. Y., et al. (2020). Creating novel *Wx* alleles with fine-tuned amylose levels and improved grain quality in rice by promoter editing using CRISPR/Cas9 system. *Plant Biotechnol. J.* 18, 2164–2166. doi: 10.1111/pbi.13391

Huang, X. Z., Qian, Q., Liu, Z. B., Sun, H. Y., He, S. Y., Luo, D., et al. (2009). Natural variation at the *DEP1* locus enhances grain yield in rice. *Nat. Genet.* 41, 494–497. doi: 10.1038/ng.352

Huang, L. Z., Wang, W. G., Zhang, N., Cai, Y. Y., Liang, Y., Meng, X. B., et al. (2021). *LAZY2* controls rice tiller angle through regulating starch biosynthesis in gravity-sensing cells. *New Phytol.* 231, 1073–1087. doi: 10.1111/nph.17426

Huang, Y. Y., Zhao, S. S., Fu, Y. C., Sun, H. D., Ma, X., Tan, L. B., et al. (2018). Variation in *FZP* regulatory region causes increases of inflorescence secondary branches and grain yield in rice domestication. *Plant J.* 96, 716–733. doi: 10.1111/tpj.14062

Huo, X., Wu, S., Zhu, Z. F., Liu, F. X., Fu, Y. C., Cai, H. W., et al. (2017). *NOG1* increases grain production in rice. *Nat. Commun.* 8, 1497. doi: 10.1038/s41467-017-01501-8

Jiang, L., Liu, X., Xiong, G. S., Liu, H. H., Chen, F. L., Wang, L., et al. (2013). *DWARF 53* acts as a repressor of strigolactone signalling in rice. *Nature* 504, 401–405. doi: 10.1038/nature12870

Jiang, J. H., Tan, L. B., Zhu, Z. F., Fu, Y. C., Liu, F. X., Cai, H. W., et al. (2012). Molecular evolution of the *TAC1* gene from rice (*Oryza sativa* l.). *J. Genet. Genomics* 39, 551–560. doi: 10.1016/j.jgg.2012.07.011

Jiang, H. P., Zhang, S. Y., Bao, J. S., Wang, B. L., and Wang, S. (2009). Gene analysis and mapping of high-tillering and dwarf mutant *htd1-2* in rice. *Hereditas (Beijing).* 31, 531–539. doi: 10.3724/SP.J.1005.2009.00531

Jiao, Y. Q., Wang, Y. H., Xue, D. W., Wang, J., Yan, M. X., Liu, G. F., et al. (2010). Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice. *Nat. Genet.* 42, 541–544. doi: 10.1038/ng.591

Khanday, I., Skinner, D., Yang, B., Mercier, R., and Sundaresan, V. (2018). A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. *Nature* 565, 91–95. doi: 10.1038/s41586-018-0785-8

Li, B., Cai, Z. J., Wang, L., Chen, J., Cao, K. R., Li, J., et al. (2022). Development and application of the combinatorial marker for the rice blast resistance gene *Pigm*. *Biotechnol. Bull.* 38, 153–159. doi: 10.13560/j.cnki.biotech.bull.1985.2021-1228

Li, Q. F., Lu, J., Zhou, Y., Wu, F., Tong, H. N., Wang, J. D., et al. (2019). Abscisic acid represses rice lamina joint inclination by antagonizing brassinosteroid biosynthesis and signaling. *Int. J. Mol. Sci.* 20, 4908. doi: 10.3390/ijms20194908

Li, R. T., Meng, D. X., Su, D., Wang, M. L., and Liu, C. H. (2020). Breed a introgression line kongyu 131 (*fgr*) of early japonica based on marker-assisted selection (MAS)-mediated backcross transformation method. *Seed* 39, 5–12. doi: 10.16590/j.cnki.1001-4705.2020.05.008

Li, Z. Z., Peng, Q. X., Qiu, X. J., Xu, J. Y., Li, Z. X., and Liu, H. Y. (2023). Development and application of functional molecular marker of rice tiller angle gene *TIG1*. *J. Plant Genet. Resour.* 24, 804–816. doi: 10.13430/j.cnki.jpgr.20221108001

Li, J. H., Wang, F., Liu, W. G., Jin, S. J., and Liu, Y. B. (2006). Genetic analysis and mapping by SSR marker for fragrance gene in rice yuefeng b. *Mol. Plant Breed.* 1, 54–58. doi: 10.3864/j.issn.0578-1752.at-2005-5886

Li, P. J., Wang, Y. H., Qian, Q., Fu, Z. M., Wang, M., Zeng, D. L., et al. (2007). *LAZY1* controls rice shoot gravitropism through regulating polar auxin transport. *Cell Res.* 17, 402–410. doi: 10.1038/cr.2007.38

Li, J. B., Xia, M. Y., Qi, H. X., He, G. C., Wan, B. L., and Zha, Z. P. (2006). Marker-assisted selection for brown planthopper (*Nilaparvata iugens stal*) resistance genes *Bph14* and *Bph15* in rice. *Scientia Agricultura Sinica.* 39, 2132–2137.

Li, N., Xu, R., and Li, Y. H. (2019). Molecular networks of seed size control in plants. *Annu. Rev. Plant Biol.* 70, 435–463. doi: 10.1146/annurev-arplant-050718-095851

Li, Y., Xu, X. Y., Yan, M., Feng, F. J., Ma, X. S., and Mei, H. W. (2016). Improvement of rice grain shape by functional molecular marker of *GS3* gene. *Acta Agriculturae Shanghai.* 32, 1–5. doi: 10.15955/j.issn1000-3924.2016.01.01

Li, Z. H., Zeng, L. H., Gao, F. Y., Lu, X. J., Liu, G. C., and Ren, G. J. (2009). Development and utilization of SSR primer in *Wx* gene of rice (*Oryza sativa* l.). *Southwest China J. Agric. Sci.* 22, 231–235. doi: 10.16213/j.cnki.scjas.2009.02.021

Liu, C. L., He, Z. X., Zhang, Y., Hu, F. Y., Li, M. Q., Liu, Q., et al. (2023). Synthetic apomixis enables stable transgenerational transmission of heterotic phenotypes in hybrid rice. *Plant Commun.* 4, 100470. doi: 10.1016/j.xplc.2022.100470

Liu, K., Wan, B. J., Zhao, S. L., Zhu, J. W., Zhang, G. Y., Dai, J. Y., et al. (2021). Pyramiding resistance gene *Pi-ta*, *Pi-b* and *Pi-9* by marker-assisted selected selection in rice (*Oryza sativa* l.). *Southwest China J. Agric. Sci.* 34, 926–931. doi: 10.16213/j.cnki.scjas.2021.5.003

Liu, Y., Zhang, A. N., Wang, F. M., Tang, J. J., Luo, L. J., Yu, X. J., et al. (2018). Molecular marker development and breeding application of brown planthopper resistance gene *Bph14* of rice. *Mol. Plant Breed.* 16, 4658–4662. doi: 10.13271/j.mpb.016.004658

Ma, J., Ma, X. D., Zhao, Z. C., Wang, S., Wang, J. L., Wang, J., et al. (2015). Development and application of a functional marker of the blast resistance gene *Pi35* in rice. *Acta Agron. Sin.* 41, 1779–1790. doi: 10.3724/SP.J.1006.2015.01779

Mao, T., Li, X., Li, Z. Y., and Xu, Z. J. (2017). Development of PCR functional markers for multiple alleles of *Wx* and their application in rice. *Acta Agron. Sin.* 43, 1715–1723. doi: 10.3724/SP.J.1006.2017.01715

Niu, F. A., Zhou, J. H., Cao, L. M., Cheng, C., Tu, R. J., Hu, X. J., et al. (2019). Development and utilization of KASP marker for low amylose content gene, Wx^{mq}, in rice (*Oryza sativa* l.). *Mol. Plant Breed.* 17, 8125–8131. doi: 10.13271/j.mpb.017.008125

Ordonio, R. L., and Matsuoka, M. (2017). New path towards a better rice architecture. *Cell Res.* 27, 1189–1190. doi: 10.1038/cr.2017.115

Peng, Y. B., Du, C. Y., Zheng, C. K., Zhou, J. J., Sun, W., He, Y. N., et al. (2022). Development and application of PARMS markers specific for rice heading date regulation gene *Hd6*. *Shandong Agric. Sci.* 54, 1–6. doi: 10.14083/j.issn.1001-4942.2022.08.001

Qiao, S. L., Sun, S. Y., Wang, L. L., Wu, Z. H., Li, C. X., Li, X. M., et al. (2017). The RLA1/SMOS1 transcription factor functions with *OsBZR1* to regulate brassinosteroid signaling and rice architecture. *Plant Cell.* 29, 292–309. doi: 10.1105/tpc.16.00611

Qu, L., Lin, L. B., and Xue, H. W. (2019). Rice miR394 suppresses leaf inclination through targeting an f-box gene, *LEAF INCLINATION 4*. *J. Integr. Plant Biol.* 61, 406–416. doi: 10.1111/jipb.12713

Ren, D. Y., Cui, Y. J., Hu, H. T., Xu, Q. K., Rao, Y. C., Yu, X. Q., et al. (2019a). *AH2* encodes a MYB domain protein that determines hull fate and affects grain yield and quality in rice. *Plant J.* 100, 813–824. doi: 10.1111/tpj.14481

Ren, D. Y., Ding, C. Q., and Qian, Q. (2023). Molecular bases of rice grain size and quality for optimized productivity. *Sci. Bull. (Beijing).* 68, 314–350. doi: 10.1016/j.scib.2023.01.026

Ren, D. Y., Hu, J., Xu, Q. K., Cui, Y. J., Zhang, Y., Zhou, T. T., et al. (2018). *FZP* determines grain size and sterile lemma fate in rice. *J. Exp. Bot.* 69, 4853–4866. doi: 10.1093/jxb/ery264

Ren, X. M., Xiang, C., Lei, D. Q., and Lei, D. Y. (2018). Improvement of BPH resistance of rice TGMS line C815S through molecular marker-assisted selection. *Hybrid Rice.* 33, 54–58; 87. doi: 10.16267/j.cnki.1005-3956.20171128.233

Ren, D. Y., Xu, Q. K., Qiu, Z. N., Cui, Y. J., Zhou, T. T., Zeng, D. L., et al. (2019b). *FON4* prevents the multi-floret spikelet in rice. *Plant Biotechnol. J.* 17, 1007–1009. doi: 10.1111/pbi.13083

Ren, D. Y., Yu, H. P., Rao, Y. C., Xu, Q. K., Zhou, T. T., Hu, J., et al. (2017). 'Two-floret spikelet' as a novel resource has the potential to increase rice yield. *Plant Biotechnol. J.* 16, 351–353. doi: 10.1111/pbi.12849

Sang, D. J., Chen, D. Q., Liu, G. F., Liang, Y., Huang, L. Z., Meng, X. B., et al. (2014). Strigolactones regulate rice tiller angle by attenuating shoot gravitropism through inhibiting auxin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 111, 11199–11204. doi: 10.1073/pnas.1411859111

Sang, M. P., Jiang, M. S., Li, G. X., and Yao, F. Y. (2009). Pyramiding *Xa21* and *fgr* in rice by marker-assisted selection. *Shandong Agric. Sci.* 1, 4–7. doi: 10.14083/j.issn.1001-4942.2009.01.019

Song, X., Huang, W., Shi, M., Zhu, M. Z., and Lin, H. X. (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39, 623–630. doi: 10.1038/ng2014

Song, X. G., Lu, Z. F., Yu, H., Shao, G. N., Xiong, J. S., Meng, X. B., et al. (2017). *IPA1* functions as a downstream transcription factor repressed by *D53* in strigolactone signaling in rice. *Cell Res.* 27, 1128–1141. doi: 10.1038/cr.2017.102

Song, X. G., Meng, X. B., Guo, H. Y., Cheng, Q., Jing, Y. H., Chen, M. J., et al. (2022). Targeting a gene regulatory element enhances rice grain yield by decoupling panicle number and size. *Nat. Biotechnol.* 40, 1403–1411. doi: 10.1038/s41587-022-01281-7

Sun, S. X. (2007). "Gene mapping and expression of aroma in rice Chuanxiang29 and molecular assisted breeding," (Yaan: Sichuan Agricultural University).

Sun, S. Y., Wang, L., Mao, H. L., Shao, L., Li, X. H., Xiao, J. H., et al. (2018). A G-protein pathway determines grain size in rice. *Nat. Commun.* 9, 851. doi: 10.1038/s41467-018-03141-y

Tian, M. X., Yu, B. X., He, Y. X., Zhang, S. L., Wang, S., and Ye, Y. Y. (2014). Design and validation for a new function marker of wide compatibility gene S_5^n in rice. *Guangdong Agric. Sci.* 41, 134–137. doi: 10.16768/j.issn.1004-874x.2014.16.025

Voytas, D. F. (2013). Plant genome engineering with sequence-specific nucleases. *Annu. Rev. Plant Biol.* 64, 327–350. doi: 10.1146/annurev-arplant-042811-105552

Wang, M. X., Bai, Y. L., Zhang, Q., Xu, D. W., Zhang, Z. Y., and Wang, P. (2021). Development and utilization of bacterial blight resistance gene *Xa23* molecular marker. *Southwest China J. Agric. Sci.* 34, 2070–2075. doi: 10.16213/j.cnki.scjas.2021.10.002

Wang, S. K., Li, S., Liu, Q., Wu, K., Zhang, J. Q., Wang, S. S., et al. (2015). The *OsSPL16-GW7* regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nat. Genet.* 47, 949–954. doi: 10.1038/ng.3352

Wang, J. K., Li, H. H., Zhang, X. C., Yin, C. B., Li, Y., Ma, Y. Z., et al. (2011). Molecular design breeding in crops in China. *Crop J.* 37, 191–201. doi: 10.3724/SP.J.1006.2011.00191

Wang, F., Liu, W. G., Liu, Z. R., Zhu, X. Y., Li, J. H., Liao, Y. L., et al. (2010). Pyramiding *Pi-1* and *fgr* genes to improve rice restorer lines by molecular marker-assisted selection. *Hybrid Rice.* 25, 237–244. doi: 10.16267/j.cnki.1005-3956.2010.s1.041

Wang, C., Liu, Q., Shen, Y., Hua, Y. F., Wang, J. J., Lin, J. R., et al. (2019). Clonal seeds from hybrid rice by simultaneous genome engineering of meiosis and fertilization genes. *Nat. Biotechnol.* 37, 283–286. doi: 10.1038/s41587-018-0003-0

Wang, Z. H., MARC, R., and Jia, Y. L. (2005). Development of the codominant marker of rice blast resistance gene *Pi-ta*. *Chin. J. Rice Sci.* 6, 483–488. doi: 10.16819/j.1001-7216.2005.06.001

Wang, Y. X., Shang, L. G., Yu, H., Zeng, L. J., Hu, J., Ni, S., et al. (2020). A strigolactone biosynthesis gene contributed to the green revolution in rice. *Mol. Plant* 13, 923–932. doi: 10.1016/j.molp.2020.03.009

Wang, B. B., and Wang, H. Y. (2017). *IPA1*: a new "green revolution" gene? *Mol. Plant* 10, 779–781. doi: 10.1016/j.molp.2017.04.011

Wang, S. K., Wu, K., Yuan, Q. B., Liu, X. Y., Liu, Z. B., Lin, X. Y., et al. (2012). Control of grain size, shape and quality by *OsSPL16* in rice. *Nat. Genet.* 44, 950–954. doi: 10.1038/ng.2327

Wang, Y. X., Xiong, G. S., Hu, J., Jiang, L., Yu, H., Xu, J., et al. (2015). Copy number variation at the *GL7* locus contributes to grain size diversity in rice. *Nat. Genet.* 47, 944–948. doi: 10.1038/ng.3346

Wang, J., Yang, J., Chen, Z. D., Fan, F. J., Zhu, J. Y., Yang, J. H., et al. (2011). Pyramiding resistance gene *Pi-ta*, *Pi-b* and *Stv-b*^{i} by marker-assisted selection in rice (*Oryza sativa* l.). *Acta Agron. Sin.* 37, 975–981. doi: 10.3724/SP.J.1006.2011.00975

Wang, J., Yu, H., Xiong, G. S., Lu, Z. F., Jiao, Y. Q., Meng, X. B., et al. (2017). Tissue-specific ubiquitination by IPA1 INTERACTING PROTEIN1 modulates IPA1 protein levels to regulate plant architecture in rice. *Plant Cell.* 29, 697–707. doi: 10.1105/tpc.16.00879

Wu, Y., Wang, Y., Mi, X. F., Shan, J. X., Li, X. M., Xu, J. L., et al. (2016). The QTL GNP1 encodes GA20ox1, which increases grain number and yield by increasing cytokinin activity in rice panicle meristems. *PloS Genet.* 12, e1006386. doi: 10.1371/journal.pgen.1006386

Wu, K., Wang, S. S., Song, W. Z., Zhang, J. Q., Wang, Y., Liu, Q., et al. (2020). Enhanced sustainable green revolution yield via nitrogen-responsive chromatin modulation in rice. *Science* 367, eaaz2046. doi: 10.1126/science.aaz2046

Xing, Y. G., Liu, Y., Xu, B., Sun, Z. G., Lu, B. G., Chen, T. M., et al. (2023). Breeding of a new glutinous rice variety through pyramiding rice blast resistance genes *Pi-ta* and *Pi-2*. *Mol. Plant Breed.* 2023.

Xu, R., Duan, P. G., Yu, H. Y., Zhou, Z. K., Zhang, B. L., Wang, R. C., et al. (2018a). Control of grain size and weight by the *OsMKKK10-OsMKK4-OsMAPK6* signaling pathway in rice. *Mol. Plant* 11, 860–873. doi: 10.1016/j.molp.2018.04.004

Xu, R., Yu, H. Y., Wang, J. M., Duan, P. G., Zhang, B. L., Li, J., et al. (2018b). A mitogen-activated protein kinase phosphatase influences grain size and weight in rice. *Plant J.* 95, 937–946. doi: 10.1111/tpj.13971

Yamamuro, C., Ihara, Y., Wu, X., Noguchi, T., Fujioka, S., Takatsuto, S., et al. (2000). Loss of function of a rice brassinosteroid insensitive1 homolog prevents internode elongation and bending of the lamina joint. *Plant Cell.* 12, 1591–1606. doi: 10.1105/tpc.12.9.1591

Yang, R. F., Sun, C. L., Bai, J. J., Luo, Z. X., Shi, B., Zhang, J. M., et al. (2012). A putative gene *sbe3-rs* for resistant starch mutated from *SBE3* for starch branching enzyme in rice (*Oryza sativa* l.). *PloS One* 7, e43026. doi: 10.1371/journal.pone.0043026

Yao, S., Chen, T., Luo, M. R., Zhang, Y. D., Zhu, Z., Zhao, Q. Y., et al. (2013). Improving resistance of japonica variety Wuyunjing7 to rice stripe virus disease by molecular marker-assisted selection. *Acta Agric. Boreali Sin.* 28, 195–203. doi: 10.3969/j.issn.1000-7091.2013.04.036

Yao, S., Chen, T., Zhang, Y. D., Zhu, Z., Zhao, L., Zhao, Q. Y., et al. (2010). Pyramiding of translucent endosperm mutant gene *Wx-mq* and rice stripe disease resistance gene *Stv-b*^{i} by marker-assisted selection in rice (*Oryza sativa*). *Chin. J. Rice Sci.* 24, 341–347. doi: 10.3969/j.issn.1001-7216.2010.04.02

Yao, S., Chen, T., Zhang, Y. D., Zhu, Z., Zhao, Q. Y., Zhou, L. H., et al. (2017). Pyramiding *Pi-ta*, *Pi-b* and *Wx-mq* genes by marker-assisted selection in rice (*Oryza sativa* l.). *Acta Agron. Sin.* 43, 1622–1631. doi: 10.3724/SP.J.1006.2017.01622

Ye, L. H. (2013). "CAPS molecular marker development and assisted selection breeding for *ipal* in rice (*Oryza sativa* l.)," (Nanchang: Jiangxi Agricultural University).

Yu, B. S., Lin, Z. W., Li, H. X., Li, X. J., Li, J. Y., Wang, Y. H., et al. (2007). *TAC1*, a major quantitative trait locus controlling tiller angle in rice. *Plant J.* 52, 891–898. doi: 10.1111/j.1365-313X.2007.03284.x

Yu, H., Wang, B., Chen, M. J., Liu, G. F., and Li, J. Y. (2018). Research advance and perspective of rice breeding by molecular design. *Chin. Bull. Life Sci.* 30, 1032–1037. doi: 10.13376/j.cbls/2018124

Zeng, D. L., Tian, Z. X., Rao, Y. C., Dong, G. J., Yang, Y. L., Huang, L. C., et al. (2017a). Rational design of high-yield and superior-quality rice. *Nat. Plants.* 3, 17031. doi: 10.1038/nplants.2017.31

Zeng, D. L., Tian, Z. X., Rao, Y. C., Dong, G. J., Yang, Y. L., Huang, L. C., et al. (2017b). Breeding high-yield superior quality hybrid super rice by rational design. *Natl. Sci. Rev.* 3, 283–294. doi: 10.1038/nplants.2017.31

Zhang, T., Li, Y. F., Ma, L., Sang, X. C., Ling, Y. H., Wang, Y. T., et al. (2017). *LATERAL FLORET 1* induced the three-florets spikelet in rice. *Proc. Natl. Acad. Sci. U.S.A.* 114, 9984–9989. doi: 10.1073/pnas.1700504114

Zhang, G., Song, X. G., Guo, H. Y., Wu, Y., Chen, X. Y., and Fang, R. X. (2016). A small G protein as a novel component of the rice brassinosteroid signal transduction. *Mol. Plant* 9, 1260–1271. doi: 10.1016/j.molp.2016.06.010

Zhang, W. F., Tan, L. B., Sun, H. Y., Zhao, X. H., Liu, F. X., Cai, H. W., et al. (2019). Natural variations at *TIG1* encoding a TCP transcription factor contribute to plant architecture domestication in rice. *Mol. Plant* 12, 1075–1089. doi: 10.1016/j.molp.2019.04.005

Zhang, Q., Xie, J. Y., Zhu, X. Y., Ma, X. Q., Yang, T., Khan, N. U., et al. (2023). Natural variation in *Tiller number 1* affects its interaction with *TIF1* to regulate tillering in rice. *Plant Biotechnol. J* 21, 104–1057. doi: 10.1111/pbi.14017

Zhang, N., Yu, H., Yu, H., Cai, Y. Y., Huang, L. Z., Xu, C., et al. (2018). A core regulatory pathway controlling rice tiller angle mediated by the *LAZY1*-dependent asymmetric distribution of auxin. *Plant Cell.* 30, 1461–1475. doi: 10.1105/tpc.18.00063

Zhao, Q., Feng, Q., Lu, H. Y., Li, Y., Wang, A. H., Tian, Q. L., et al. (2018). Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. *Nat. Genet.* 50, 278–284. doi: 10.1038/s41588-018-0041-z

Zhao, S. Q., Hu, J., Guo, L. B., Qian, Q., and Xue, H. W. (2010). Rice leaf inclination2, a VIN3-like protein, regulates leaf angle through modulating cell division of the collar. *Cell Res.* 20, 935–947. doi: 10.1038/cr.2010.109

Zhao, L., Tan, L. B., Zhu, Z. F., Xiao, L. T., Xie, D. X., and Sun, C. Q. (2015). *PAY1* improves plant architecture and enhances grain yield in rice. *Plant J.* 83, 528–536. doi: 10.1111/tpj.12905

Zhou, F., Lin, Q. B., Zhu, L. H., Ren, Y. L., Zhou, K. N., Shabek, N., et al. (2013). *D14-SCF*(*D3*)-dependent degradation of *D53* regulates strigolactone signalling. *Nature* 504, 406–410. doi: 10.1038/nature12878

Zhou, G. L., Liu, Q. Y., Fan, J. M., Feng, R. T., Xia, Z. J., Wang, X. J., et al. (2017). High-efficient technology system on marker-assistant selection of *Pi-1* and *Pi-2* by SSR. *J. Anhui Agric. Sci.* 45, 115–117. doi: 10.13989/j.cnki.0517-6611.2017.18.036

Zhou, Y., Tao, Y. J., Tang, D. N., Wang, J., Zhong, J., Wang, Y., et al. (2017). Identification of QTL associated with nitrogen uptake and nitrogen use efficiency using high throughput genotyped CSSLs in rice (*Oryza sativa l.*). *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01166

Zhou, X. G., Wang, J., Peng, C. F., Zhu, X. B., Yin, J. J., Li, W. T., et al. (2016). Four receptor-like cytoplasmic kinases regulate development and immunity in rice. *Plant Cell Environ.* 39, 1381–1392. doi: 10.1111/pce.12696

Zou, J. H., Zhang, S. Y., Zhang, W. P., Li, G., Chen, Z. X., Zhai, W. X., et al. (2006). The rice *HIGH-TILLERING DWARF1* encoding an ortholog of arabidopsis MAX3 is required for negative regulation of the outgrowth of axillary buds. *Plant J.* 48, 687–698. doi: 10.1111/j.1365-313X.2006.02916.x

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 11 August 2023
DOI 10.3389/fpls.2023.1236251

OPEN ACCESS

EDITED BY
Ryuji Ishikawa,
Hirosaki University, Japan

REVIEWED BY
Longbiao Guo,
Chinese Academy of Agricultural Sciences, China
Weihua Qiao,
Chinese Academy of Agricultural Sciences, China

*CORRESPONDENCE
Zhikang Li
lizhikang@caas.cn
Yingyao Shi
Shiyy123@163.com

RECEIVED 07 June 2023
ACCEPTED 26 July 2023
PUBLISHED 11 August 2023

CITATION
Zeng W, Li H, Zhang F, Wang X, Rehman S, Huang S, Zhang C, Wu F, Li J, Lv Y, Zhang C, Li M, Li Z and Shi Y (2023) Functional characterization and allelic mining of *OsGLR* genes for potential uses in rice improvement.
Front. Plant Sci. 14:1236251.
doi: 10.3389/fpls.2023.1236251

Functional characterization and allelic mining of *OsGLR* genes for potential uses in rice improvement

Wei Zeng[1,2], Hua Li[1], Fanlin Zhang[1], Xinchen Wang[1], Shamsur Rehman[2], Shiji Huang[1], Chenyang Zhang[1], Fengcai Wu[1], Jianfeng Li[1], Yamei Lv[1], Chaopu Zhang[1], Min Li[1], Zhikang Li[1,3]* and Yingyao Shi[1]*

[1]School of Agronomy, Anhui Agricultural University, Hefei, China, [2]Peking University Institute of Advanced Agricultural Sciences, Shandong Laboratory of Advanced Agriculture Sciences in Weifang, Weifang, China, [3]Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China

Glutamate-like receptor (GLR) genes are a group of regulatory genes involved in many physiological processes of plants. With 26 members in the rice genome, the functionalities of most rice GLR genes remain unknown. To facilitate their potential uses in rice improvement, an integrated strategy involving CRISPR-Cas9 mediated knockouts, deep mining and analyses of transcriptomic responses to different abiotic stresses/hormone treatments and gene CDS haplotype (gcHap) diversity in 3,010 rice genomes was taken to understand the functionalities of the 26 rice GLR genes, which led us to two conclusions. First, the expansion of rice GLR genes into a large gene family during evolution had gone through repeated gene duplication events occurred primarily in two large GLR gene clusters on rice chromosomes 9 and 6, which was accompanied with considerable functional differentiation. Secondly, except for two extremely conserved ones (*OsGLR6.2* and *OsGLR6.3*), rich gcHap diversity exists at the remaining GLR genes which played important roles in rice population differentiation and rice improvement, evidenced by their very strong sub-specific and population differentiation, by their differentiated responses to day-length and different abiotic stresses, by the large phenotypic effects of five GLR gene knockout mutants on rice yield traits, by the significant association of major gcHaps at most GLR loci with yield traits, and by the strong genetic bottleneck effects and artificial selection on the gcHap diversity in populations *Xian* (*indica*) and *Geng* (*japonica*) during modern breeding. Our results suggest the potential values of the natural variation at most rice GLR loci for improving the productivity and tolerances to abiotic stresses. Additional efforts are needed to determine the phenotypic effects of major gcHaps at these GLR loci in order to identify 'favorable' alleles at specific GLR loci specific target traits in specific environments to facilitate their application to rice improvement in future.

KEYWORDS

glutamate-like receptor (GLR) genes, gene-CDS-haplotype (gcHap) diversity, functional alleles, knockout mutants, rice improvement

Introduction

Glutamate receptors (GLRs) are an important gene family that are known to play important roles in a variety of physiological processes of eukaryotes. In animals, GLRs are known to act as neurotransmitters that allow animals to respond external stimuli through ion channels (Lam et al., 1998). In plants, GLRs are known to play a role in a variety of physiological processes, including growth and development (Kang and Turano, 2003; Li et al., 2006; Michard et al., 2011; Vincill et al., 2013; Kong et al., 2015; Singh et al., 2016; Wudick et al., 2018), stress responses (Lu et al., 2014; Cheng et al., 2016; Cheng et al., 2018; Li et al., 2019; Philippe et al., 2019; Wang et al., 2019; Zheng et al., 2021), and pathogen defense (Mousavi et al., 2013; Shao et al., 2020; Liu et al., 2021; Xue et al., 2022; Yu et al., 2022). However, since GLR genes represent a large gene family that encodes membrane proteins embedded within the cell membrane, studying their functionalities in plants through traditional biochemical methods has been proven difficult.

In recent years, there has been growing interest to understand the molecular structure, expression patterns and functional roles of GLR genes in plants. Specially, technical advances in structural biology have enabled researchers to gain insights into the molecular structure of plant GLRs. One breakthrough study in this regard was the high-resolution structure of a glutamate receptor-like *AtGLR3.4* resolved from the model plant *Arabidopsis thaliana* using a cryo-electron microscopy (cryo-EM) in which the tetrameric assembly of *AtGLR3.4* subunits was revealed in a three-layer domain architecture that implies how GLRs transmit signals in plants (Green et al., 2021). In parallel, the expression patterns of plant GLR genes in response to different environmental cues such as abiotic stresses, pathogen infections and developmental transitions have recently been investigated. GLR genes have been shown to be essential for proper root growth and branching in *Arabidopsis thaliana* and rice, while plants with reduced GLR activity exhibit abnormal root growth (Li et al., 2006; Vincill et al., 2013; Singh et al., 2016). Plant GLR genes are also involved in plant responses to biotic and abiotic stresses. In particular, they have been shown to be vital for plant responses to salt, drought and chilling (cold) stresses (Lu et al., 2014; Cheng et al., 2016; Cheng et al., 2018; Zheng et al., 2021). In addition to their roles in abiotic stress tolerances, GLR genes also play an important role in plant defense systems against their pathogens. Cotton lines with CRISPR/Cas9-mediated *GhGLR4.8* knockout become highly susceptible to the fungal pathogen, *Fusarium oxysporum* f. sp. *vasinfectum* (*Fov*) and lost the ability to induce the calcium influx in response to the secreted proteins of *Fov* (Liu et al., 2021). *AtGLR3.3* and *AtGLR3.6* were shown to trigger leaf-to-leaf systemic signaling and induce systemic defenses against its piercing-sucking insect, *myzus persicae* (green peach aphid) or its chewing insect, *spodoptera litura* (cotton leafworm) when under insect attack (Xue et al., 2022). These findings highlight the importance of GLR genes in sensing and responding to pest and disease erosions and suggest their potential uses in developing new strategies for enhancing plant resistance to pests and diseases. While the importance of GLR genes in plant growth/development, stress responses and pathogen defense are becoming increasingly clear, how different GLR genes are involved in these complex processes remains unclear. Thus, it remains an interesting subject regarding how these GLR receptors can be used as targets for improving crop production and sustainability.

As the most important cereal crop grown worldwide, rice is also known to be the model species of cereals for functional and population genomic research. Today, more than 4,600 rice genes have been cloned and functionally characterized at the molecular level (https://funricegenes.github.io/). Meanwhile, the 3010 diverse rice accessions (3KRG) from 89 countries worldwide have been sequenced, revealing extremely rich genomic diversity in the primary gene pool of this important species (Project, 2014; Wang et al., 2018; Zhang et al., 2021). Unfortunately, the rapid progresses in rice functional and population genomic research have not yet been widely applied to development of more efficient breeding technologies. This is because, three pieces of information are essential to most plant breeders, but generally lacking for virtually all cloned rice genes. First, the current information on phenotypic effect(s) of a cloned gene is incomplete because this information was obtained from typical molecular biology experiments using near isogenic lines in specific genetic backgrounds and environments and thus is irrelevant to target environments of most breeding programs. In other words, the genotype by environment interaction and genetic background effects on important agronomic traits are largely uncharacterized for most cloned rice genes. Secondly, although there is rich natural allelic variation at most gene loci in rice populations (Zhang et al., 2021), but it remains a huge challenge regarding how to determine and mine desirable allele(s) at the cloned rice gene loci from rice germplasm accessions for improving specific target trait(s). Third, the functionalities of most rice genes, such as most GLR genes, remain unknown. Thus, given the fact that the functionalities of most rice genes remain unknown, one important challenge is to obtain the required information on these uncharacterized genes for breeding application without going through the typical time-consuming gene cloning processes.

In this study, we employed an integrated approach that involved phylogenetic analysis of rice GLR genes alongside their counterparts in closely related plants, phenotypic and functional characterization of important rice GLR genes using CRISPR technology and transcriptomic analyses, as well as comprehensive population genetic analyses of their gene-CDS-haplotype (gcHap) diversity in rice populations. Collectively, we demonstrated a more efficient strategy to generate important information for large numbers of gene loci required for developing novel breeding technologies in future.

Materials and methods

The phylogenetic relations between plant GLRs and mammals iGluRs (ionotropic glutamate receptors)

To understand the relationships between the rice GLR genes and those in animals and other plants, we compared all GLR genes in the Os-Nipponbare-Reference-IRGSP-1.0 reference genome with those in other flowering plants (Simon et al., 2023), including

Arabidopsis thaliana (*At*), *Zea mays* (*Zm*), *Solanum lycopersicum* (*Sl*), *Gossypium hirsutum* (*Gh*), and *Raphanus sativus* (*Rs*), as well as basal land plants such as the moss *Physcomitrium patens* (*Pp*), the liverwort *Marchantia polymorpha* (*Mp*), and the lycophyte *Selaginella moellendorffii* (*Sm*), compared to the invertebrate *Caenorhabditis elegans* (*Ce*) and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors (AMPARs), N-methyl-d-aspartate receptors (NMDARs), kainate receptors (KARs), and δ-receptors from mammals. Also included is the bacterial GluR0 from *Synechocystis PCC 6803* and AvGluR1 from the freshwater rotifer *Adineta vaga* (*Av*). Sequences are available in the Supplementary Data 1. Phylogenetic analysis of all plant GLRs and mammals iGluRs was conducted using the MEGA7.0 software (Shi et al., 2022). Muscle sequence comparison, which is a logarithmic expected multi-sequence comparison method, was performed to compare the protein sequences of GLR proteins of the relevant species. Then, a Bootstrap method with 1,000 replicates using the neighbor-joining (NJ) method (with p-distance correction) was used to assess the robustness of all nodes in the tree. The resulting phylogenetic tree was visualized using the website iTOL (https://itol.embl.de/), which allowed for the interactive exploration and customization of phylogenetic trees.

The gene-CDS-haplotype (gcHap) diversity analyses

The construction method of gcHap dataset in the 3,010 rice genomes (3KRG, Wang et al., 2018) was referred to the *Molecular Plant* article of Zhang (Zhang et al., 2021). A total of 913,360 single nucleotide polymorphisms (SNPs) were identified within the coding sequence (CDS) regions of 42,497 of the 45,963 annotated genes in the Os-Nipponbare-Reference-IRGSP-1.0 reference genome. In the rice genome, no SNPs were found within the CDS regions of 3,466 genes, which were defined as house-keeping (HK) genes. Shannon's equitability (E_H) was used to evaluate the level of gcHap diversity at all rice GLR loci in specific rice populations or the whole species (Sheldon, 1969). Nei's genetic identity ((I_{Nei}) was used to measure the genetic similarity between two populations or individuals based on their gcHap frequencies at different GLR loci (Nei, 1972). The formula for calculating Shannon's equability (E_H) and Nei's genetic identity (I_{Nei}) were referred to the method used in this article by Zhang (Zhang et al., 2021). The GraphPad Prism 9 software was used for statistical analysis and picture plotting of the obtained data.

The gcHaps diversity of rice GLR genes in modern varieties (MVs) and landraces (LANs)

To understand how modern breeding during past decades has affected gcHap diversity of rice GLR genes, detailed information was gathered for the total of 3,010 rice accessions of 3KRG. Of these, 732 were identified as *Xian* (*indica*) landraces (LANs-*Xian*), 358 were identified as *Geng* (*japonica*) landraces (LANs-*Geng*), 328 were identified as modern *Xian* varieties (MVs-*Xian*), and 139 were identified as modern *Geng* varieties (MVs-*Geng*). The first step was to download the predominant gcHaps (with the highest frequency in 3KRG) of each GLR gene from the RFGB webpage (https://www.rmbreeding.cn/Index). The frequency shift of the predominant gcHap for each GLR gene between modern varieties (MVs) and landraces (LANs) was calculated based on an R script (available at website https://github.com/isaac-Tsang/gcHap_diversity_LANs_MVs-). Then, A comparison was made between the distribution of gcHaps in modern *Xian* and *Geng* varieties and their respective landraces. At the same time, the loss and emergence of gcHap diversity were also analyzed in modern *Xian*/*Geng* varieties. Finally, the above data was plotted using GraphPad Prism 9 software.

Determination of phenotypic effects of the major gcHaps of rice GLR genes

Firstly, we compiled a large-scale phenotypic data for 15 agronomic traits across 3010 Asia cultivated rice accessions. The study examined 15 agronomic traits, including Days to Heading (DTH, day), Plant Height (PH, cm), Flag Leaf Length (FLL, cm), Flag Leaf Width (FLW, cm), Panicle Number (PN, count), Panicle Length (PL, cm), Culm Number (CN, count), Culm length (CL, cm), Grain Length (GL, mm), Grain Width (GW, mm), Grain Length/Width Ratio (GLWR, ratio), Thousand Grain Weight (TGW, g), Leaf Rolling Index (LRI, %), Seedling Height (SH, cm) and Ligule Length (LL, mm). Phenotypic data for 15 rice traits were downloaded from the RFGB website (https://www.rmbreeding.cn/Index/). Next, the major gcHaps (with frequency ≥1% in 3KRG) of all rice GLR genes were obtained using an R script (https://github.com/isaac-Tsang/haplotype_network). Finally, the associations of major gcHaps with these agronomic traits across the diverse 3,010 rice accessions were achieved using an R script (https://github.com/isaac-Tsang/functional-importance-). The significance was calculated using one-way analysis of variance, and Tukey's multiple comparison method was utilized to compare the significance among major gcHaps. The layout of the image is done in the Adobe Illustrator software (https://www.adobe.com/cn/products/illustrator.html).

Expression profiles analysis of rice GLR genes under normal and abiotic stress conditions

Firstly, we downloaded a rice expression database (RED), a repository of gene expression profiles for rice (Xia et al., 2017). The database includes rice gene expression data derived from RNA-Seq analysis of eight different tissues (aleurone, anthers, callus, leaves, panicles, pistils, roots, seeds and shoot) at different growth stages, as well as under a variety of biotic and abiotic treatments, including ABA, JA, Cd, drought, and cold treatment conditions. The expression profiles of all GLR genes are available in the Supplementary Data 2. The expression patterns of all rice GLR

genes from the dataset were analyzed using the heatmap.2 function of the R language and the layout of the image was done in the Adobe Illustrator software (https://www.adobe.com/cn/products/illustrator.html).

Knockout and field phenotypic experiments of rice GLR genes

Based on their strong expression in more rice tissues and/or their strong responses to abiotic stresses, five rice *GLR* genes, *OsGLR2.2*, *OsGLR9.8*, *OsGLR6.8*, *OsGLR4.1* and *OsGLR7.1*, were chosen for functional analysis using the CRISPR-Cas9 technology in three steps. First, we designed two target sequences and corresponding detection primers for each of the five GLR genes (Supplementary Data 3). Cloning and transformation of these GLR genes were commissioned by Wuhan Boyuan Biotechnology Co., Ltd, using a *Geng* variety, Nipponbare as a background material for genetic transformation. Screening of the transformants *via* PCR amplification and sequencing to determine those carrying the desired mutations for the five target GLR genes were described previously (Zeng et al., 2020). 2×Spark Taq PCR Master Mix (with dye) (Shandong Sparkjade Biotechnology Co., Ltd.) is used for PCR amplification. The PCR products were sent to Sangon Biotech (Shanghai, China) for Sanger sequencing. DNASTAR Lasergene software was used for sequence analysis and alignment (https://www.dnastar.com/software/lasergene/).

Field phenotypic experiments of these GLR gene knockout mutants were conducted in the short-day environment of Hainan (18°15' N and 109°30' E) from Nov. 25 of 2021 – April 10 of 2022 and in the long-day environment of Hefei (31°52' N and 117°17' E) from May 27 – Oct. 7 of 2022. Each homozygous GLR knockout mutant was planted in a 3-row plot with 8 plants per row at a spacing of 15 cm between individual plants and 20 cm between rows, with three replications for each of the knockout mutant lines and Nipponbare as the wild type check in the field experiments. Four agronomic traits, 1000-grain weight (TGW), seed setting rate (SR), panicle number (PN), panicle length (PL), and plant height (PH) were measured on the five middle plants in each plot. Statistical analyses of the phenotypic data of the experiments were performed using the GraphPad Prism 9 software, while the layout of the obtained picture plotting data was done using the Adobe Illustrator software (https://www.adobe.com/cn/products/illustrator.html).

Construction of gene CDS haplotype (gcHap) networks of rice GLR genes and association analyses of the major gcHaps at the GLR loci with yield traits in the 3KRG accessions

The construction of the gcHap networks containing major gcHaps with frequencies ≥1% in the 3KRG accessions or specific populations for each GLR gene was performed using the R package pegas (Paradis, 2010). The network of gcHaps for each GLR gene was produced using statistical parsimony algorithm, which is a method that connects the most closely related haplotypes first *via* the smallest number of mutations (Templeton et al., 1992). For more detailed steps, please refer to this *Molecular Plant* article of Zhang (Zhang et al., 2021). Finally, the layout of the image is done in the Adobe Illustrator software (https://www.adobe.com/cn/products/illustrator.html).

Results

The phylogenetic relations between plant GLRs and mammals iGluRs (ionotropic glutamate receptors)

We identified 26 GLR genes in the Os-Nipponbare-Reference-IRGSP-1.0 reference genome clustered primarily on chromosomes 2, 6, and 9. To understand the evolution and functional diversification of GLR genes, we performed a phylogenetic analysis to compare the rice GLR genes with their corresponding ones from various flowering plants, including *Arabidopsis thaliana* (*At*), *Zea mays* (*Zm*), *Solanum lycopersicum* (*Sl*), *Gossypium hirsutum* (*Gh*), and *Raphanus sativus* (*Rs*), three basal land plants, moss *Physcomitrium patens* (*Pp*), liverwort *Marchantia polymorpha* (*Mp*), and lycophyte *Selaginella moellendorffii* (*Sm*), one invertebrate *Caenorhabditis elegans* (*Ce*) and four mammals, AMPARs, NMDARs, KARs, and δ-receptors (Figure 1). Obviously, there is a clear distinction between GLRs in plants and animals, indicating a distant genetic relationship between the two groups. This is not unexpected as plants and animals represent distinct evolutionary lineages that diverged from a common ancestor more than a billion years ago. Consistent with our previous results (Shi et al., 2022), the 26 GLR genes could be roughly classified into four groups (Figure 1).

Group I contains a single *OsGLR* gene, *OsGLR6.1*, related to two *SlGLR* genes. Group II comprises 10 *OsGLR* genes in three subgroups: subgroup II-1 (*OsGLR9.1* – *OsGLR9.6*) related to maize *ZmGLR16*, subgroup II-2 (*OsGLR9.8*) closely related to maize *ZmGLR4* and *ZmGLR17*, and subgroup II-3 (*OsGLR9.7*, *OsGLR2.2* and *OsGLR6.2*) related to maize *ZmGLR3*, *ZmGLR6*, *ZmGLR12* and *ZmGLR13*. Group III has seven rice *OsGLR* genes in three subgroups: subgroup III-1 (*OsGLR6.11*, *OsGLR7.1* and *OsGLR9.9*) closely related to maize *ZmGLR14* and *ZmGLR15*, subgroup III-2 (*OsGLR6.12* and *OsGLR7.2*) related to maize *ZmGLR5*, *ZmGLR9*, *ZmGLR8* and *ZmGLR10*, and subgroup III-3 (*OsGLR4.1* and *OsGLR2.1*) related to maize *ZmGLR1*, *ZmGLR2*, *ZmGLR7* and *ZmGLR11*. Group IV is unique and has eight GLRs (*OsGLR6.4* – *OsGLR6.10*) found only in rice. The sequence differentiation of the 26 rice GLR genes suggested that the rice GLR genes had undergone unique evolutionary processes, including gene duplication events and functional differentiation that distinguished them from GLR genes in other species. The close relationships of most rice GLR genes with maize GLR genes suggest these rice GLRs might share similar functionalities with the maize ones, while the unique group IV GLRs clustered on rice chromosome 6 are the most interesting ones for future research focuses.

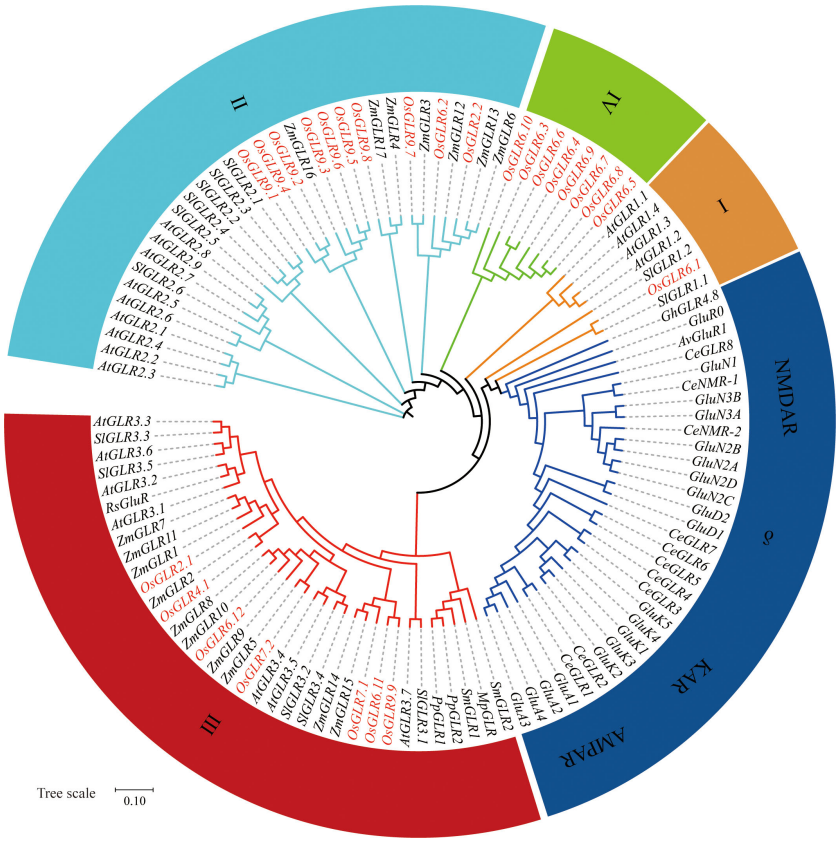


FIGURE 1
The phylogenetic relations between plant GLRs and mammals iGluRs (ionotropic glutamate receptors). The phylogenetic relations of glutamate receptors from *Oryza sativa Geng* (*Os*) based on the Os-Nipponbare-Reference-IRGSP-1.0 reference genome and other flowering plants that are described to show a conserved phenotype, including *Arabidopsis thaliana* (*At*), *Zea mays* (*Zm*), *Solanum lycopersicum* (*Sl*), *Gossypium hirsutum* (*Gh*), and *Raphanus sativus* (*Rs*), as well as basal land plants such as the moss *Physcomitrium patens* (*Pp*), the liverwort *Marchantia polymorpha* (*Mp*), and the lycophyte *Selaginella moellendorffii* (*Sm*), compared to the invertebrate *Caenorhabditis elegans* (*Ce*) and AMPARs, NMDARs, KARs, and δ-receptors from mammals (without prefix). Also included is the bacterial GluR0 from *Synechocystis PCC 6803* and AvGluR1 from the freshwater rotifer *Adineta vaga* (*Av*). *OsGLRs* are shown in colored text for clarity. Sequences were aligned using MUSCLE software, and the phylogenetic tree was constructed using the neighbor-joining method. Sequences are available in the Supplementary Data 1. AMPAR, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor; KAR, kainate receptor; NMDAR, N-methyl-d-aspartate receptor; δ, δ-receptor.

Functional differentiation and tissue specificity of rice GLR genes

To understand the functionalities of different rice GLR genes, we analyzed the expression levels of all GLR genes in eight different tissues of rice, including aleurone, anther, callus, leaf, panicle, pistil, root, seed and shoot tissues. Under the normal conditions, different GLR genes showed varied expression levels in different tissues of rice (Figure 2). Notably, *OsGLR2.1*, *OsGLR4.1*, *OsGLR6.5*, *OsGLR6.8*, *OsGLR6.10*, *OsGLR6.12*, *OsGLR7.1*, *OsGLR7.2*, and *OsGLR9.8* were each expressed strongly (mean FPKM > 5) in more than one tissue, indicating their importance in rice growth and development. Among these genes, *OsGLR4.1* had the highest expression levels in leaf and seed tissues (Figure 2). *OsGLR2.2*, *OsGLR6.3* and *OsGLR6.9* showed moderate levels of expression in one or more tissues, while the remaining ones (*OsGLR6.1*, *OsGLR6.2*, *OsGLR6.4*, *OsGLR6.6*, *OsGLR6.9*, *OsGLR6.11*, *OsGLR9.7* and *OsGLR9.9*) showed low levels of expression (<1) in all tissues. Interestingly, four GLRs (*OsGLR9.3* - *OsGLR9.6*) showed a moderate expression specifically in the pistil tissue but extremely low expression in other tissues. These results indicated that one important aspect in the functional differentiation of rice GLR genes was reflected in their tissue specificity, though it remains a mystery regarding how the expression pattern of each rice GLR gene related its functions in its specifically expressed tissues.

Expression profiles of rice GLR genes in response to abiotic stresses

In order to gain insights into the roles of rice GLR genes under different abiotic stress conditions, we conducted an integrative analysis of their expression profiles under abscisic acid (ABA), JA, cadmium (Cd), drought, and cold treatments. When treated with ABA, rice *OsGLR6.5*, *OsGLR6.8*, *OsGLR6.10*, *OsGLR7.2* and *OsGLR9.8* showed increased expression in both shoots and roots (Figure 3A), indicating that they might be involved in the ABA regulated pathways. Interestingly, *OsGLR4.1* and *OsGLR9.8* showed

10.3389/fpls.2023.1236251

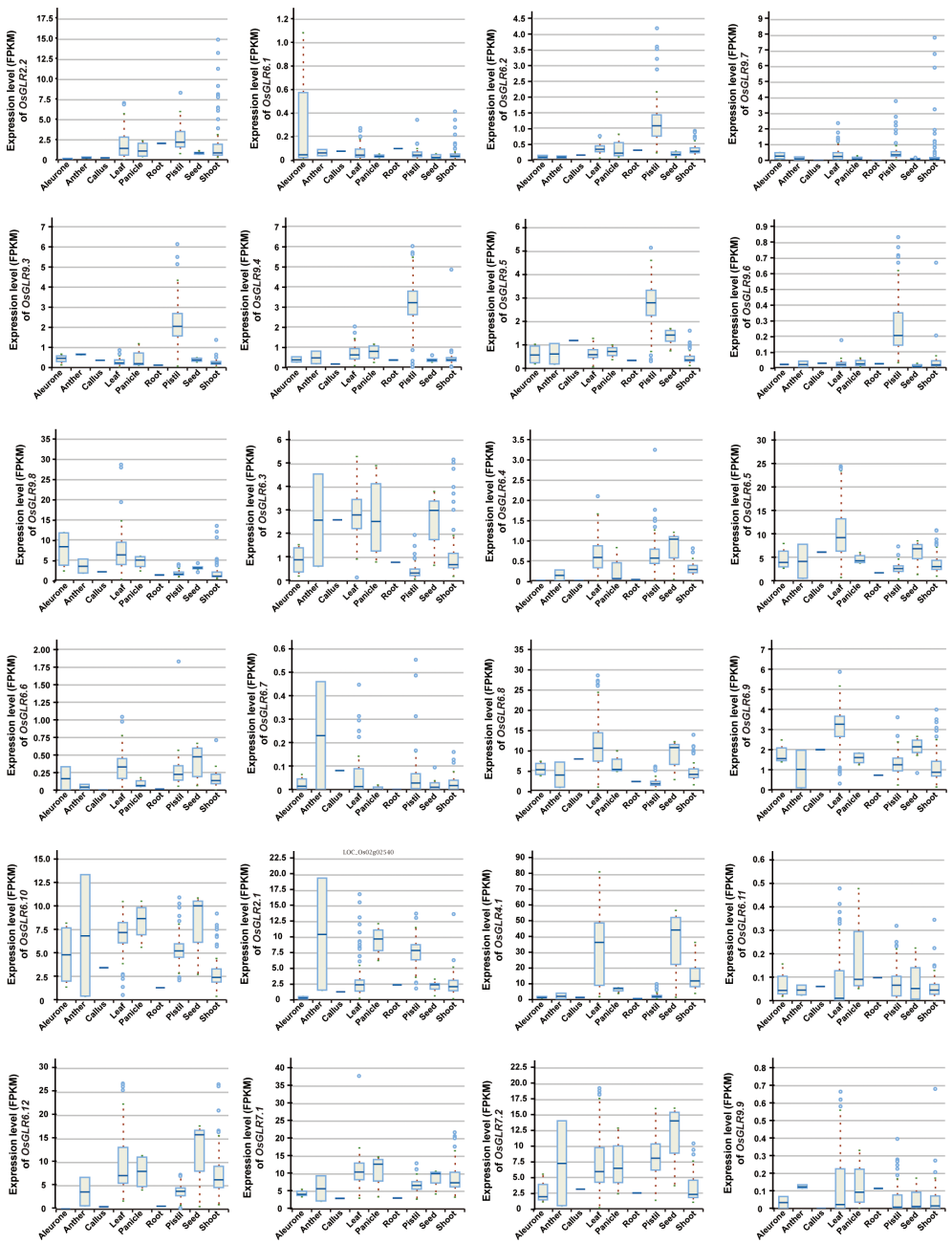


FIGURE 2
The expression level of the *OsGLR* genes in different rice tissues under the normal conditions. The plant tissues mainly include aleurone, anther, callus, leaf, panicle, pistil, root, seed and shoot.

opposite expression patterns in shoot under the ABA treatment (Figure 3A), in which *OsGLR4.1* showed gradually decreased expression in response to the ABA treatment over time, whereas *OsGLR9.8* showed gradually increased expression over time (Figure 3A). Under three different Cd (cadmium) ion stress conditions, *OsGLR2.1*, *OsGLR4.1*, *OsGLR6.5*, *OsGLR6.8*, *OsGLR7.1*, *OsGLR7.2* and *OsGLR9.8* showed increased but varied expression (Figures 3B, C). Among them, *OsGLR4.1* and *OsGLR7.1* genes had the highest expression in both roots and shoots under different concentrations of Cd stress. Under the JA treatment, *OsGLR2.1*, *OsGLR7.1*, *OsGLR7.2*, *OsGLR6.5*, *OsGLR6.8* and *OsGLR6.10* showed significantly increased expression in either or both roots and shoots (Figure 3D), indicating they might be involved in the JA signaling pathways. Interestingly, *OsGLR4.1* showed a similarly decreasing expression in shoots under the JA treatment as it did in response to the ABA treatment. Under the cold stress, *OsGLR2.1*, *OsGLR2.2*, *OsGLR7.1*, *OsGLR7.2*, *OsGLR6.5*, *OsGLR6.10*, *OsGLR6.12*, and *OsGLR9.8*, showed significantly increased but varied expression in either or both tissues (Figure 3E). Similarly, the expression of *OsGLR2.1*, *OsGLR2.2*, *OsGLR7.1*, *OsGLR4.1*, *OsGLR6.12* and *OsGLR9.8* was significantly upregulated by the drought treatment (Figure 3F). Taking together,

10.3389/fpls.2023.1236251

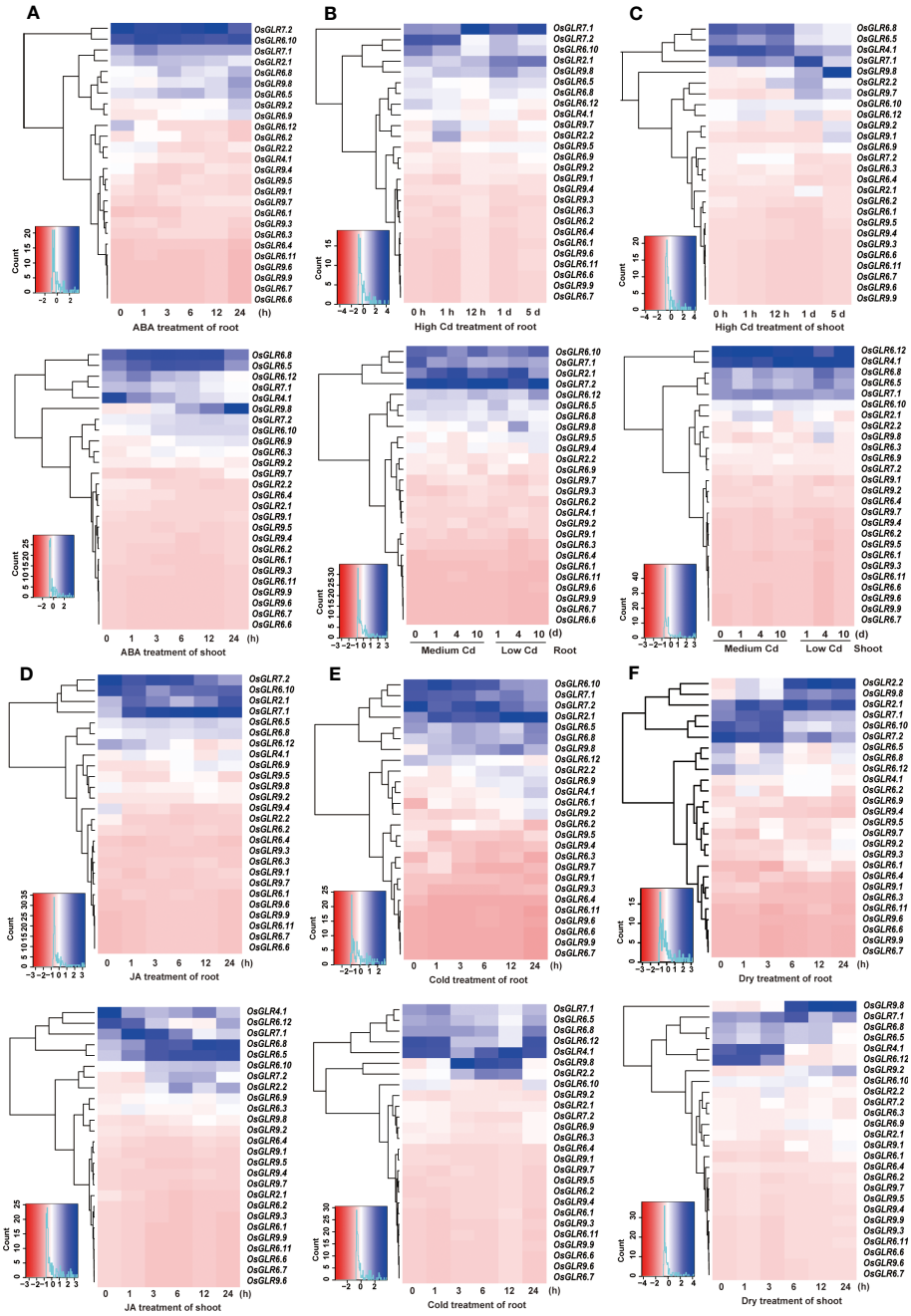


FIGURE 3
Integrative analysis of the rice *OsGLR* gene family expression profiles under ABA **(A)**, Cd **(B, C)**, JA **(D)**, cold **(E)**, and dry **(F)** treatment conditions.

ten rice *GLR* genes showed significantly increased expression in response to the five external treatments. Of these, *OsGLR7.1*, *OsGLR2.1* and *OsGLR9.8* were the most noteworthy ones. The expression of *OsGLR7.1* in both roots and shoots and expression of *OsGLR2.1* in roots were strongly upregulated in response to JA, Cd, drought and cold, suggesting their important roles in the JA signaling pathways involved in abiotic stresses. In contrast, the expression of *OsGLR9.8* in shoots was strongly upregulated under the Cd, drought and cold stresses, indicating its important role in rice responses to abiotic stresses.

Functional determination of five rice GLR genes by knockout and phenotypic assessment

To assess the functional importance of rice GLR genes, five important *OsGLR* genes, *OsGLR2.2*, *OsGLR9.8*, *OsGLR6.8*, *OsGLR4.1* and *OsGLR7.1* were selected based on previous transcriptomic analyses for further gene knockout experiments using the CRISPR-Cas9 technology. Specifically, three independent homozygous mutants carrying different mutation

10.3389/fpls.2023.1236251

sites for each of the GLR genes were obtained (Supplementary Figure 1) and assessed for 1000-grain weight (TGW), seed setting rate (SR), panicle number (PN), panicle length (PL) and plant height (PH) under the short-day and long-day environments of Hainan and Hefei. Three interesting observations were noted regarding the phenotypic differences for the measured traits between the GLR knockout mutants and the wild type (Figure 4). First, significant differences were observed between the knockout mutants of each GLR gene and the wild type for all the measured traits, but these differences varied greatly depending on the environments where these materials were assessed (Figures 4A–E). In other words, all five GLR genes showed huge genotype x environment interactions for their phenotypic effects on the measured traits, particularly for SR, PN and PH. When compared with the wild type, the mutant lines of all five GLR genes showed greatly reduced SR under the short-day environment of Hainan, but their SR was either significantly increased (*OsGLR7.1* and *OsGLR9.8*), unchanged (*OsGLR4.1* and *OsGLR6.8*) or varied

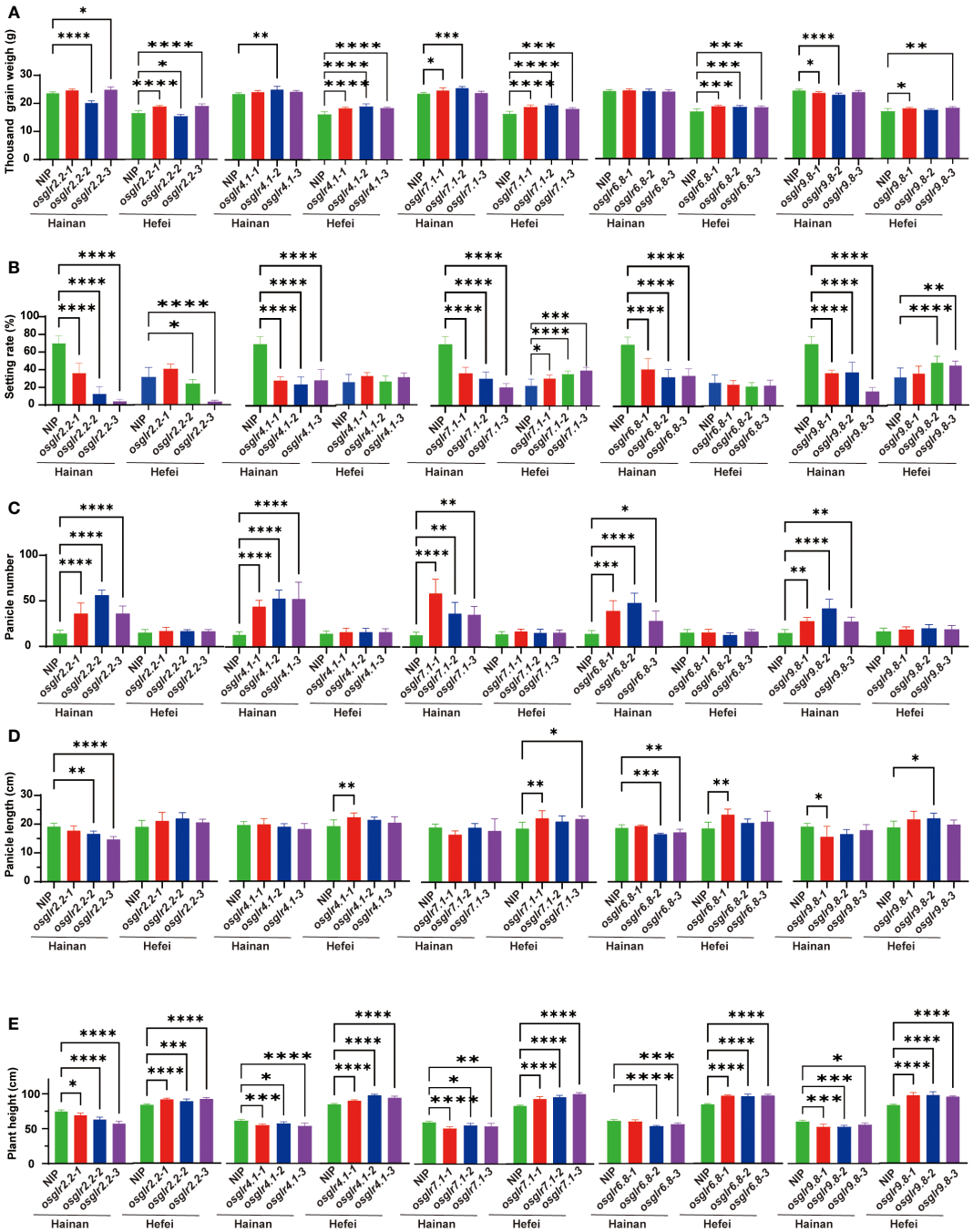


FIGURE 4
Phenotypic comparisons for five agronomic traits of knockout mutants of five rice *OsGLR* genes evaluated in two environments. **(A-E)** The agronomic traits represented by **(A-E)** are 1000-grain weight (g), seed setting rate (%), panicle number, panicle length (cm) and plant height (cm), respectively. *, **, *** and **** meant statistically significant at a level of 0.05, 0.01, 0.001 and 0.0001, respectively.

(*OsGLR2.2*) in the long-day environment of Hefei (Figure 4B). The mutant lines of all five GLR genes showed greatly increased PN under the short-day environment of Hainan, but not in the long-day environment of Hefei (Figure 4C). More interestingly, the mutant lines of all five GLR genes showed significantly reduced PH under the short-day environment of Hainan, but significantly increased PH in the long-day environment of Hefei (Figure 4E). Secondly, one or more mutant lines of the GLR genes, except for *OsGLR6.8* in Hainan, showed significant increased TGW in both environments (Figure 4A). Thirdly, the phenotypic differences between the mutant lines and wild type varied among different mutant lines of the same GLR genes, particularly for PL (Figure 4D), indicating that mutation sites in different GLR genes did make differences in generating phenotypically detectable mutants. These results led us to an important conclusion, i.e. obvious functional differentiation and functional redundancy regarding their phenotypic effects on the measured traits, the five rice GLR genes appeared to play an important role in contributing to the significant G x E interaction of agronomic traits in rice across the short-day and long-day environments.

The rich genetic and allelic diversity at rice GLR loci in rice populations

Given the functional importance of the rice GLR genes, it is important to understand their genetic diversity in rice populations in order to facilitate their applications in rice improvement. Using the gene CDS haplotype (gcHap) data from the 3,010 rice genome project (3KRG), we obtained the Shannon's equitability, E_H, and number of gcHaps and major gcHaps (with frequency ≥1% in 3KRG) of the 26 GLR genes in four major rice populations (Supplementary Data 4, Figure 5). Of the 26 GLR genes, *OsGLR6.2* and *OsGLR6.3* are extremely conserved house-keeping (HK) genes with $E_H = 0$, and gcHapN = 1 (Figure 5A), indicating all rice varieties share the same alleles at the two loci. For the remaining 24 polymorphic GLR genes, the average numbers of gcHaps (major gcHaps) and E_H were 476.4 (8.4) and 0.368 in the 3KRG accessions. However, different GLR genes vary considerably in their genetic diversity. *OsGLR9.4* has the highest genetic diversity with E_H of 0.821 and 2,182 gcHaps (Figures 5C, D), followed by *OsGLR9.2* with E_H of 0.691 and 1,681 gcHaps, and *OsGLR9.5* with E_H of 0.649 and 1,374 gcHaps. *OsGLR6.1* has the lowest diversity with E_H of 0.050 and six gcHaps (Supplementary Data 4).

Different rice populations also vary considerably in their diversity of the GLR genes. The mean E_H of the 24 GLR genes were 0.367 in population *Xian*, 0.213 in population *Geng*, 0.409 in population *Aus*, and 0.352 in population *Bas*. The numbers of detected gcHaps and major gcHaps were 342.1 and 7.7 in population *Xian*, 76.2 and 4.8 in population *Geng*, 47.8 and 5.2 in population *Aus*, and 17.8 and 5.0 in population *Bas* (Supplementary Data 4). Clearly, the numbers of gaHaps and major gcHaps in different rice populations were largely determined by their population sizes, implying that most gcHaps are rare ones with very low frequencies in different rice populations (Figures 5C, D).

To understand how different GLR genes were differentiated among major rice populations, we estimated the Nei's genetic identity (I_{Nei}) of all GLR genes using the gcHap data of the 24 polymorphic *OsGLR* genes, except for *OsGLR6.2*, *OsGLR6.3*, between all pairwise populations. Except for *OsGLR6.1*, *OsGLR6.5* and *OsGLR6.11*, strong *Xian-Geng* sub-specific differentiation (I_{Nei}<0.35) was observed at 21 of the GLR genes, 20, 17 and 18 of which also showed strong *Xian-Bas*, *Aus-Geng* and *Aus-Bas* differentiation (Figure 6, Supplementary Data 5), indicating that the allelic differentiation at most GLR loci had contributed very strongly to the differentiation of major rice populations and expectedly to the adaptations of different rice populations to their environments.

Impact of modern breeding on gcHap diversity of rice GLR genes

To understand how modern breeding during past decades has affected gcHap diversity of rice GLR genes, we analyzed the diversity parameters of all the GLR genes in modern varieties (MVs) and landraces (LANs), including 732 *Xian* landraces (LANs-*Xian*), 358 *Geng* landraces (LANs-*Geng*), 328 modern *Xian* varieties (MVs-*Xian*), and 139 modern *Geng* varieties (MVs-*Geng*) (Supplementary Data 6). In population *Xian*, the average E_H of the 24 GLR genes was 0.427 in MVs-*Xian*, a 13.3% increase compared to 0.377 in LANs-*Xian* (Table 1). Increased diversity in MVs-*Xian* was observed in eight GLR genes (*OsGLR6.4*, *OsGLR6.5*, *OsGLR6.6*, *OsGLR6.9*, *OsGLR6.10*, *OsGLR9.7*, *OsGLR7.1* and *OsGLR7.2*) out of the 24. No GLR loci showed reduced diversity. Interestingly, MVs-*Xian* had an average of 85 gcHaps/locus, significantly lower than the 156.7 gcHaps/locus in LANs-*Xian*. Detailed examination revealed that the MVs-*Xian* lost average 134.4 gcHaps/locus and 1.6 major gcHaps/locus that were present the LANs-*Xian*, but absent in MVs-*Xian*, which resulted apparently from the genetic bottlenecks during modern breeding (Table 1, Figure 7A).

Additionally, the proportion of newly emerged gcHaps observed in modern varieties is significantly lower than that of lost gcHaps (Figures 7B, C). MVs-*Xian* gained average 62.8 new gcHaps/locus and 2.2 new major gcHaps that were absent in the LANs-*Xian*. The new gcHaps were apparently generated by intragenic recombination during breeding, while all new major gcHaps were rare ones in the LANs-*Xian* but became major ones with significantly increased frequencies in the MVs-*Xian* (Table 1). We noted four GLR loci (*OsGLR9.2* - *OsGLR9.5*) in a large cluster on chromosome 9 (Figure 7A), each of which had a huge number of gcHaps and showed more dramatically changes in gcHapN, suggesting genetic drift and unequal crossing-over occurred more frequently among duplicated GLR genes in the GLR gene clusters. The increased diversity and reduced gcHapN observed at many of the GLR loci suggested that significant frequency shifts in major gcHaps had occurred at these GLR loci in population *Xian* during breeding. Indeed, significant shifts in frequencies of the predominant gcHaps, $F_{(P)}$, were observed at 13 of the 24 GLR loci, including significantly reduced $F_{(P)}$ at nine of the GLR loci and significantly increased $F_{(P)}$ at four of the GLR loci (Supplementary

10.3389/fpls.2023.1236251

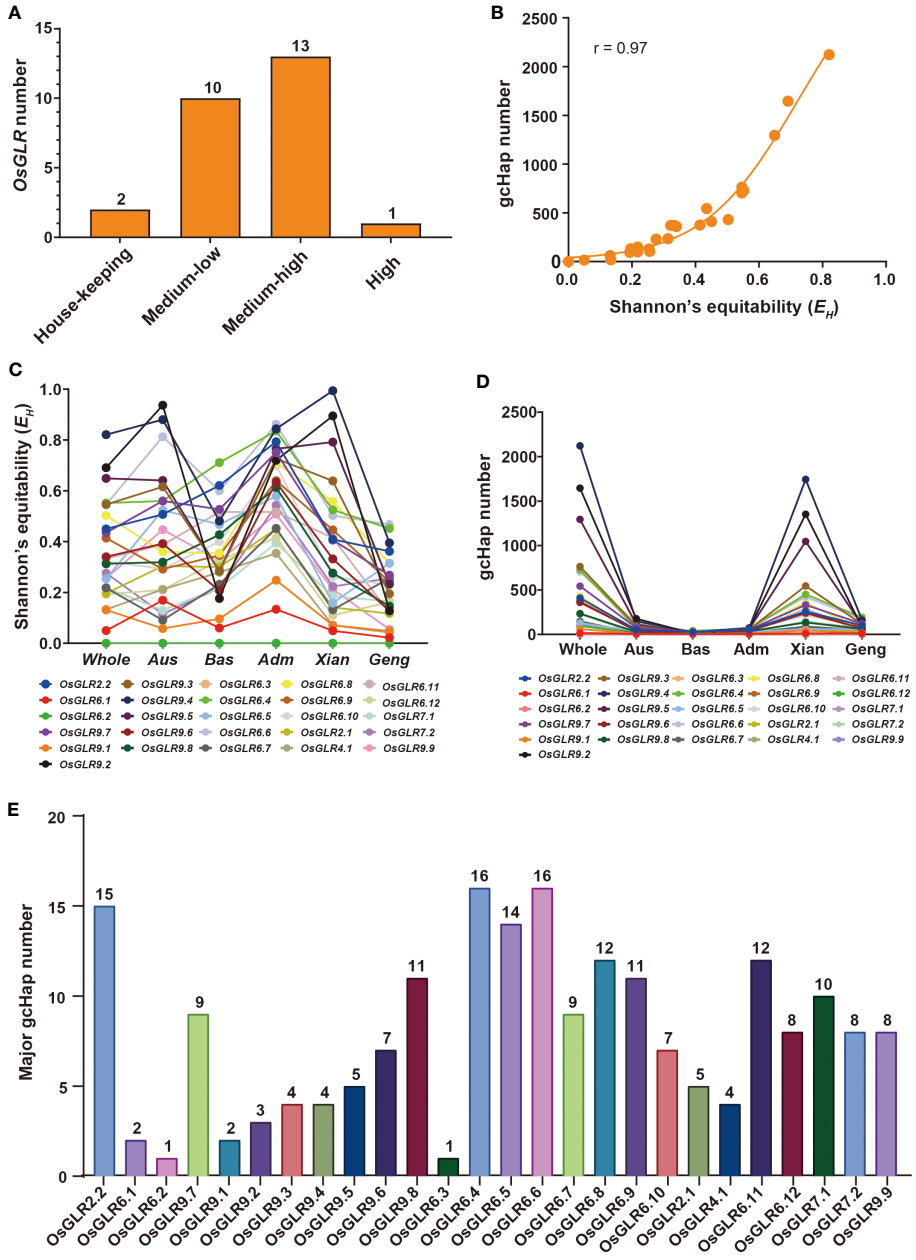


FIGURE 5
Population diversity of *OsGLR* genes in 3010 Asia cultivated rice accessions. **(A)** *OsGLR* genes distribution of five diversity gene types (house-keeping (HK) genes without SNP variants, low-diversity genes ($0 < E_H < 0.05$, $E_H = 0.020 \pm 0.015$, and gcHapN = 12 ± 10), medium- to low-diversity genes ($0.05 \leq E_H < 0.3$, $E_H = 0.171 \pm 0.065$, and gcHapN = 82 ± 67), medium- to high-diversity genes ($0.3 \leq E_H < 0.7$, $E_H = 0.444 \pm 0.110$, and gcHapN = 498 ± 305) and a high-diversity gene ($0.7 \leq E_H \leq 1$, $E_H = 0.804 \pm 0.075$, and gcHapN = 1767 ± 458), Shannon's equitability (E_H) takes into account both the number of species present (species richness) and the relative abundance of each species within the community. It ranges from 0 to 1, with values closer to 1 indicating a more even distribution of species abundance, and values closer to 0 indicating a less even distribution. **(B)** The relationship between Shannon's equitability (E_H) and gcHapN of *OsGLR* genes. **(C)** E_H values of *OsGLR* genes in different population organization. **(D)** gcHap number (gcHapN) of *OsGLR* genes in different population organization. **(E)** The major gcHap (≥ 30 rice accessions) number of *OsGLR* genes.

Data 7). Notably, at *OsGLR7.1* locus, the predominant allele (Hap1) had an $F_{(P)}$ of 0.60 and 0.34 in populations LANs-*Xian* and MVs-*Xian*, or a major reduction in $F_{(P)}$ by 0.25 in MVs-*Xian* (Supplementary Figure 3), accompanied by significantly increased $F_{(P)}$ of five newly major gcHaps, all of which were rare gcHaps with frequencies <0.01 in the LANs-*Xian* but became major gcHaps in the MVs-*Xian*. Similar observations were noted on four other loci (*OsGLR6.11*, *OsGLR9.1*, *OsGLR6.5* and *OsGLR6.4*) where significantly reduced $F_{(P)}$ was accompanied by emergences of 1-2 new major gcHaps in the MVs-*Xian* from rare ones in the in the LANs-*Xian*. The remaining 11 loci showing no significant differences in $F_{(P)}$ between LANs-*Xian* and MVs-*Xian* included the four GLR loci (*OsGLR9.2* - *OsGLR9.5*) where no or few predominant gcHaps were present.

In population *Geng*, the average E_H of the 24 GLR genes was similar in the MVs-*Geng* (0.262) to 0.257 of the LANs-*Geng*

10.3389/fpls.2023.1236251

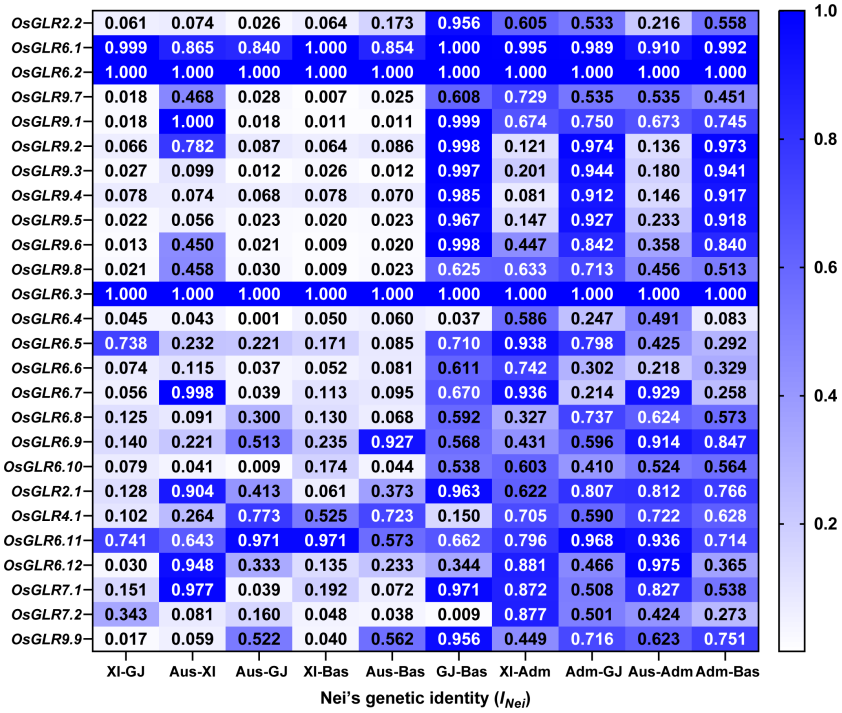


FIGURE 6
The Nei's genetic identity (I_{Nei}) of rice *OsGLR* genes between all pairwise populations calculated from the gcHap data.

(Supplementary Data 6). However, significant reduced diversity in the MVs-*Geng* was observed at only two GLR loci (*OsGLR6.4* and *OsGLR6.6*), while significantly increased diversity was observed at *OsGLR2.2*, *OsGLR7.1* and *OsGLR9.7* genes. Similarly, the MVs-*Geng* had an average of 20.6 gcHaps/locus, significantly less than 42.6 gcHaps/locus in the LANs-*Geng* (Figure 7; Table 1). In fact, the MVs-*Geng* lost an average 35.3 gcHaps/locus and 2.8 major gcHaps/locus that were present in the LANs-*Geng* but absent in the MVs-*Geng*, while the MVs-*Geng* gained an average 13.5 new gcHaps/locus and 3.0 new major gcHaps that were present in the MVs-*Geng* but absent in the LANs-*Geng* (Table 1). The new gcHaps were apparently generated by intra-genic recombination during breeding, while the new major gcHaps were all come from minor rare gcHaps in the population LANs-*Geng*. Significant shifts in frequencies of the predominant gcHaps, $F_{(P)}$, were observed at 12 of the 24 *OsGLR* loci, including significantly reduced $F_{(P)}$ at *OsGLR6.7*, *OsGLR6.9*, *OsGLR7.1* and *OsGLR9.7* and significantly increased $F_{(P)}$ at *OsGLR6.4*, *OsGLR6.5*, *OsGLR6.6*, *OsGLR6.8*, *OsGLR6.10*, *OsGLR9.2*, *OsGLR9.5* and *OsGLR7.2* (Supplementary Data 7). It should be pointed out that populations *Xian* and *Geng* have different predominant gcHaps at all polymorphic GLR loci, except for *OsGLR6.1* and *OsGLR9.4*. This was because there is only a single predominant in both populations *Xian* and *Geng* at *OsGLR6.1* while there is absence of any gcHaps at *OsGLR9.4* whose frequency was higher than 1% in population *Xian*.

Associations of the major gcHaps at rice GLR genes with important agronomic traits

To provide additional evidence for the functional importance of the 24 polymorphic GLR genes, we constructed the gcHap networks of the major alleles at 24 GLR genes in the four major rice populations, and analyzed the associations between the major gcHaps at each of the 24 *GLR* gene loci with phenotypic values of four agronomic traits, plant height (PH), panicle length (PL), panicle number per plant (PN) and thousand grain weight (TGW) in 3KRG (Figure 8; Supplementary Figures 4-7). Strong ($P < 10^{-7}$) associations were observed in 61 (63.5%) of the 96 (24 x 4) cases and major alleles at many of the GLR gene were strongly associated with trait values of one or more traits. Notably, major alleles at *OsGLR2.2*, *OsGLR4.1*, *OsGLR7.1*, *OsGLR6.8* and *OsGLR9.8* were strongly associated with trait values of all four traits, except for PL in two cases (Figure 8), consistent with the results from above knock-out experiments (Figure 4). In particular, major alleles at ten GLR loci (*OsGLR6.4*, *OsGLR6.5*, *OsGLR6.6*, *OsGLR6.7*, *OsGLR6.8*, *OsGLR6.9*, *OsGLR6.10*, *OsGLR6.12 OsGLR7.1* and *OsGLR7.2*) were strongly associated with trait values of all four traits (Figure 8; Supplementary Figures 4-7). In contrast, no associations were detected for major alleles at *OsGLR6.1* and *OsGLR9.1* with the trait values of the four traits, largely because of too few major gcHaps present at the three loci (Supplementary Data 4).

TABLE 1 Comparisons in genetic diversity of the 26 GLR genes between the landraces and modern varieties.

Gene name	*Xian* (*indica*)							*Geng* (*japonica*)						
	LANs		MVs		Change of gcHapN (major gcHapN)			LANs		MVs		Change of gcHapN (major gcHapN)		
	E_H	gcHapN (major gcHapN)	E_H	gcHapN (major gcHapN)	Lost	New	Retained	E_H	gcHapN (major gcHapN)	E_H	gcHapN (major gcHapN)	Lost	New	Retained
OsGLR2.1	0.177	32 (4)	0.233	29 (5)	19 (2)	16 (1)	13 (3)	0.098	17 (2)	0.151	13 (4)	12 (3)	8 (1)	5 (1)
OsGLR2.2	**0.386**	**89 (13)**	**0.440**	**62 (12)**	**33 (1)**	**60 (2)**	**29 (11)**	**0.349**	**44 (7)**	**0.477**	**35 (8)**	**31(4)**	**22 (3)**	**13 (4)**
OsGLR4.1	**0.077**	**7 (2)**	**0.081**	**6 (3)**	1 (1)	0	6 (2)	**0.044**	**4 (4)**	**0.035**	**3 (2)**	1 (1)	2 (3)	2 (1)
OsGLR6.1	0.031	4 (2)	0.027	3 (3)	1 (1)	0	3 (2)	0.012	3 (1)	0.000	1 (1)		0	1(1)
OsGLR6.4	0.392	109 (12)	0.554	98 (13)	77 (3)	66 (2)	32 (10)	0.636	103 (19)	0.475	41 (11)	86 (4)	24 (12)	17 (7)
OsGLR6.5	0.255	35 (11)	0.389	34 (11)	14 (2)	13 (2)	21 (9)	0.368	33 (10)	0.395	24 (12)	21 (4)	12 (2)	12 (8)
OsGLR6.6	0.588	196 (13)	0.709	140 (15)	149 (5)	93 (3)	47 (10)	0.603	105 (13)	0.490	44 (10)	85 (2)	24 (5)	20 (8)
OsGLR6.7	0.115	25 (4)	0.164	21 (4)	10 (1)	6 (1)	15 (3)	0.263	19 (7)	0.264	12 (5)	11	4 (2)	8 (5)
OsGLR6.8	**0.312**	**60 (8)**	**0.367**	**35 (9)**	40 (2)	15 (1)	20 (7)	**0.305**	**21 (8)**	**0.270**	**11 (8)**	15 (4)	5 (4)	6 (4)
OsGLR6.9	0.423	113 (12)	0.507	74 (15)	78 (4)	39 (1)	35 (11)	0.268	46 (3)	0.314	27 (5)	42 (2)	23	4 (3)
OsGLR6.10	0.230	91 (6)	0.326	61 (5)	75 (1)	45 (2)	16 (4)	0.299	55 (3)	0.271	24 (4)	49 (2)	18 (1)	6 (2)
OsGLR6.11	0.502	112 (17)	0.474	66 (14)	76 (1)	31 (4)	35 (13)	0.109	12 (3)	0.150	8 (3)	9	5	3 (3)
OsGLR6.12	0.202	35 (5)	0.166	24 (5)	21 (1)	10 (1)	14 (4)	0.167	26 (2)	0.190	12 (8)	20 (6)	6	6 (2)
OsGLR7.1	**0.266**	**51 (7)**	**0.385**	**38 (10)**	32 (5)	19 (2)	19 (5)	**0.335**	**33 (8)**	**0.420**	**24 (8)**	22 (2)	13 (2)	11 (6)
OsGLR7.2	0.232	80 (5)	0.325	52 (7)	66 (2)	38	14 (5)	0.343	53 (4)	0.326	20 (6)	47 (2)	14	6 (4)
OsGLR9.1	0.019	2 (2)	0.037	2 (2)	0	0	2 (2)	0.012	3 (2)	0.042	3 (2)	0	0	3 (2)
OsGLR9.2	0.895	587 (2)	0.930	275 (3)	549 (2)	237 (1)	38 (1)	0.188	52 (4)	0.120	15 (1)	51	14 (3)	1 (1)
OsGLR9.3	0.862	538 (2)	0.873	249 (4)	500 (3)	211 (1)	38 (1)	0.324	89 (3)	0.271	33 (1)	85	29 (2)	4 (1)
OsGLR9.4	0.998	728 (0)	0.998	325 (0)	727	324	1	0.477	102 (6)	0.450	38 (6)	92 (1)	28 (1)	10 (5)
OsGLR9.5	0.862	515 (7)	0.872	233 (8)	470 (3)	188 (2)	45 (5)	0.337	74 (4)	0.275	27 (2)	68	21 (2)	6 (2)
OsGLR9.6	0.367	132 (5)	0.408	70 (7)	107 (3)	45 (1)	25 (4)	0.175	49 (1)	0.180	19 (5)	44 (4)	14	5 (1)

(Continued)

10.3389/fpls.2023.1236251

10.3389/fpls.2023.1236251

TABLE 1 Continued

Gene name	*Xian (indica)*							*Geng (japonica)*						
	LANs		MVs		Change of gcHapN (major gcHapN)			LANs		MVs		Change of gcHapN (major gcHapN)		
	E_H	gcHapN (major gcHapN)	E_H	gcHapN (major gcHapN)	Lost	New	Retained	E_H	gcHapN (major gcHapN)	E_H	gcHapN (major gcHapN)	Lost	New	Retained
OsGLR9.7	0.347	110 (8)	0.468	77 (10)	82 (3)	49 (1)	28 (7)	0.243	39 (6)	0.366	32 (6)	28 (2)	21 (2)	11 (4)
OsGLR9.8	**0.300**	**72 (7)**	**0.337**	**46 (8)**	48 (2)	22 (1)	24 (6)	**0.166**	**31 (3)**	**0.235**	**20 (5)**	24 (2)	13	7 (3)
OsGLR9.9	0.205	38 (7)	0.186	21 (6)	23 (1)	6 (2)	15 (5)	0.056	9 (2)	0.115	9 (4)	3 (2)	3	6 (2)
Mean	**0.377**	**157 (6.7)**	**0.427**	**85 (7.5)**	134.4 (2.2)	62.8 (1.6)	22.3 (5.7)	**0.257**	**43 (5.2)**	**0.262**	**21 (5.3)**	35.3 (2.8)	13.5 (3.0)	7.2 (3.3)

E_H and gcHapN (major gcHapN) are the Shannon's equitability, and number of gcHaps (≥ 1% varieties) identified. The bold GLR genes were those used in functional analyses by the CRISPR-Case9 knockout experiments.

Comparison of trait values between the predominant and 'unfavorable' gcHaps of GLR genes in rice

In contrast to the predominant gcHap (with highest frequency) at a GLR locus in a rice population would be considered to be favored by natural selection during evolution, the major gcHap with the lowest frequencies in the population would more likely to be the 'unfavorable' one. To test if this scenario was holding true, we compared the phenotypic differences between the predominant gcHap and 'unfavorable' one at each of the 24 GLR loci for 15 agronomic traits. Of the total 24 x 15 = 360 comparisons, significant differences were detected in 149 (41.4%) cases between the predominant gcHaps and unfavorable at all GLR gene loci except for *OsGLR6.8*, *OsGLR9.7* and *OsGLR9.8* (Supplementary Figure 8-31; Supplementary Data 8). Of the 15 traits, DTH was detected at 16 (66.7%) of the GLR loci, followed by PN at 14 (58.3%) of the GLR loci, consistent with the result of the gene knockout experiments that the five knockout GLR genes had large and consistent effects on PN and their phenotypic effects were influenced strongly by the two environments of contrasting day-lengths that are known to great influence on DTH of different rice varieties (Figure 4). Across the 24 GLR loci, significant differences in mean trait values between the predominant gcHap and unfavorable one were detected for almost all traits for *OsGLR9.1* (14 traits), *OsGLR9.9* (14 traits), and *OsGLR9.2* (13 traits), indicating these three GLR genes were more likely affecting more traits than other GLR loci (Supplementary Data 8).

Mining 'favorable' allele(s) of rice GLR loci for improved productivity

As described above, there were significant differences between or among the major gcHaps at most rice GLR loci regarding their effects on traits of agronomic importance. Thus, it is reasonable to assume the presence of 'favorable' gcHap(s) at most GLR loci for traits leading to increased productivity, even though the 'favorable' gcHap(s) may differ across different target environments. Thus, there should be one or more 'favorable' alleles, defined as the gcHap (s) associated with trait values for high productivity. Then, the frequencies of these most 'favorable' gcHaps in the MVs and different rice populations would of great interest for rice breeders. Figure 9 and Supplementary Figure 32 show the frequencies of the 'favorable' gcHaps at 24 rice GLR genes that affects PN, TGW, PH and PL in populations MVs-*Xian* and MVs-*Geng* and the four major rice populations of 3KRG. The general observation was that the 'favorable' allele frequency varied considerably across different yield traits, different GLR loci, and across populations MVs-*Xian* and MVs-*Geng* or different rice populations. For example, the Hap3 at *OsGLR9.2* and Hap5 at *OsGLR6.12* were associated with the highest PN in 3KRG accessions, which reached fixation in MVs-*Xian*, but had low frequencies in MVs-*Geng* (Figure 9). The same was true for eight GLR loci (*OsGLR9.2* - *OsGLR9.6*, *OsGLR6.7*, *OsGLR6.8* and *OsGLR6.12*) with the highest TGW. In contrast, Hap1 at *OsGLR2.2*, Hap12 at *OsGLR6.5*, Hap6 at *OsGLR6.9*, and

10.3389/fpls.2023.1236251

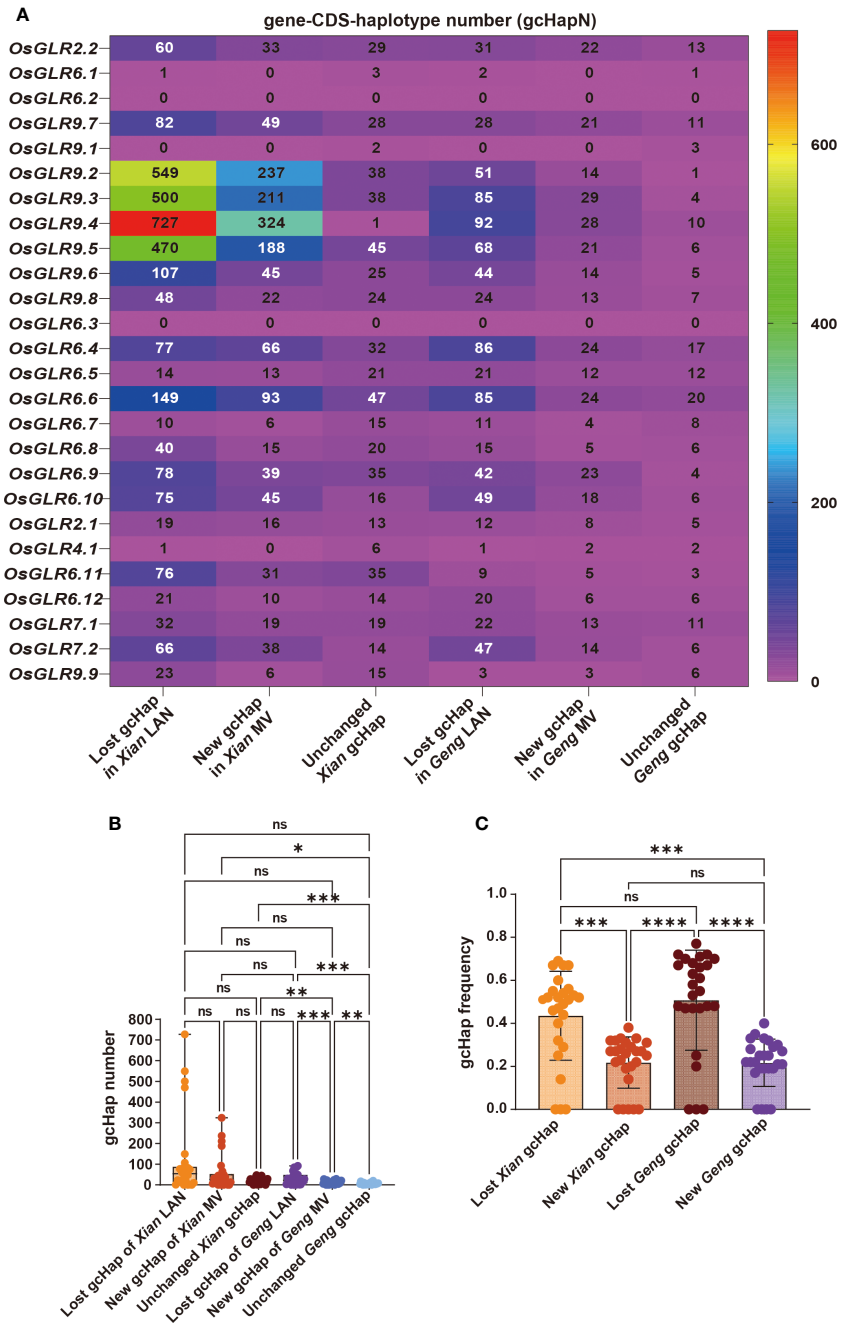


FIGURE 7
The lost, new and unchanged gcHap variation of rice 26 *OsGLR* genes during modern breeding. **(A)** Heatmap of the numbers of lost, newly emerged, and unchanged gcHaps. **(B)** The comparative analysis of the numbers of lost, newly emerged, and unchanged gcHaps. **(C)** The frequency comparison of lost, newly emerged, and unchanged gcHaps. The Tukey's multiple comparison method was used. ****P < 0.0001, ***P < 0.001, **P < 0.01, *P < 0.05, and n.s., not significant.

Hap5 at *OsGLR6.10* were associated with the highest PN in 3KRG accessions, which reached fixation in MVs-*Geng*, but had low frequencies in MVs-*Xian*. This was consistent with the empirical observation in rice breeding that the extremely high values for any specific yield components may not generally result in the highest productivity. This result also suggested that the 'favorable' alleles for different yield traits could be very different across the two rice subspecies or genetic backgrounds and possibly across different environments.

Discussion

Although glutamate receptors are specialized proteins that perform a crucial role in signal transduction and communication in both animals and plants. Present as a large gene family in the rice genome, the molecular properties and functional characteristics of most rice GLR genes remain largely unknown. The phylogenetic analysis based on sequence similarity classified the 26 rice GLR genes into four major groups with groups II and III closely related

10.3389/fpls.2023.1236251

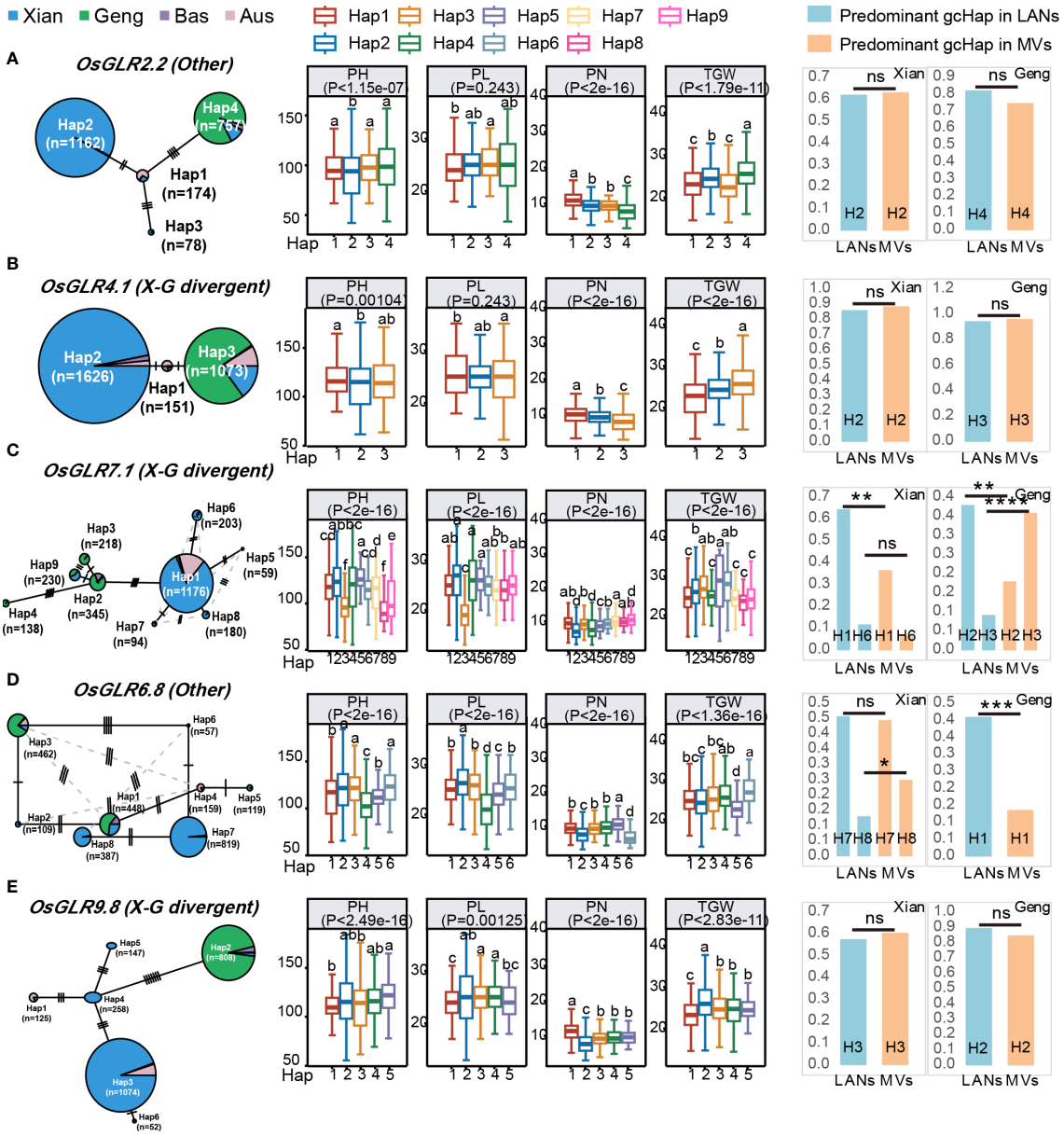


FIGURE 8
Haplotype networks of five GLR genes with CRSPR-9 mediated knocking-out and their associations with four agronomic traits in 3KRG. **(A)** Based on gcHap diversity and contribution to subspecific/population differentiation, the rice *OsGLR2.2* is classified as the other gene. **(B)** The *OsGLR4.1* is classified as a *X-G* divergent gene. **(C)** The *OsGLR7.1* is classified as a X-G divergent gene. **(D)** The rice *OsGLR6.8* is classified as the other gene. **(E)** The *OsGLR9.8* is classified as a *X-G* divergent gene. Within each haplotype network, two adjacent gcHaps are separated by mutational changes with hatching line indicating differences between the two most related haplotypes. The right side of each gene haplotype network corresponds to the phenotypic variation among the haplotypes. Boxplots are shown for the following four traits: plant height (PH), panicle length (PL), panicle number (PN) and 1000-grain weight (TGW). The *P* values under trait names indicate differences between the haplotypes assessed by two-way ANOVA, with different letters on the boxplots indicating statistically significant differences at $P < 0.05$ based on Duncan's multiple range test. The bar charts on the right show the differences in frequency of the predominant gcHaps between landraces (LANs) and modern varieties (MVs) in *Xian* and *Geng*. Chi-square tests were used to determine significant differences in the proportions of the same gcHap between the different populations with ****$P < 0.0001$, ***$P < 0.0001$, **$P < 0.01$, *$P < 0.05$, and n.s., not significant.

to maize GLR genes, while group IV containing eight GLR genes clustered on rice chromosome 6 was unique ones only present rice only (Figure 1). In this study, we demonstrated an integrated strategy by limited experimental efforts plus comprehensive data mining and analyses to understand the functionalities of the rice GLR genes, focusing on the phenotypic characterization and genetic diversity of rice GLR genes because of their importance for potential utilization in rice improvement. In this regard, our results revealed several important properties of the rice GLR genes that led us to several conclusions on the functionalities and evolutionary history of rice GLR genes.

Our first conclusion was that the expansion of the rice GLR genes into a large gene family during evolution had gone through repeated gene duplication events evidenced by the presence of two

10.3389/fpls.2023.1236251

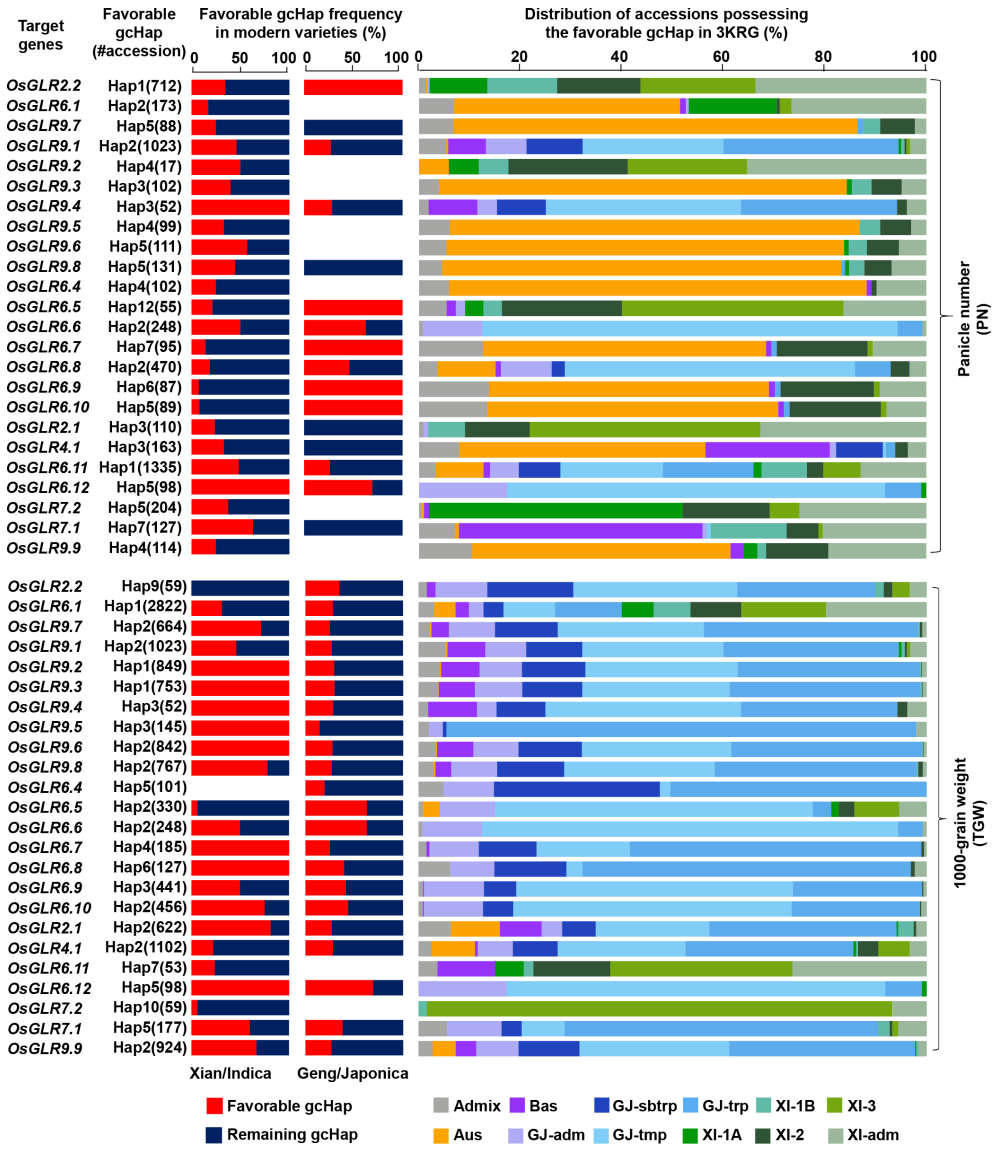


FIGURE 9
Frequencies of the favorable gcHaps of all rice GLR genes affecting important agronomic traits (PN and TGW) in *Xian/indica* (*XI*) and *Geng/japonica* (*GJ*) modern varieties and different rice subpopulations. The favorable gcHap of a gene is defined as the gcHap associated with the highest trait value. "#accession" indicates the number of accessions that possess the favorable gcHap. Five subpopulations of *XI* (*XI-1A*, *XI-1B*, *XI-2*, *XI-3*, and *XI-adm*) and four subpopulations of *GJ* (temperate *GJ*: *GJ-tmp*, subtropical *GJ*: *GJ-sbtrp*, tropical *GJ*: *GJ-trp*, and *GJ-adm*) were classified by Wang et al. (Wang et al., 2018).

large GLR gene clusters on rice chromosomes 9 and 6, and their greater gcHap diversity and rapid evolutionary rates (Table 1), and this expansion of the rice GLR genes had been accompanied with considerable functional differentiation, evidenced by their different tissue expression specificities (Figure 2), by their differentiated responses to different combinations of abiotic stresses and hormone treatments (Figure 3), and by their associations with different combinations of agronomic traits (Supplementary Figures 8-31).

Secondly, rice GLR genes may have played an important roles in rice sub-specific and population differentiation during evolution, evidenced by their very strong sub-specific and population differentiation at most rice GLR loci (Figure 6), by the great differentiated responses to day-length, the major factor determining rice subspecific adaptations to their environments (Figure 4), and by their differentiated responses to different abiotic stresses (Figure 3). In fact, the observed population differentiation at most GLR loci were far stronger than most of the 3,000+ cloned rice genes or genomewide SNPs analyzed previously (Wang et al., 2018; Zhang et al., 2021).

Thirdly, many rice GLR genes may have played an important role in rice improvement, suggested by the strong and large phenotypic effects of the five GLR gene knockout mutants on rice yield traits (Figure 4), by the significant shifts of the predominant gcHaps and diversity changes at many rice GLR loci (Figures 5–7, Table 1), and by their strong associations of the major gcHaps with

agronomic traits at many GLR loci (Figures 8–9; Supplementary data 8). This is not surprising since acting as signal receptors, GLR genes are expected to play important regulatory roles involved in many processes of rice growth and development and their responses to environmental cues. However, it remains to be elucidated in future how different rice GLR genes are involved in these important processes at the molecular level.

Fourthly, the observed diversity increases at some of the GLR loci during breeding were somehow unexpected since modern rice breeding is believed by many to cause reduced diversity in MVs. For example, the convergent phenotypic selection during breeding for similar phenotypes associated for high productivity in virtually all breeding programs, which include more compact plant type (erect leaves and smaller culm and leaf angles), large panicles (more), moderate sizes of TGW, high SR, more PN, and tolerance/resistance to the common abiotic/biotic stresses, etc. Also, the fact that only a very small portion of germplasm accessions maintained in genebanks worldwide have been utilized as parents in specific breeding programs would expected to result in severe genetic bottleneck and thus reduced genetic diversity at most gene loci. Indeed, we observed average losses of 39.3% (20.8%) of gcHaps (major gcHaps) at the 24 GLR loci during *Xian* breeding, which were accompanied with the emergences of 62.8 (18.4%) and new gcHaps per locus and 2.2 (28.6%) new major gcHaps/locus in MVs-*Xian* (Figure 7). The average losses of gcHaps (major gcHaps) at the 24 GLR loci reached were pronounced during *Geng* breeding, reaching 46.3% (58.3%), which were accompanied with an average emergences of 13.5 (17.7%) new gcHaps/locus and 3.0 (62.5%) new major gcHaps/locus during *Geng* breeding. In other words, the losses of gcHaps from the genetic bottleneck effects during modern breeding were more severe for MVs-*Geng* than MVs-*Xian*. This was expected since line hybridization is a common practice of modern breeding, which would result in greatly increased recombination as compared to the extremely low outcrossing rate in rice landraces. The much greater numbers of gcHap losses/gains observed in the two large GLR clusters on chromosomes 9 and 6 (Table 1) strongly suggest that more frequently occurred unequal crossing-over events in the large gene clusters would not only generate gene copy number variation but also new alleles of the duplicated GLR genes. More interestingly, significant reduced frequencies of the predominant gcHaps, losses of a few major gcHaps and the emergences of new major gcHaps in MVs-*Xian* and MVs-*Geng* from the rare ones in LANs-*Xian* and LANs-*Geng* at specific GLR loci would suggest cases that the predominant and/or major gcHaps favored by natural selection were against by the artificial selection during modern breeding. This also implies that many rare gcHaps might be of value in rice improvement. Taken together, modern breeding activities in the past had the greater impact on the GLR gene loci than the 3,000 + cloned genes analyzed previously (Zhang et al., 2021), implying again the greater importance of the GLR genes in rice improvement.

Finally, our results would suggest the potential values of the natural variation at most rice GLR loci for improving the productivity and tolerances to abiotic stresses. All rice GLR genes, except for *OsGLR6.2*, *OsGLR6.3*, *OsGLR6.1* and *OsGLR9.1*, are potential values to be utilized in rice trait improvement. Specifically, *OsGLR2.1*, *OsGLR2.2*, *OsGLR4.1*, *OsGLR6.12*, *OsGLR7.1* and *OsGLR9.8* could potentially be used for improving rice tolerances to abiotic stresses, while others can potentially be used for improving rice yield traits. However, challenges remain in the identification of 'favorable' alleles at specific GLR loci for improving specific yield traits because of the greater G x E interaction of most rice GLR genes on yield traits observed in this study. Thus, additional efforts are needed to obtain the required information in order to facilitate their application to improving rice productivity in future.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

YS and ZL designed the experiments. WZ, HL, FZ, XW, SR, SH, CheZ, FW, JL, YL, ChaZ, ML, and YS performed the experiments and data analyses. ZL, WZ, and YS wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Scientific Research Plan Major Projects of Anhui Province (grant number 2022AH040126), the Science and Technology Major Project of Anhui Province (grant number 2021d06050002), the Improved Varieties Joint Research (Rice) Project of Anhui Province (the 14th five-year plan) and the National Natural Science Foundation of China (grant number U21A20214).

Acknowledgments

We would like to extend special thanks to Professor Fan Zhang from the Institute of Crop Science, Chinese Academy of Agricultural Sciences for his invaluable assistance with the data analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1236251/full#supplementary-material

SUPPLEMENTARY DATA SHEET 1
Sequence information of plant and animal GLR genes.

SUPPLEMENTARY FIGURE 1
Gene structures of knockout mutants of five rice *OsGLR* genes. **(A-E)** The genes represented by **(A-E)** are *OsGLR2.2*, *OsGLR9.8*, *OsGLR6.8*, *OsGLR4.1* and *OsGLR7.1*.

SUPPLEMENTARY FIGURE 2
Classification of gene categories of rice *OsGLR* genes based on their gcHap diversity and contribution to subspecific/population differentiation, with conserved genes (E_H < 0.3 in 3KRG and I_{Nei} ≥ 0.8 between pairwise populations) and the *Xian-Geng* (*X-G*) divergent genes (E_H < 0.3 in both Xian and Geng and I_{Nei} ≤ 0.2 between *Xian* and *Geng*).

SUPPLEMENTARY FIGURE 3
Frequency shifts in the predominant gcHaps at 24 *OsGLR* genes between the landraces and modern varieties.

SUPPLEMENTARY FIGURE 4
Haplotype networks of five overexpressed *OsGLR* genes and their associations with four agronomic traits in 3KRG. Within each haplotype network, two adjacent gcHaps are separated by mutational changes with hatching line indicating differences between the two most related haplotypes. The right side of each gene haplotype network corresponds to the phenotypic variation among the haplotypes. Boxplots are shown for the following four traits: plant height (PH), panicle length (PL), panicle number (PN) and 1000-grain weight (TGW). The *P* values under trait names indicate differences between the haplotypes assessed by two-way ANOVA, with different letters on the boxplots indicating statistically significant differences at $P < 0.05$ based on Duncan's multiple range test. The bar charts on the right show the differences in frequency of the predominant gcHaps between landraces (LANs) and modern varieties (MVs) in *Xian* and *Geng*. Chi-square tests were used to determine significant differences in the proportions of the same gcHap between the different populations with ****$P < 0.0001$, **$P < 0.01$, *$P < 0.05$, and n.s., not significant.

SUPPLEMENTARY FIGURE 5
Haplotype networks of four GLR genes and their associations with four agronomic traits in 3KRG. Within each haplotype network, two adjacent gcHaps are separated by mutational changes with hatching line indicating differences between the two most related haplotypes. The right side of each gene haplotype network corresponds to the phenotypic variation among the haplotypes. Boxplots are shown for the following four traits: plant height (PH), panicle length (PL), panicle number (PN) and 1000-grain weight (TGW). The *P* values under trait names indicate differences between the haplotypes assessed by two-way ANOVA, with different letters on the boxplots indicating statistically significant differences at $P < 0.05$ based on Duncan's multiple range test. The bar charts on the right show the differences in frequency of the predominant gcHaps between landraces (LANs) and modern varieties (MVs) in *Xian* and *Geng*. Chi-square tests were used to determine significant differences in the proportions of the same gcHap between the different populations with ****$P < 0.0001$, **$P < 0.01$, *$P < 0.05$, and n.s., not significant.

SUPPLEMENTARY FIGURE 6
Haplotype networks of five *OsGLR* genes and their associations with four agronomic traits in 3KRG. Within each haplotype network, two adjacent gcHaps are separated by mutational changes with hatching line indicating differences between the two most related haplotypes. The right side of each gene haplotype network corresponds to the phenotypic variation among the haplotypes. Boxplots are shown for the following four traits: plant height (PH), panicle length (PL), panicle number (PN) and 1000-grain weight (TGW). The *P* values under trait names indicate differences between the haplotypes assessed by two-way ANOVA, with different letters on the boxplots indicating statistically significant differences at $P < 0.05$ based on Duncan's multiple range test. The bar charts on the right show the differences in frequency of the predominant gcHaps between landraces (LANs) and modern varieties (MVs) in *Xian* and *Geng*. Chi-square tests were used to determine significant differences in the proportions of the same gcHap between the different populations with ****$P < 0.0001$, **$P < 0.01$, *$P < 0.05$, and n.s., not significant.

SUPPLEMENTARY FIGURE 7
Haplotype networks of remaining GLR genes and their associations with four agronomic traits in 3KRG. Within each haplotype network, two adjacent gcHaps are separated by mutational changes with hatching line indicating differences between the two most related haplotypes. The right side of each gene haplotype network corresponds to the phenotypic variation among the haplotypes. Boxplots are shown for the following four traits: plant height (PH), panicle length (PL), panicle number (PN) and 1000-grain weight (TGW). The *P* values under trait names indicate differences between the haplotypes assessed by two-way ANOVA, with different letters on the boxplots indicating statistically significant differences at $P < 0.05$ based on Duncan's multiple range test. The bar charts on the right show the differences in frequency of the predominant gcHaps between landraces (LANs) and modern varieties (MVs) in *Xian* and *Geng*. Chi-square tests were used to determine significant differences in the proportions of the same gcHap between the different populations with ****$P < 0.0001$, **$P < 0.01$, *$P < 0.05$, and n.s., not significant.

SUPPLEMENTARY FIGURE 8
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR2.1*.

SUPPLEMENTARY FIGURE 9
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR2.2*.

SUPPLEMENTARY FIGURE 10
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR4.1*.

SUPPLEMENTARY FIGURE 11
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.1*.

SUPPLEMENTARY FIGURE 12
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.4*.

SUPPLEMENTARY FIGURE 13
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.5*.

SUPPLEMENTARY FIGURE 14
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.6*.

SUPPLEMENTARY FIGURE 15
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.7*.

SUPPLEMENTARY FIGURE 16
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.8*.

SUPPLEMENTARY FIGURE 17
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.9*.

SUPPLEMENTARY FIGURE 18
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.10*.

SUPPLEMENTARY FIGURE 19
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.11*.

SUPPLEMENTARY FIGURE 20
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR6.12*.

SUPPLEMENTARY FIGURE 21
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR7.1*.

SUPPLEMENTARY FIGURE 22
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR7.2*.

SUPPLEMENTARY FIGURE 23
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR9.1*.

SUPPLEMENTARY FIGURE 24
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR9.2*.

SUPPLEMENTARY FIGURE 25
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR9.3*.

SUPPLEMENTARY FIGURE 26
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR9.4*.

SUPPLEMENTARY FIGURE 27
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR9.5*.

SUPPLEMENTARY FIGURE 28
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR9.6*.

SUPPLEMENTARY FIGURE 29
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR9.7*.

SUPPLEMENTARY FIGURE 30
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR9.8*.

SUPPLEMENTARY FIGURE 31
Comparison and analysis of 15 agronomic traits among the predominant gcHap, unfavorable gcHap, and major gcHaps of *OsGLR9.9*.

SUPPLEMENTARY FIGURE 32
Frequencies of the favorable gcHaps of all rice *OsGLR* genes affecting important agronomic traits (PH and PL) in *Xian/indica* (*XI*) and *Geng/japonica* (*GJ*) modern varieties and different rice subpopulations. The favorable gcHap of a gene was defined as the gcHap associated with the highest trait value. "#accession" indicates the number of accessions that possess the favorable gcHap. Five subpopulations of *XI* (*XI-1A*, *XI-1B*, *XI-2*, *XI-3*, and *XI-adm*) and four subpopulations of *GJ* (temperate *GJ*: *GJ-tmp*, subtropical *GJ*: *GJ-sbtrp*, tropical *GJ*: *GJ-trp*, and *GJ-adm*) were classified by Wang et al. (Wang et al., 2018).

References

Cheng, Y., Tian, Q., and Zhang, W. H. (2016). Glutamate receptors are involved in mitigating effects of amino acids on seed germination of *Arabidopsis thaliana* under salt stress. *Environ. Exp. Bot.* 130, 68–78. doi: 10.1016/j.envexpbot.2016.05.004

Cheng, Y., Zhang, X., Sun, T., Tian, Q., and Zhang, W. H. (2018). Glutamate receptor homolog3.4 is involved in regulation of seed germination under salt stress in *Arabidopsis*. *Plant Cell Physiol.* 59 (5), 978–988. doi: 10.1093/pcp/pcy034

Green, M. N., Gangwar, S. P., Michard, E., Simon, A. A., Portes, M. T., Barbosa-Caro, J., et al. (2021). Structure of the *Arabidopsis thaliana* glutamate receptor-like channel GLR3.4. *Mol. Cell* 81 (15), 3216–3226.e3218. doi: 10.1016/j.molcel.2021.05.025

Kang, J., and Turano, F. J. (2003). The putative glutamate receptor 1.1 (AtGLR1.1) functions as a regulator of carbon and nitrogen metabolism in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 100 (11), 6872–6877. doi: 10.1073/pnas.1030961100

Kong, D., Ju, C., Parihar, A., Kim, S., Cho, D., and Kwak, J. M. (2015). Arabidopsis glutamate receptor homolog3.5 modulates cytosolic Ca2+ level to counteract effect of abscisic acid in seed germination. *Plant Physiol.* 167 (4), 1630–1642. doi: 10.1104/pp.114.251298

Lam, H. M., Chiu, J., Hsieh, M. H., Meisel, L., Oliveira, I. C., Shin, M., et al. (1998). Glutamate-receptor genes in plants. *Nature* 396 (6707), 125–126. doi: 10.1038/24066

Li, H., Jiang, X., Lv, X., Ahammed, G. J., Guo, Z., Qi, Z., et al. (2019). Tomato GLR3.3 and GLR3.5 mediate cold acclimation-induced chilling tolerance by regulating apoplastic H(2) O(2) production and redox homeostasis. *Plant Cell Environ.* 42 (12), 3326–3339. doi: 10.1111/pce.13623

Li, J., Zhu, S. H., Song, X. W., Shen, Y., Chen, H. M., Yu, J., et al. (2006). A rice glutamate receptor-like gene is critical for the division and survival of individual cells in the root apical meristem. *Plant Cell* 18 (2), 340–349. doi: 10.1105/tpc.105.037713

Liu, S., Zhang, X., Xiao, S., Ma, J., Shi, W., Qin, T., et al. (2021). A single-nucleotide mutation in a *GLUTAMATE RECEPTOR-LIKE* gene confers resistance to fusarium wilt in *Gossypium hirsutum*. *Adv. Sci. (Weinh)* 8 (7), 2002723. doi: 10.1002/advs.202002723

Lu, G. H., Wang, X. P., Liu, J. H., Yu, K., Gao, Y., Liu, H. Y., et al. (2014). Application of T-DNA activation tagging to identify *glutamate receptor-like* genes that enhance drought tolerance in plants. *Plant Cell Rep.* 33 (4), 617–631. doi: 10.1007/s00299-014-1586-7

Michard, E., Lima, P. T., Borges, F., Silva, A. C., Portes, M. T., Carvalho, J. E., et al. (2011). *Glutamate receptor-like* genes form Ca2+ channels in pollen tubes and are regulated by pistil D-serine. *Science* 332 (6028), 434–437. doi: 10.1126/science.1201101

Mousavi, S. A. R., Chauvin, A., Pascaud, F., Kellenberger, S., and Farmer, E. E. (2013). *GLUTAMATE RECEPTOR-LIKE* genes mediate leaf-to-leaf wound signalling. *Nature* 500 (7463), 422–426. doi: 10.1038/nature12478

Nei, M. (1972). Genetic distance between populations. *Am. Nat.* 106 (949), 283–292. doi: 10.1086/282771

Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26 (3), 419–420. doi: 10.1093/bioinformatics/btp696

Philippe, F., Verdu, I., Morere-Le Paven, M. C., Limami, A. M., and Planchet, E. (2019). Involvement of Medicago truncatula glutamate receptor-like channels in nitric oxide production under short-term water deficit stress. *J. Plant Physiol.* 236, 1–6. doi: 10.1016/j.jplph.2019.02.010

Project, r.g (2014). The 3,000 rice genomes project. *Gigascience* 3, 7. doi: 10.1186/2047-217X-3-7

Shao, Q., Gao, Q., Lhamo, D., Zhang, H., and Luan, S. (2020). Two glutamate- and pH-regulated Ca(2+) channels are required for systemic wound signaling in Arabidopsis. *Sci. Signal* 13 (640), eaba1453. doi: 10.1126/scisignal.aba1453

Sheldon, A. L. (1969). Equitability indices: dependence on the species count. *Ecology* 50 (3), 466–467. doi: 10.2307/1933900

Shi, Y., Zeng, W., Xu, M., Li, H., Zhang, F., Chen, Z., et al. (2022). Comprehensive Analysis of *Glutamate Receptor-like* Genes in Rice (Oryza sativa L.): Genome-Wide Identification, Characteristics, Evolution, Chromatin Accessibility, gcHap Diversity, Population Variation and Expression Analysis. *Curr. Issues Mol. Biol.* 44 (12), 6404–6427. doi: 10.3390/cimb44120437

Simon, A. A., Navarro-Retamal, C., and Feijó, J. A. (2023). Merging signaling with structure: functions and mechanisms of plant glutamate receptor ion channels. *Annu. Rev. Plant Biol* 74, 415–452. doi: 10.1146/annurev-arplant-070522-033255

Singh, S. K., Chien, C. T., and Chang, I. F. (2016). The Arabidopsis glutamate receptor-like gene *GLR3.6* controls root development by repressing the Kip-related protein gene KRP4. *J. Exp. Bot.* 67 (6), 1853–1869. doi: 10.1093/jxb/erv576

Templeton, A. R., Crandall, K. A., and Sing, C. F. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132 (2), 619–633. doi: 10.1093/genetics/132.2.619

Vincill, E. D., Clarin, A. E., Molenda, J. N., and Spalding, E. P. (2013). Interacting glutamate receptor-like proteins in Phloem regulate lateral root initiation in Arabidopsis. *Plant Cell* 25 (4), 1304–1313. doi: 10.1105/tpc.113.110668

Wang, P. H., Lee, C. E., Lin, Y. S., Lee, M. H., Chen, P. Y., Chang, H. C., et al. (2019). The glutamate receptor-like protein GLR3.7 interacts with 14-3-3ω and participates in salt stress response in *Arabidopsis thaliana*. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01169

Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., et al. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557 (7703), 43–49. doi: 10.1038/s41586-018-0063-9

Wudick, M. M., Portes, M. T., Michard, E., Rosas-Santiago, P., Lizzio, M. A., Nunes, C. O., et al. (2018). CORNICHON sorting and regulation of GLR channels underlie pollen tube Ca(2+) homeostasis. *Science* 360 (6388), 533–536. doi: 10.1126/science.aar6464

Xia, L., Zou, D., Sang, J., Xu, X., Yin, H., Li, M., et al. (2017). Rice Expression Database (RED): An integrated RNA-Seq-derived gene expression database for rice. *J. Genet. Genomics* 44 (5), 235–241. doi: 10.1016/j.jgg.2017.05.003

Xue, N., Zhan, C., Song, J., Li, Y., Zhang, J., Qi, J., et al. (2022). The glutamate receptor-like 3.3 and 3.6 mediate systemic resistance to insect herbivores in Arabidopsis. *J. Exp. Bot.* 73 (22), 7611–7627. doi: 10.1093/jxb/erac399

Yu, B., Wu, Q., Li, X., Zeng, R., Min, Q., and Huang, J. (2022). GLUTAMATE RECEPTOR-like gene *OsGLR3.4* is required for plant growth and systemic wound signaling in rice (*Oryza sativa*). *New Phytol.* 233 (3), 1238–1256. doi: 10.1111/nph.17859

Zeng, W., Shi, J., Qiu, C., Wang, Y., Rehman, S., Yu, S., et al. (2020). Identification of a genomic region controlling thermotolerance at flowering in maize using a combination of whole genomic re-sequencing and bulked segregant analysis. *Theor. Appl. Genet.* 133 (10), 2797–2810. doi: 10.1007/s00122-020-03632-x

Zhang, F., Wang, C., Li, M., Cui, Y., Shi, Y., Wu, Z., et al. (2021). The landscape of gene-CDS-haplotype diversity in rice: Properties, population organization, footprints of domestication and breeding, and implications for genetic improvement. *Mol. Plant* 14 (5), 787–804. doi: 10.1016/j.molp.2021.02.003

Zheng, Y., Luo, L., Wei, J., Chen, Q., Yang, Y., Hu, X., et al. (2021). Corrigendum to: The glutamate receptors AtGLR1.2 and AtGLR1.3 increase cold tolerance by regulating jasmonate signaling in *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* 566, 211–213. doi: 10.1016/j.bbrc.2021.06.061

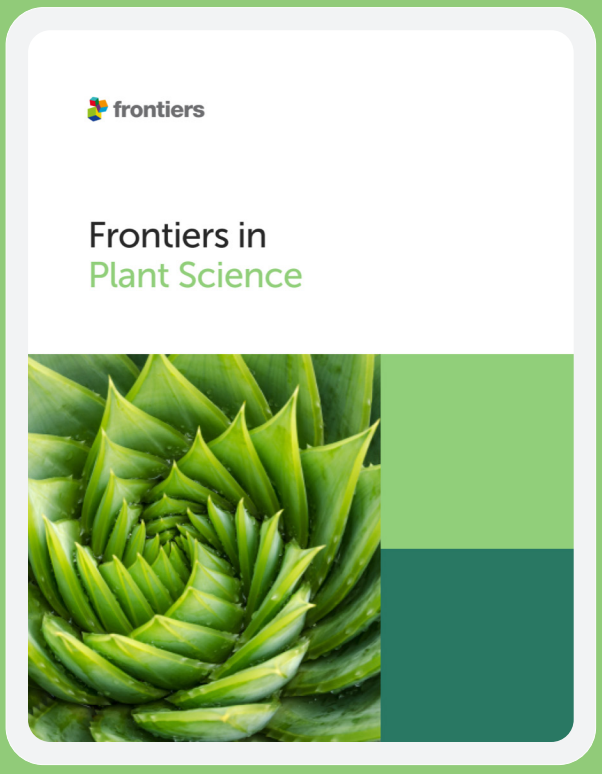
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact