REPRODUCTIVE BARRIERS AND GENE INTROGRESSION IN RICE SPECIES, VOLUME II

EDITED BY: Yohei Koide, Kazuki Matsubara, Dayun Tao and Kenneth Lee McNally

PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-840-0
DOI 10.3389/978-2-88976-840-0

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

REPRODUCTIVE BARRIERS AND GENE INTROGRESSION IN RICE SPECIES, VOLUME II

Topic Editors:
Yohei Koide, Hokkaido University, Japan
Kazuki Matsubara, Institute of Crop Science (NARO), Japan
Dayun Tao, Yunnan Academy of Agricultural Sciences, China
Kenneth Lee McNally, International Rice Research Institute (IRRI), Philippines

Citation: Koide, Y., Matsubara, K., Tao, D., McNally, K. L., eds. (2022). Reproductive Barriers and Gene Introgression in Rice Species, Volume II.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-840-0

Table of Contents

05 ***Editorial: Reproductive Barriers and Gene Introgression in Rice Species, Volume II***
Yohei Koide, Kazuki Matsubara, Dayun Tao and Kenneth L. McNally

08 ***Use of RNAi With* OsMYB76R *as a Reporter for Candidate Genes Can Efficiently Create and Verify Gametophytic Male Sterility in Rice***
Yun Chen, Wenping Zhu, Shudan Shi, Lina Wu, Shuanglin Du, Liangshen Jin, Kuan Yang, Wenjia Zhao, Jiaxin Yang, Longbiao Guo, Zhongwei Wang and Yi Zhang

16 ***Loss of OsEAF6, a Subunit of the Histone Acetyltransferase Complex, Causes Hybrid Breakdown in Intersubspecific Rice Crosses***
Takahiko Kubo, Atsushi Yoshimura and Nori Kurata

27 ***Mapping of QTLs for Yield Traits Using $F_{2:3:4}$ Populations Derived From Two Alien Introgression Lines Reveals* qTGW8.1 *as a Consistent QTL for Grain Weight From* Oryza nivara**
Kavitha Beerelli, Divya Balakrishnan, Krishnam Raju Addanki, Malathi Surapaneni, Venkateswara Rao Yadavalli and Sarla Neelamraju

44 ***Reticulate Evolution in AA-Genome Wild Rice in Australia***
Sharmin Hasan, Agnelo Furtado and Robert Henry

57 ***The Next Generation of Rice: Inter-Subspecific* Indica-Japonica *Hybrid Rice***
Guiquan Zhang

64 ***Fine Mapping of Two Interacting Loci for Transmission Ratio Distortion in Rice (*Oryza sativa *L.)***
Chaopu Zhang, Jilin Wang, Xiongfeng Xiao, Dianwen Wang, Zhiyang Yuan, Xiaodan Zhang, Wenqiang Sun and Sibin Yu

73 ***A Genetic Network Underlying Rhizome Development in* Oryza longistaminata**
Wenfei Li, Shilai Zhang, Guangfu Huang, Liyu Huang, Jing Zhang, Zheng Li and Fengyi Hu

85 ***Introgression Lines: Valuable Resources for Functional Genomics Research and Breeding in Rice (*Oryza sativa *L.)***
Bo Zhang, Ling Ma, Bi Wu, Yongzhong Xing and Xianjin Qiu

101 ***Development of Wide-Compatible* Indica *Lines by Pyramiding Multiple Neutral Alleles of* Indica–Japonica *Hybrid Sterility Loci***
Jie Guo, Yun Li, Liang Xiong, Tingxian Yan, Jinsong Zou, Ziju Dai, Guang Tang, Kangli Sun, Xin Luan, Weifeng Yang, Quanya Tan, Haitao Zhu, Ruizhen Zeng, Shaokui Wang and Guiquan Zhang

111 ***Phenotypic Variation and the Impact of Admixture in the* Oryza rufipogon *Species Complex (*ORSC*)***
Georgia C. Eizenga, HyunJung Kim, Janelle K. H. Jung, Anthony J. Greenberg, Jeremy D. Edwards, Maria Elizabeth B. Naredo, Maria Celeste N. Banaticla-Hilario, Sandra E. Harrington, Yuxin Shi, Jennifer A. Kimball, Lisa A. Harper, Kenneth L. McNally and Susan R. McCouch

126 ***Understanding the Nature of Hybrid Sterility and Divergence of Asian Cultivated Rice***

Yu Zhang, Jie Wang, Qiuhong Pu, Ying Yang, Yonggang Lv, Jiawu Zhou, Jing Li, Xianneng Deng, Min Wang and Dayun Tao

139 ***A Novel Combination of Genes Causing Temperature-Sensitive Hybrid Weakness in Rice***

Than Kutay Soe, Mai Kunieda, Hidehiko Sunohara, Yoshiaki Inukai, Vincent Pamugas Reyes, Shunsaku Nishiuchi and Kazuyuki Doi

153 ***Influence of Gender Bias on Distribution of Hybrid Sterility in Rice***

Zin Mar Myint and Yohei Koide

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 22 July 2022
DOI 10.3389/fpls.2022.974613

OPEN ACCESS

EDITED AND REVIEWED BY
Inaki Hormaza,
La Mayora Experimental Station (CSIC), Spain

*CORRESPONDENCE
Yohei Koide
ykoide@abs.agr.hokudai.ac.jp
Kazuki Matsubara
vague@affrc.go.jp
Dayun Tao
taody12@aliyun.com
Kenneth L. McNally
K.McNally@irri.org

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

RECEIVED 21 June 2022
ACCEPTED 04 July 2022
PUBLISHED 22 July 2022

CITATION
Koide Y, Matsubara K, Tao D and McNally KL (2022) Editorial: Reproductive Barriers and Gene Introgression in Rice Species, Volume II. *Front. Plant Sci.* 13:974613. doi: 10.3389/fpls.2022.974613

Editorial: Reproductive Barriers and Gene Introgression in Rice Species, Volume II

Yohei Koide[1]*, Kazuki Matsubara[2]*, Dayun Tao[3]* and Kenneth L. McNally[4]*

[1]Laboratory of Plant Breeding, Research Faculty of Agriculture, Hokkaido University, Sapporo, Japan, [2]Institute of Crop Science, National Agriculture and Food Research Organization, Tsukuba, Japan, [3]Yunnan Key Laboratory for Rice Genetic Improvement, Food Crops Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, China, [4]T. T. Chang Genetic Resources Center, International Rice Research Institute, Los Baños, Philippines

KEYWORDS

breeding, evolutionary biology, genetic resource, genome variation, hybridization, introgression, reproductive barriers

Editorial on the Research Topic
Reproductive Barriers and Gene Introgression in Rice Species, Volume II

Distant hybridization and introgression undoubtedly play a crucial role in Asian rice (*Oryza sativa* L.) domestication and diversification and will continue to play an important role as humans face to climate change challenges. The limited genetic variation within current rice varieties necessitates the use of various approaches to improve production under increasingly harsher and more frequent abiotic and biotic stresses resulting from climate change. The wild relatives of Asian cultivated rice represent valuable reservoirs of variation for both traditional breeding and advanced technology breeding such as *de novo* domestication, redomestication, and gene editing improvement.

To achieve the complex task mentioned above, the use of distant relatives of rice as donors of genetic diversity has been attempted to improve resistance together with productivity in rice varieties. However, reproductive barriers, which appear in prezygotic and/or postzygotic phases, are often observed in interspecific or intersubspecific crosses and prevent the introgression of useful genes. Therefore, reproductive barriers and gene introgression are major concerns in rice breeding, and are also of concern in rice evolutionary and developmental genetics.

This Research Topic is the second of a series meant to gather current knowledge in this field and share it with the scientific community to accelerate rice genetic improvement.

In this Research Topic, two articles described the phenotypic and genotypic nature of wild rice species. Eizenga et al. characterized the phenotypic variation of the *Oryza rufipogon* species complex (ORSC), the wild progenitor of Asian rice (*O. sativa* L.).

The ORSC comprises perennial, annual, and intermediate forms historically designated as *O. rufipogon*, *O. nivara*, and *O. sativa* f. *spontanea*, respectively, based on non-standardized morphological, geographical, and/or ecological-based species definitions and boundaries. A collection of 240 diverse ORSC accessions, previously characterized by genotyping-by-sequencing (113,739 SNPs), was phenotyped for a total of 57 traits in three different locations. Combined with information from morphology-based and genetic identity-based species definition analyses, one phenotypic group contained predominantly *O. rufipogon* accessions characterized as perennial and largely out-crossing and one contains predominantly *O. nivara* accessions characterized as annual and largely self-crossing. The third group was identified as *Oryza* spp. and admixture levels of *O. sativa* accounted for more than 50% of the genome. Meanwhile, Hasan et al. discuss the issue of the two distinct taxa found growing together in northern Australia, *O. meridionalis* (including annual and perennial forms) and *O. rufipogon*-like taxa shown to have a chloroplast genome sequence closer to *O. meridionalis* than to *O. rufipogon* from Asia. A comparison of chloroplast and nuclear genome sequences indicated hybridization between these taxa. Individuals with intermediate morphology had high nuclear genome heterozygosity consistent with a hybrid origin. An examination of specific genes suggested that some wild plants were early generation hybrids.

For reproductive barriers, two articles reviewed the latest issues regarding hybrid sterility in rice. Zhang Y. et al. reviewed systematically the relationship between hybrid sterility and divergence of Asian cultivated rice, finding that more than 40 conserved and specific loci have been shown responsible for hybrid sterility of subgroup crosses. Most studies have focused on the sterility barriers between indica and japonica crosses, ignoring hybrid sterility among other subgroup crosses, leading to neither a systematic understanding of hybrid sterility and subgroup divergence nor an effective understanding of how to use the strong heterosis between subgroups in Asian cultivated rice. Future studies will aim to create a blueprint to identify intraspecific hybrid sterility loci for (1) overcoming the intraspecific hybrid sterility according to the parent subgroup type identification, (2) allowing the use of heterosis among subgroups, and (3) unlocking the relationships between hybrid sterility and Asian cultivated rice divergence. Myint and Koide reviewed the distribution of hybrid sterility genes, from the viewpoint of transmission ratio distortion in male (*m*TRD), female (*f* TRD), or *si*TRD (sex-independent transmission ratio distortion). Among 49 hybrid sterility loci surveyed, the number of loci for *m*TRD is the largest (33), in contrast to that of *f* TRD (10) and *si*TRD (7). Additionally, the sterility based on gamete specificity is distributed disproportionately between interspecific and intraspecific hybrids. This bias in the frequency of sex-specificity in the TRD system might reflect different evolutionary pressures acting on the *Oryza* system, suggesting that they are not driven only by mutation and genetic drift.

Identification of genes for reproductive barriers is also ongoing. Soe et al. reported their research progress regarding rice hybrid weakness, which was observed in F_1 hybrids. Genetic analysis for low temperature-dependent hybrid weakness was conducted in a rice F_2 population derived from Taichung 65 (T65, *japonica*) and Lijiang-Xin-Tuan-Heigu (LTH, *japonica*). They observed that growth at 24°C enhanced hybrid weakness, whereas recovery occurred at 34°C. Two major QTLs were detected on chromosome 1, named *hybrid weakness j 1* (*hwj1*), and on chromosome 11, termed *hybrid weakness j 2* (*hwj2*). Further genotyping indicated that the hybrid weakness was due to an incompatible interaction between the T65 allele of *hwj1* and the LTH allele of *hwj2*.

Kubo et al., using a map-based cloning strategy, found that *HWE1* and *HWE2* encode the Esa1-associated factor 6 (EAF6) protein, a component of histone acetyltransferase complexes. The indica *hwe1* and japonica *hwe2* alleles lacked functional EAF6, demonstrating that double recessive homozygote causes hybrid breakdown in rice. These findings suggest that EAF6 plays a pivotal role in transcriptional regulation of essential genes during the vegetative and reproductive development of rice.

Zhang C. et al. undertook fine mapping and characterization of two loci (*TRD4.1* and *TRD4.2*) for transmission ratio distortion (TRD) using large F_2 segregating populations. The two loci exhibited a preferential transmission of the ZS97 alleles in the derived progeny. Reciprocal crossing experiments using near-isogenic lines harboring three different alleles at *TRD4.1* suggested that male gametic selection occurred. Moreover, the transmission bias of *TRD4.2* was diminished in heterozygotes when they carried homozygous $TRD4.1^{ZS97}$. These findings broaden the understanding of the genetic mechanisms of TRD and offer an approach to overcome the barrier of gene flow between the groups of varietal types in rice.

Chen et al. reported their preliminary results using RNAi and *OsMYB76R* as a reporter to efficiently create and verify gametophytic male sterility in rice. Guo et al. reported their research on the development of wide-compatible indica lines by pyramiding multiple neutral alleles of indica–japonica hybrid sterility loci. The results showed that wide-compatible indica lines (WCILs) by pyramiding multiple neutral (n) alleles of five loci showed wide compatibility with both indica and japonica rice varieties. Therefore, the WCILs can be used to develop intraspecific indica–japonica hybrid rice with normal fertility.

For gene introgression from the wild relatives of Asian cultivated rice, Li et al. reported preliminary results of the genetic network underlying rhizome development in *O.*

longistaminata. Beerelli et al. mapped *qTGW8.1* to a 2.6 Mb region in all three generations with PV 6.1 to 9.8%. This stable and consistent *qTGW8.1* allele from *O. nivara* can be finely mapped for the identification of causal genes for higher grain weight.

Zhang B. et al. reviewed recent progress on the development of introgression lines (ILs). They classified rice improvement into five generations: semi-dwarf rice, intra-subspecific hybrid rice, inter-subspecific introgression rice, and inter-subspecific indica-japonica hybrid rice. They found that indica-japonica hybrid sterility was mainly controlled by six loci. The indica-japonica hybrid sterility can be overcome by developing indica-compatible japonica lines (ICJLs) or WCILs using genes at the six hybrid sterility loci. With the understanding of the genetic and molecular basis of indica-japonica hybrid sterility and the development of molecular breeding technology, the development of indica-japonica hybrid rice has become possible.

The fine pieces of research collected in this topic facilitated a comprehensive understanding of issues in the introgression of traits of interest from the wild genetic resources to breeding programs in rice. We believe that this platform for enhancing exchange and promoting development has merit to be continued with other aspects such as gametophytes gene and non-random recombination of independent genes.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Acknowledgments

The authors wish to acknowledge the contributors to this Research Topic.

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

frontiers in Plant Science

ORIGINAL RESEARCH
published: 06 September 2021
doi: 10.3389/fpls.2021.728193

Use of RNAi With *OsMYB76R* as a Reporter for Candidate Genes Can Efficiently Create and Verify Gametophytic Male Sterility in Rice

Yun Chen[1†], Wenping Zhu[1†], Shudan Shi[1†], Lina Wu[2], Shuanglin Du[1], Liangshen Jin[1], Kuan Yang[1], Wenjia Zhao[1], Jiaxin Yang[1], Longbiao Guo[3], Zhongwei Wang[4]* and Yi Zhang[1]**

[1] State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Research Center for Perennial Rice Engineering and Technology in Yunnan, School of Agriculture, Yunnan University, Kunming, China, [2] College of Agronomy and Biotechnology, Southwest University, Chongqing, China, [3] State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou, China, [4] Biotechnology Research Center, Chongqing Academy of Agricultural Sciences, Chongqing, China

OPEN ACCESS

Edited by:
Kazuki Matsubara,
Institute of Crop Science (NARO), Japan

Reviewed by:
Tomohiko Kazama,
Kyushu University, Japan
Junxing Lu,
Chongqing Normal University, China

***Correspondence:**
Longbiao Guo
guolongbiao@caas.cn
Zhongwei Wang
shengran4@126.com
Yi Zhang
zhangyi6116@ynu.edu.cn

[†] These authors have contributed equally to this work

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

Received: 21 June 2021
Accepted: 13 August 2021
Published: 06 September 2021

Citation:
Chen Y, Zhu W, Shi S, Wu L, Du S, Jin L, Yang K, Zhao W, Yang J, Guo L, Wang Z and Zhang Y (2021) Use of RNAi With OsMYB76R as a Reporter for Candidate Genes Can Efficiently Create and Verify Gametophytic Male Sterility in Rice. Front. Plant Sci. 12:728193. doi: 10.3389/fpls.2021.728193

Gametophytic male sterility (GMS) plays an important role in the study of pollen development and seed propagation of recessive nuclear male sterile lines insensitive to the environmental conditions in hybrid rice breeding. Since the inherent phenotypic and genetic characteristics of GMS, it is very difficult to find and identify the GMS mutants. However, due to the abundance of gene transcription data, a large number of pollen-specific genes have been found, and most of them may be associated with GMS. To promote the study of these genes in pollen development and heterosis utilization, in this study, an easy and efficient method of creating and identifying GMS was established using RNAi and *OsMYB76R* as a reporter. First, the *OsC1/OsMYB76* gene involved in anthocyanin synthesis was modified, and we have validated that the modified *OsMYB76R* is workable as the same as the pre-modified *OsMYB76* gene. Then, the ascorbic acid oxidase gene *OsPTD1* was downregulated using RNAi, driven by its own promoter that resulted in abnormal pollen tube growth. Finally, the RNAi elements were linked with *OsMYB76R* and transformed into an *osmyb76* mutant, and the distortion of purple color segregation was found in T_1 and F_1 generations. This indicates that the *OsPTD1* GMS was prepared successfully. Compared to current methods, there are several advantages to this method. First, time is saved in material preparation, as one generation less needs to be compared than in the conventional method, and mutation screening can be avoided. In addition, for identification, the cost is lower; PCR, electrophoresis, and other processes are not needed; and no expensive chemicals or instruments are required. Finally, the results are more accurate, with much lower background effects, and no damage to the plant. The result is an easy, efficient, low-cost, and accurate method of preparing and identifying GMS genes.

Keywords: *Oryza sativa*, gametophytic male sterility, *OsMYB76R*, RNAi, GMS preparation and identification

Abbreviations: GMS, gametophytic male sterility; SMS, sporophytic male sterility; MYB, myeloblastosis; AAO, ascorbic acid oxidase.

INTRODUCTION

Plant male sterility plays an important role in the study of the development of the pollen mechanism and heterosis utilization. This includes sporophytic male sterility (SMS) and gametophytic male sterility (GMS). These two types are quite different in their heredity and phenotypes. SMS is caused by the sterility of pollen grains, as determined by sporophyte genotypes. In addition, the SMS gene can be transmitted between generations by either female gametes or male gametes in heterozygotes. All pollens in an SMS anther are sterile. In the field, the SMS anthers are usually abnormal in shape, color, and size, and the sterile plant has few seed settings. It is easy to find and identify homozygous SMS mutants. Due to the accessibility of SMS mutants, many studies have been carried out on the mechanism of pollen development and SMS using abundant SMS mutants (Ma, 2005; Wilson and Zhang, 2009; Ariizumi and Toriyama, 2011; Shi et al., 2015). By contrast, GMS is caused by the sterility of the pollen grains, as determined by gametophyte genotypes. In addition, the GMS gene can only be transmitted through the female gamete in heterozygotes, which leads to a very low probability of a homozygous mutant. Half of the pollen grains in a heterozygous anther are fertile, and the other half are sterile. There is no significant difference in appearance between the anthers of mutants and wild types in field observations, and the seed setting radio for the GMS mutant is normal. It is very difficult to locate and identify GMS mutants directly (Han et al., 2006, 2011; Boavida et al., 2009; Chen et al., 2012; Yang et al., 2012). For these reasons, compared with SMS, the location and identification of GMS are more difficult than for SMS. In officially published reports, GMS mutants have usually been created through spontaneous mutation, induced mutation, T-DNA insertion mutation, and so on. After mutation, the GMS candidates were verified in their crossing or self-crossing generations using PCR or resistance to screening agents, or their traits for detecting the segregation distortion of mutated genes or selectable markers or reporters (SariGorla et al., 1996; Feldmann et al., 1997; Howden et al., 1998; Boavida et al., 2009; Yang et al., 2009; Moon et al., 2013; Liu et al., 2016). These strategies, however, require special equipment, expensive chemicals, or invasive treatment, which entails time-consuming, low-efficiency, low-accuracy, high-cost, sample-damaging practices. Further, fewer GMS mutants are produced by these means, and the study of GMS has fallen behind the SMS research.

Pollen development is regulated not only through sporophytic gene expression but also through a large number of pollen genes (Ma, 2005). The genes associated with GMS are usually expressed in microspores and pollen at a late stage of anther development. Due to the development of the transcriptome and other detection methods for gene expression, an increasing number of pollen-specific genes have been identified (Ma, 2005; Suwabe et al., 2008; Wang et al., 2020). It is speculated that most of these genes are related to gametophyte development, which provides us with an opportunity to understand the mechanism of pollen development and GMS. To identify GMS-related genes from all the pollen-specific genes, it is necessary to develop a simple, rapid, economical, and accurate method for the preparation and identification of GMS.

Anthocyanin pigmentation is visible to the naked eye. The rice coleoptile purple line is a clear purple line that appears on each side of the coleoptile after the seed germination. Due to its early performance stage (budding stage) and stability, it is suitable to use as reporter traits (Xi, 1997; Zhang et al., 2004). *OsC1*/*OsMYB76* is a key gene encoding MYB transcription factor for anthocyanin biosynthesis in almost all organs, such as coleoptile in rice (Reddy, 1998; Zhang et al., 2004; Zhao et al., 2016; Sun et al., 2018; Zheng et al., 2019; Hu et al., 2020). Hence, *OsC1*/*OsMYB76* is an ideal reporter for the indication of transgenic events.

In view of the low efficiency and high cost of the preparation and identification of GMS in commonly used methods, this study aimed to design and validate a new method of creating and verifying GMS using RNAi technology and *OsMYB76R* (the revised *OsMYB76*) as a reporter gene. The GMS-related gene *OsPTD1* was identified in rice using our new method. We provide an easy, efficient, and low-cost method of creating and identifying the gene for GMS.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

The *Japonica* rice line Zhonghua 11 was kindly provided by Dr. Li Ping at Shanghai University. The *Indica* line R25, with multiple organs displaying purple color, is a self-breeding restorer, and YR25, without purple color, is an *osmyb76* mutant from R25. Zhongjiu B, another *osmyb76* mutant without purple color in all organs, is a breeding material that features a shorter growth period than Zhonghua 11, and it was kindly provided by the China National Rice Research Institute. Zhongjiu B-*osabcg15* is nuclear male sterility material, an *osabcg15* near-isogenic line of Zhongjiu B, derived from crosses and backcrosses between Zhongjiu B (the recurrent parent) and *osabcg15* (Wu et al., 2014). All rice plants used in this study were grown in the fields in Jinghong, Yunnan, China.

Vector Construction

OsMYB76R was synthesized by Shanghai Generay Biotech Co., Ltd. (Shanghai, China), after eliminating the restriction sites of *OsMYB76*, and its code was optimized according to the rice preference. The target fragment was amplified using the synthesized *OsMYB76R* as a template. The GUS gene in the pCAMBIA1301 vector was substituted with the PCR product of *OsMYB76R* using a homologous recombination strategy described in the user manual of the In-Fusion HD Cloning Kit (Takara, Otsu, Japan), which generated the *pP35S-OsMYB76R*-OE vector. This vector was transformed into the Zhongjiu B calli using an agrobacterium-mediated method. The sequences of *OsMYB76R* and primers 76F/R are shown in the **Supplementary Material 3**.

For the RNAi of the *OsPTD1* gene, all of the RNAi elements, such as the promoter of *POsPTD1* or *P35S*, target sequence (left arm), stem-loop structure (first intron of *OsMYB76*),

complementary sequence of target (right arm), and Tnos, were synthesized as one fusion sequence by Shanghai Generay Biotech Co., Ltd. The whole RNAi sequence was loaded between the restriction sites *Sal*I and *Sbf*I of pCAMBIA1301, which generated the vectors *pPOsPTD1-OsPTD1*-RNAi and *pP35S-OsPTD1-RNAi_P35S-OsC1*. These two vectors were separately transformed into Zhonghua 11.

For the preparation and identification of *OsPTD1* GMS, the synthesized RNAi sequences were inserted into *pP35S-OsMYB76R*-OE using *Sal*I and *Sbf*I, which generated *pPOsPTD1-OsPTD1_P35S-OsMYB76R* and *pP35S-OsPTD1_P35S-OsMYB76R*. These were transformed into Zhongjiu B.

Pollen Germination *in vitro*

The germination medium was composed of 18% sucrose, 5% potato starch, and 0.005% H_3BO_3. The solid medium was prepared in advance following these steps: evenly mixing the ingredients, heating and melting them in a microwave oven, cooling, and evenly spreading them on the slide. Immediately after the floret opened, pollen grains were shed onto the solid germination medium. After 30 min incubation at 28–30°C, the slides were observed with a Nikon (Tokyo, Japan) SMZ1500 stereoscope and photographed with a Nikon DS-5Mc digital camera.

Coleoptile Purple Color Observation and Inheritance Analyses

The transgenic positive plants with anthocyanin color were self-crossed to obtain the T_1 generations. The same positive plants were used as male parents to pollinate the male sterile of Zhongjiu B-*osabcg15* to obtain F_1 generations. After being soaked and pregerminated, the seeds were evenly planted in a mud tray, covered with plastic cling to keep them wet, and placed under natural light. The purple color of the coleoptile was investigated after 2–3 days of growth. The members of the segregated groups purple and no-purple were counted and calculated, and a chi-square test was conducted to establish whether the observed values were consistent with the theoretical values.

RESULTS

OsMYB76R, Revised From *OsC1/OsMYB76*, Is Workable When Replacing the *GUS* Gene as a Reporter in Rice

The *OsC1*/*OsMYB76* gene is a functional gene for anthocyanin biosynthesis in rice. In the previous studies, the CDS (coding sequence) of the *OsC1*/*OsMYB76* gene was used to replace the *GUS* gene in the vector pCAMBIA1301. The function of the *OsC1*/*OsMYB76* gene was verified with a complement test (data not shown). Transgenic positive plants showed a purple color in many organs, such as the stem base of the seedling, the leaf sheath, the auricle, the stigma apiculus, and the coleoptile, which shows that it can be used as a reporter gene to indicate positive transgenic events. The CDS of *OsMYB76* contains common restriction sites, such as *Pst*I, *Sac*I, and *Sal*I, but if *OsMYB76* is directly used as a reporter, it must influence the freedom of restriction site selection when the target gene is loaded. To remove these restriction sites, we optimized and modified the gene with artificial synthesis and obtained *OsMYB76R*, after that we replaced the *GUS* gene in pCAMBIA1301 with homologous recombination, which resulted in the functional verification vector *pP35S-OsMYB76R*-OE (**Figure 1A**). The plasmid was transformed into Zhongjiu B without any purple in any organs, and the results showed that the stem base of the seedling, the leaf sheath, the auricle, the stigma and apiculus, and the coleoptile of the transgenic positive plants were all purple in color (**Figures 1B–E**). The distribution of purple coloration of the coleoptile of the transgenic-positive plants was wider and easier to observe than that of most wild-type lines, such as R25 (an *Indica* restorer line) (**Figures 1E,F**). The results showed that the revised *OsMYB76R* gene has the equivalent of a genetic function for *OsMYB76* in rice anthocyanin biosynthesis, and *OsMYB76R* is suitable for reporting transformants.

RNA Interference of *OsPTD1* Damages Pollen Tube Development

The ascorbate acid oxidase (AAO) gene *NPT303* plays an important role in the growth of the tobacco pollen tube, and the silencing of *NPT303* results in the abnormal growth of the pollen tube (de Groot et al., 2004). There are more than 10 AAO genes in the rice genome,[1] and most of them have an unclear function. Among these AAO genes, *Os05g0485800* encodes a protein that is highly homologous to NPT303, may be related to pollen tube growth, and is named pollen tube development gene 1 (*OsPTD1*). To determine whether it plays a role in pollen germination and growth, two vectors for *OsPTD1* RNA interference were constructed, promoted by the constitutive 35S CaMV promoter (*P35S*) and the specific promoter of *POsPTD1* (**Figures 2A,B**), and were transferred into Zhonghua 11. The T_0 pollen germination assay showed that all the pollen from the RNAi promoted by *P35S* and the wild type could be germinated and grown normally, but only about half of the pollen from the RNAi promoted by *POsPTD1* germinated and grew normally. The germination of other pollen grains was abnormal, with most of the abnormally germinated pollen tubes rupturing without growing (**Figures 2C–E**). These results show that the RNAi under the *OsPTD1* promoter broke the growth of the pollen tube, and this may result in GMS.

Construction and Verification of a System for GMS Preparation and Identification

To confirm *OsPTD1* as a gene related to GMS, two vectors for the preparation and identification of GMS were designed and constructed, based on the *OsPTD1* gene. The *P35S*- and *POsPTD1*-promoted *OsPTD1* RNAi elements were linked with *OsMYB76R* that generated the RNAi vectors of

[1] http://www.gramene.org/

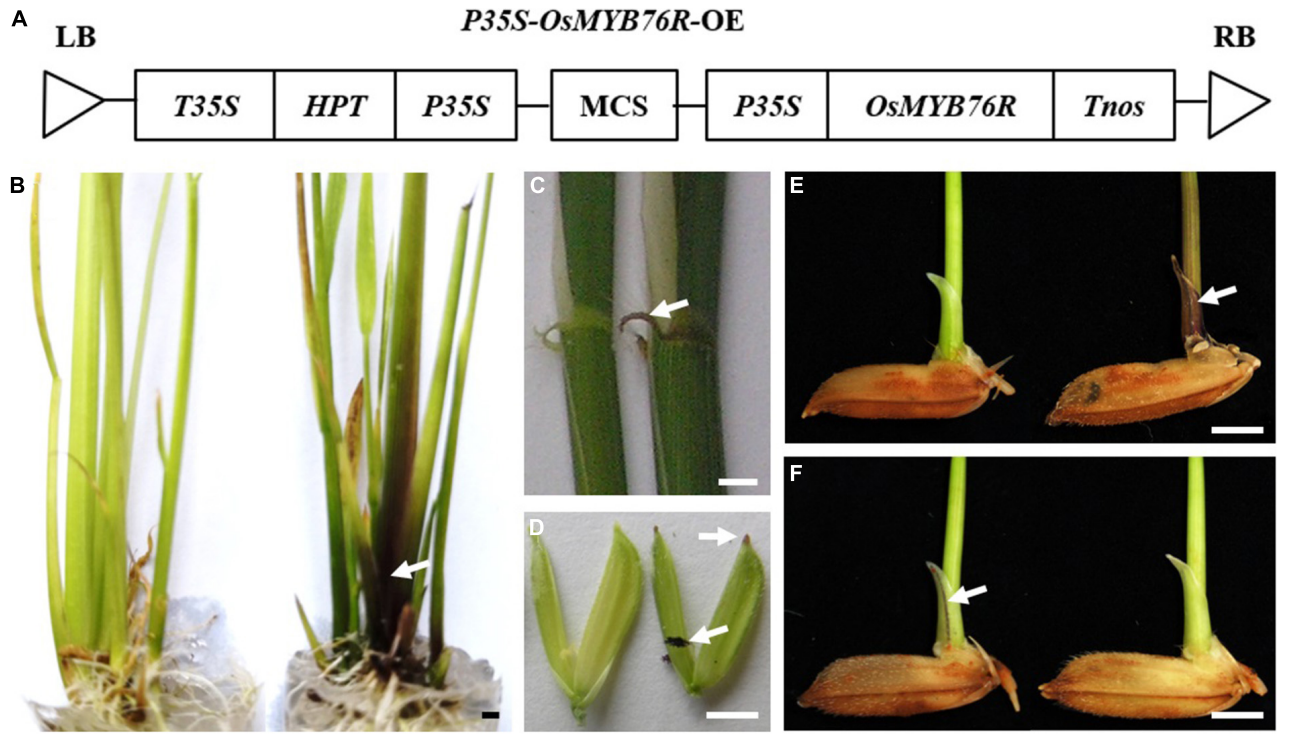

FIGURE 1 | Phenotype of the *OsMYB76R* gene. **(A)** The structure of the *pP35S-OsMYB76R*-OE vector. The left sides of panels **(B–E)** refer to the color of different organs in Zhongjiu B, and the right sides refer to the color of different organs in a plant that is positive for *OsMYB76R* overexpression. The traits in panels **(B–D)** are from T_0 plants. The trait in panel **(E)** is from the T_1 plant. The left side of the panel **(F)** is purple wild type R25, and the right side of the panel **(F)** shows *osmyb76* mutant YR25. The arrows point to the purple tissue. Scale bars, 0.2 cm.

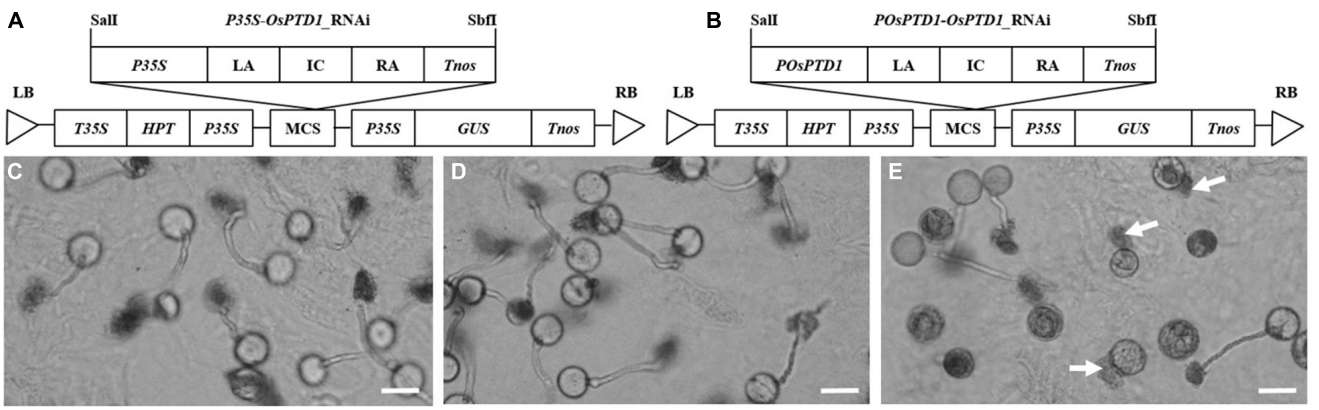

FIGURE 2 | RNAi pollen germination *in vitro*. **(A)** The structure of *pP35S-OsPTD1*-RNAi. **(B)** The structure of *pPOsPTD1-OsPTD1*-RNAi. **(C)** The germination of wild pollen. **(D)** The germination of the *P35S*-promoted RNAi pollen. **(E)** The germination of *POsPTD1*-promoted RNAi pollen; arrows indicate the rupture of RNAi pollen. All images were created 30 min after germination. Scale bars, 20 μm.

P35S-OsPTD1_P35S-OsMYB76R and *POsPTD1-OsPTD1_P35S-OsMYB76R*, respectively (**Figures 3A,E**). In addition, they were separately transferred into Zhongjiu B (**Figures 3B,F**). Two types of RNAi-positive plants (purple) were crossed as male parents with a male-sterile Zhongjiu B-*osabcg15* mutant with no purple color. The male parents were also allowed to self-cross. Purple segregations in the coleoptile of F_1 and T_1 were investigated 3 days after the seed germination in wet soil in a disk (**Figures 3C,D,G,H**). The results showed that the T_1 and F_1 of *P35S* promoted RNAi exhibiting 3:1 and 1:1 segregation (purple: no purple), respectively, which did not match with the genetic characteristics of GMS. However, the color segregation ratio of the T_1 of the *POsPTD1* promoted RNAi that did not show a 3:1 ratio but a 1:1 ratio (**Table 1**). Only a few seeds, namely, F_1

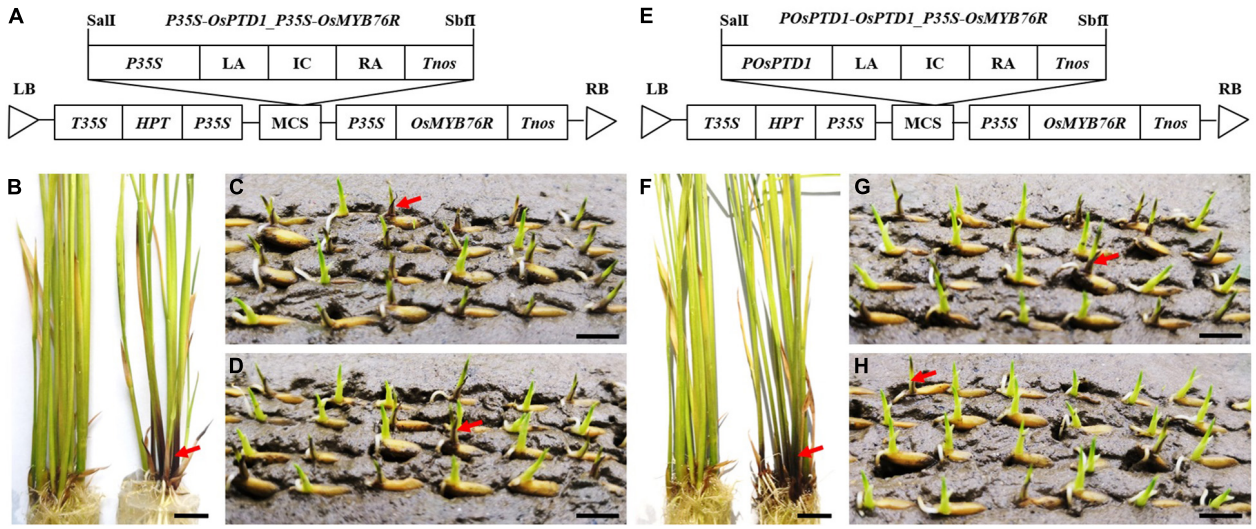

FIGURE 3 | Inheritance characteristics of *OsPTD1* RNAi. Panels **(A,E)** depict the structure of the RNAi vectors *pP35S-OsPTD1_P35S-OsMYB76R* and *pPOsPTD1-OsPTD1_P35S-OsMYB76R*, respectively. Panels **(B–D)** show a T_0-positive plant, and a T_1 and F_1 (positive plants as the male parents) of *P35S-OsPTD1_P35S-OsMYB76R*. Panels **(F–H)** show the T_0-positive plants and T_1 and F_1 (positive plants as the male parent) of *POsPTD1-OsPTD1_P35S-OsMYB76R*. Scale bars, 0.8 cm.

produced by *POsPTD1*-promoted RNAi male plants, displayed a purple coleoptile, which indicated a distorted segregation ratio, compared with the expected 1:1 (**Table 1**). The distorted segregation ratio of the *POsPTD1*-promoted RNAi generation is just the genetic characteristics of GMS, which indicates that *POsPTD1*-driven RNAi produced a partial gametophytic defect. The above results indicate that the preparation of *OsPTD1*-related GMS materials using the newly designed method was successful. Hence, this method can identify GMS easily, rapidly, and economically, 3 days after the seed germination.

DISCUSSION

OsMYB76R Is a Favorable Reporter Gene

It is best to identify the positive transformants as soon as possible after transformation. The reporter gene is a visible indicator of whether the target gene is transferred or not, and it can also be widely used for identifying positive transgenic plants. An ideal reporting gene should have the following characteristics: stable phenotype, easy detection, less background affect, low cost, and no damage to plants, the environment, or health (Ziemienowicz, 2001; Rosellini, 2012; He et al., 2020).

The endogenous genes of some plants have been screened and used as reporter genes. For example, anthocyanin-related genes are among the most widely used endogenous reporters in plants, such as maize, wheat, *Arabidopsis*, tobacco, and tomato (Rosellini, 2012; He et al., 2020). The anthocyanin biosynthesis pathway involves many genes, and at least 16 are involved in the anthocyanin biosynthesis in rice (Sharma and Dixon, 2005; Oshima et al., 2019). Due to the limited carrying capacity of the vector, it is impossible to load all genes controlling the anthocyanin synthesis into the same vector as the reporter genes. Scientists often use a compromise strategy to solve this problem: selecting the appropriate genetic background material, one in which most anthocyanin synthesis genes are functional as the transgenic host, and using one to several anthocyanin genes as the reporter (Lloyd et al., 1992; Goldsbrough et al., 1996; Rosellini, 2012; Oshima et al., 2019).

OsC1/*OsMYB76* is a homologous gene of the anthocyanin biosynthesis gene *C1* in maize, and it is a functional gene for anthocyanin biosynthesis in rice. The purple color phenotype controlled by *OsC1*/*OsMYB76* is visible and stable, which indicates that it is suitable for use as a reporter gene (Xi, 1997; Reddy, 1998; Zhang et al., 2004; Zheng et al., 2019; Hu et al., 2020). To make it easier to use this gene, *OsMYB76R* was optimized from *OsMYB76* and the usual restriction sites in *OsMYB76* CDS were removed. The revised *OsMYB76R* had normal functions, such as those of *OsMYB76,* and could indicate positive transformants quickly, directly, accurately, and harmlessly in the early T_0 generation. Additionally, it was easy to separate the positive plants from the next generations of transgenic plants. The *OsMYB76R* sequence was only 819 bp that saved space for the loading of more and longer target genes. Moreover, the usual restriction sites in *OsMYB76R* were removed, which did not affect the freedom of restriction-site selection, when the target gene is loaded. For this reason, *OsMYB76R* is a desirable reporter gene.

In addition, Zhongjiu B is an *Indica-type* rice with a shorter growth period than Zhonghua 11 (a commonly used *Japonica* variety in rice transgene). The transgenic efficiency of Zhongjiu B is higher than the two popular *Indica* rice types, Minghui 63 and 9311. Zhongjiu B, transformed by only one reporter gene *OsMYB76R*, displays purple color in multiple organs, which indicates that other anthocyanin pigment genes in Zhongjiu B meet genetic conditions sufficient to anthocyanin expression in

TABLE 1 | The segregation ratio of coleoptile purple color in generations of *OsPTD1* RNAi.

Parents or combinations	Generation	Purple color: purple colorless		
		Obs ratio	Exp ratio	$\chi^2_{0.05,1}$
*P35S-OsPTD1*RNAi_*P35S-OsMYB76R-1*(A)	T_1	143:57	3:1	1.13
*P35S-OsPTD1*RNAi_*P35S-OsMYB76R-2*(B)	T_1	152:48	3:1	0.06
Zhongjiu B-*osabcg15* × A	F_1	98:87	1:1	0.54
Zhongjiu B-*osabcg15* × B	F_1	82:72	1:1	0.53
*POsPTD1-OsPTD1*RNAi_*P35S-OsMYB76R-1*(C)	T_1	112:88	1:1	2.65
*POsPTD1-OsPTD1*RNAi_*P35S-OsMYB76R-2*(D)	T_1	103:88	1:1	1.03
Zhongjiu B-*osabcg15* × C	F_1	29:227	0:256	–
Zhongjiu B-*osabcg15* × D	F_1	18:170	0:188	–

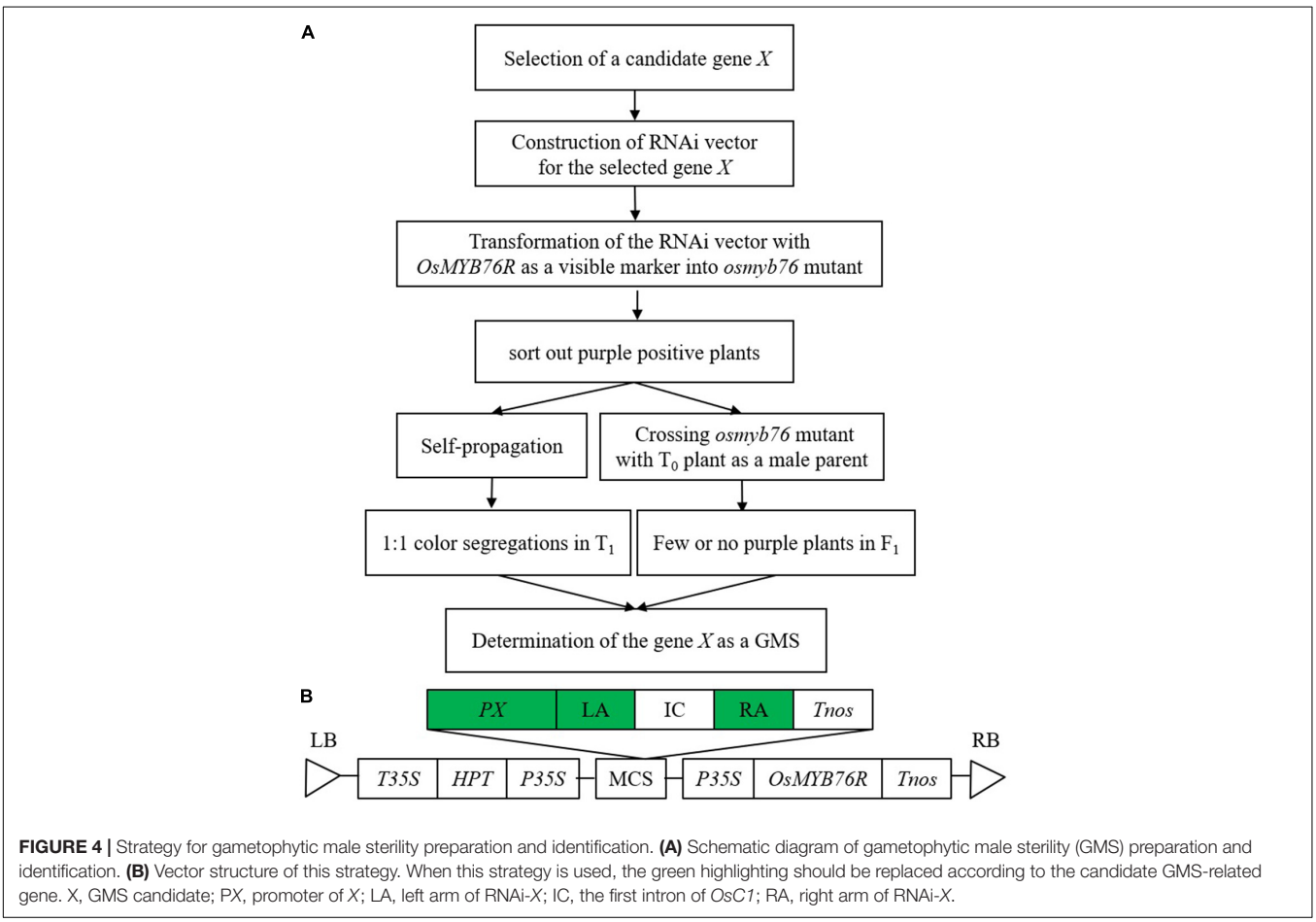

FIGURE 4 | Strategy for gametophytic male sterility preparation and identification. **(A)** Schematic diagram of gametophytic male sterility (GMS) preparation and identification. **(B)** Vector structure of this strategy. When this strategy is used, the green highlighting should be replaced according to the candidate GMS-related gene. X, GMS candidate; P*X*, promoter of *X*; LA, left arm of RNAi-*X*; IC, the first intron of *OsC1*; RA, right arm of RNAi-*X*.

multiple organs. Therefore, Zhongjiu B is a good *Indica* rice host for the transformation of the *OsMYB76R* gene as a reporter.

Combining Candidate Gene RNAi and *OsMYB76R* to Enable High-Efficiency, Accurate, and Low-Cost Preparation and Identification of GMS

Gametophytic male sterility mutants have traditionally been obtained through natural mutation, induced mutation, T-DNA insertion mutation, and gene editing. After mutagenesis, it is necessary to determine whether the plants are GMSs according to the genetic characteristics of their offspring. This can be done with genotype analyses through PCR or chemical resistance analyses using screening-marker genes in self-crossing generation and the F_1 generation (mutant as the male parent). Thereafter, segregation distortion is used to determine whether the mutant is GMS. In other words, if the segregation ratio of the self-bred offspring is 1:1 instead of 3:1, and the mutated alleles do not transfer through pollen learned from F_1, it is GMS (Feldmann et al., 1997; Howden et al., 1998; Boavida et al., 2009). For example, *TMS1* (Yang et al., 2009), *OsGT1*

(Moon et al., 2013), and *gaMS-1* (SariGorla et al., 1996) were prepared and determined as GMS using the above methods. These methods are cumbersome and involve the use of multiple reagents and equipment, which inevitably leads to the problems of low efficiency, high cost, low accuracy, and sample damage. For the gene-editing strategy, scientists must spend another generation to remove the editing component to overcome continuous editing and screen out mutants with no transgenic elements that undoubtedly lengthen the identification time by at least one generation.

In this study, using a strategy that combines RNAi technology and *OsMYB76R* reporting technology, a method was designed that allows GMS preparation and identification at the same time by transferring the linked *OsPTD1*-RNAi element and *OsMYB76R* into Zhongjiu B with no purple color in any organ, the GMS genetic resources of *OsPTD1* were prepared successfully. The genetic identification of GMS was completed only 3 days after the seed germination, using purple segregation distortion in the T_1 and F_1 progenies, derived from T_0 positive plants of RNAi.

There are three strategies used to create GMS: random mutations (such as spontaneous mutation, induced mutation, and T-DNA insertion mutation), gene editing (ZFNs, TALENs, CRISPR-Cas, and so forth), and downregulation of gene expression (with RNAi and anti-RNA). The third strategy was used in this study. RNAi is a much more efficient, easier, and cheaper means of GMS preparation than the random mutation strategy, because the burdensome mutant screening work can be avoided in RNAi. This step is necessary for random mutation because there is no definite target gene. Furthermore, the RNAi strategy is quicker than gene editing, where editing elements must be removed in generation T_1 to avoid any influence on the flowing study by persistent editing. So, if a gene-editing strategy is used to produce GMS, the genetic population for segregation distortion analysis could only be constructed in T_1, not in T_0, while it could be finished in T_0 when the RNAi strategy is adopted. This means that the time required for one generation will be saved. In identification, the first two strategies require PCR, resistance detection, or fluorescence testing to analyze the inheritance behavior, which entails the use of specialized equipment and expensive chemicals and invasive sampling procedures. However, in the RNAi-linked *OsMYB76R* strategy, we use anthocyanin-related genes as reporter genes to analyze their genetic characteristics, in a more direct, simple, timeless, harmless, low-cost, and high-accuracy way than commonly used methods based on PCR detection, resistance testing, or widely used reporting genes, such as *GUS*, *FPs*, and *Luc*.

In conclusion, our strategy for the preparation and identification of GMS has the advantages of being harmless, quick, efficient, cheap, accurate, and reliable. In addition, pollen lethal factor and fluorescent protein are the two key components in the present technology of nuclear sterility seed production. If GMS produced by this method is entirely aborted (no transgenic element transfer through pollen), the RNAi elements and linked *OsMYB76R* can be used to replace the lethal pollen factor and the fluorescent protein gene in the current system that is useful for developing a new nuclear-sterility seed-preparation system. Obviously, this strategy can also be used to prepare and identify GMS in other crops. A simple sketch map for understanding this method is shown in **Figure 4**.

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

YC, WZ, and SS designed and performed the experiments. LJ, KY, WZ, and JY performed the experiments and analyzed the data. LW and SD designed the experiments. LG, ZW, and YZ designed the experiments and wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (31701390 and 31370349), Special Project on Performance Incentive Guidance of Chongqing Scientific Research Institution (cstc2018jxjl80021), Fund of Agriculture Development (NKY-2021AC003), and Recruitment Announcement for High-level Talents of Yunnan University (KL180018).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.728193/full#supplementary-material

Supplementary Material 1 | Sequence of synthesized *OsMYB76R*.

Supplementary Material 2 | Sequence of synthesized *OsPTD1* RNAi elements.

Supplementary Material 3 | Primer sequences for *OsMYB76R* amplification.

REFERENCES

Ariizumi, T., and Toriyama, K. (2011). Genetic regulation of sporopollenin synthesis and pollen exine development. *Annu. Rev. Plant Biol.* 62, 437–460. doi: 10.1146/annurev-arplant-042809-112312

Boavida, L. C., Shuai, B., Yu, H. J., Pagnussat, G. C., Sundaresan, V., and Mccormick, S. (2009). A collection of ds insertional mutants associated with defects in male gametophyte development and function in *Arabidopsis thaliana*. *Genetics* 181, 1369–1385. doi: 10.1534/genetics.108.090852

Chen, R., Fa-Ke, Y. U., Liu, H. Q., Yang, S. H., and Wang, F. (2012). Characterization of male gametophytic sterile mutant collections in rice by t-dna insertion. *Chin. J. Rice Sci.* 26, 173–181.

Feldmann, K. A., Coury, D. A., and Christianson, M. L. (1997). Exceptional segregation of a selectable marker (kanr) in *Arabidopsis* identifies genes

important for gametophytic growth and development. *Genetics* 147, 1411–1422. doi: 10.1093/genetics/147.3.1411

Goldsbrough, A. P., Tong, Y., and Yoder, J. I. (1996). Lc as a nondestructive visual reporter and transposition excision marker gene for tomato. *Plant J*. 9, 927–933. doi: 10.1046/j.1365-313x.1996.9060927.x

Groot, P. D., Weterings, K., Been, M. D., Wittink, F., and Wullems, G. (2004). Silencing of the pollen-specific gene ntp303 and its family members in tobacco affects in vivo pollen tube growth and results in male sterile plants. *Plant Mol. Biol*. 55, 715–726. doi: 10.1007/s11103-004-1964-6

Han, M. J., Jung, K. H., Yi, G., and An, G. (2011). Rice importin β1 gene affects pollen tube elongation. *Mol. Cells* 31, 523–530. doi: 10.1007/s10059-011-2321-7

Han, M. J., Jung, K. H., Yi, G., Lee, D. Y., and An, A. G. (2006). Rice immature pollen 1 (rip1) is a regulator of late pollen development. *Plant Cell Physiol*. 47, 1457–1472. doi: 10.1093/pcp/pcl013

He, Y., Zhang, T., Sun, H., Zhan, H., and Zhao, Y. (2020). A reporter for noninvasively monitoring gene expression and plant transformation. *Hortic. Res*. 7:152.

Howden, R., Park, S. K., Moore, J. M., Orme, J., and Twell, D. (1998). Selection of t-dna-tagged male and female gametophytic mutants by segregation distortion in *Arabidopsis*. *Genetics* 149, 621–631. doi: 10.1093/genetics/149.2.621

Hu, W., Zhou, T., Han, Z., Tan, C., and Xing, Y. (2020). Dominant complementary interaction between osc1 and two tightly linked genes, rb1 and rb2, controls the purple leaf sheath in rice. *Theor. Appl. Genet*. 133, 2555–2566. doi: 10.1007/s00122-020-03617-w

Liu, L., Zheng, C., Kuang, B., Wei, L., and Tai, W. (2016). Receptor-like kinase rupo interacts with potassium transporters to regulate pollen tube growth and integrity in rice. *PLoS Genet*. 12:e1006085. doi: 10.1371/journal.pgen.1006085

Lloyd, A., Walbot, V., and Davis, R. (1992). *Arabidopsis* and nicotiana anthocyanin production activated by maize regulators r and c1. *Science* 258, 1773–1775. doi: 10.1126/science.1465611

Ma, H. (2005). Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. *Annu. Rev. Plant Biol*. 56, 393–434.

Moon, S., Kim, S. R., Zhao, G., Yi, J., Yoo, Y., Jin, P., et al. (2013). Rice glycosyltransferase1 encodes a glycosyltransferase essential for pollen wall formation. *Plant Physiol*. 161, 663–675.

Oshima, M., Taniguchi, Y., Akasaka, M., Abe, K., Ichikawa, H., Tabei, Y., et al. (2019). Development of a visible marker trait based on leaf sheath-specific anthocyanin pigmentation applicable to various genotypes in rice. *Breeding Sci*. 69, 244–254.

Reddy, V. (1998). Cloning and characterization of the rice homologue of the maize C1 anthocyanin regulatory gene. *Plant Mol. Biol*. 36, 497–498.

Rosellini, D. (2012). Selectable markers and reporter genes: a well furnished toolbox for plant science and genetic engineering. *Crit. Rev. Plant Sci*. 31, 401–453.

SariGorla, M., Ferrario, S., Villa, M., and Pe, M. E. (1996). gaMS-1: a gametophytic male sterile mutant in maize. *Sex Plant Reprod*. 9, 216–220. doi: 10.1007/BF02173101

Sharma, S. B., and Dixon, R. A. (2005). Metabolic engineering of proanthocyanidins by ectopic expression of transcription factors in *Arabidopsis thaliana*. *Plant J*. 44, 62–75. doi: 10.1111/j.1365-313X.2005.02510.x

Shi, J., Cui, M., Yang, L., Kim, Y. J., and Zhang, D. (2015). Genetic and biochemical mechanisms of pollen wall development. *Trends Plant Sci*. 20, 741–753. doi: 10.1016/j.tplants.2015.07.010

Sun, X., Zhang, Z., Chen, C., Wu, W., Ren, N., Jiang, C., et al. (2018). The C-S-A gene system regulates hull pigmentation and reveals evolution of anthocyanin biosynthesis pathway in rice. *J. Exp. Bot*. 69, 1485–1498. doi: 10.1093/jxb/ery001

Suwabe, K., Suzuki, G., Takahashi, H., Shiono, K., Endo, M., Yano, K., et al. (2008). Separated transcriptomes of male gametophyte and tapetum in rice: validity of a laser microdissection (LM) microarray. *Plant Cell Physiol*. 49, 1407–1416. doi: 10.1093/pcp/pcn124

Wang, M. L., Yan, W., Peng, X. Q., Chen, Z. F., Xu, C. J., Wu, J. X., et al. (2020). Identification of late-stage pollen-specific promoters for construction of pollen-inactivation system in rice. *J. Integr. Plant Biol*. 62, 1246–1263. doi: 10.1111/jipb.12912

Wilson, Z. A., and Zhang, D. B. (2009). From *Arabidopsis* to rice: pathways in pollen development. *J. Exp. Bot*. 60, 1479–1492. doi: 10.1093/jxb/erp095

Wu, L. N., Guan, Y. S., Wu, Z. G., Yang, K., Lv, J., Converse, R., et al. (2014). OsABCG15 encodes a membrane protein that plays an important role in anther cuticle and pollen exine formation in rice. *Plant Cell Rep*. 33, 1881–1899. doi: 10.1007/s00299-014-1666-8

Xi, J. (1997). The discovery and primary research of coleoptile purple lines in rice. *Hybrid Rice* 12:41.

Yang, K. Z., Xia, C., Liu, X. L., Dou, X. Y., Wang, W., Chen, L. Q., et al. (2009). A mutation in THERMOSENSITIVE MALE STERILE 1, encoding a heat shock protein with DnaJ and PDI domains, leads to thermosensitive gametophytic male sterility in *Arabidopsis*. *Plant J*. 57, 870–882. doi: 10.1111/j.1365-313X.2008.03732.x

Yang, S. H., Chen, R., Liu, H. Q., Zhou, S. F., and Wang, F. (2012). Identification and preliminary analysis of a male gametophytic mutant mga36 in rice. *Fujian J. Agric. Sci*. 27, 1039–1043.

Zhang, Y., Li, Y. F., Liu, X. F., Lin, M. X., Shen, F. C., He, G. H., et al. (2004). Analysis on the inheritance of coleoptile purple line in rice. *Sci. Agric. Sin*. 37, 1693–1698.

Zhao, S. S., Wang, C. H., Ma, J., Wang, S., Tian, P., Wang, J. L., et al. (2016). Map-based cloning and functional analysis of the chromogen gene C in rice (*Oryza sativa L.*). *J. Plant Biol*. 59, 496–505. doi: 10.1007/s12374-016-0227-9

Zheng, J., Wu, H., Zhu, H. B., Huang, C. Y., Liu, C., Chang, Y. S., et al. (2019). Determining factors, regulation system, and domestication of anthocyanin biosynthesis in rice leaves. *New Phytol*. 223, 705–721. doi: 10.1111/nph.15807

Ziemienowicz, A. (2001). Plant selectable markers and reporter genes. *Acta Physiol. Plant*. 23, 363–374. doi: 10.1007/s11738-001-0045-6

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 in Plant Science

ORIGINAL RESEARCH
published: 08 March 2022
doi: 10.3389/fpls.2022.866404

Loss of OsEAF6, a Subunit of the Histone Acetyltransferase Complex, Causes Hybrid Breakdown in Intersubspecific Rice Crosses

Takahiko Kubo[1]†, Atsushi Yoshimura[2] and Nori Kurata[1]*

[1] Plant Genetics Laboratory, National Institute of Genetics, Mishima, Japan, [2] Plant Breeding Laboratory, Faculty of Agriculture, Kyushu University, Fukuoka, Japan

OPEN ACCESS

Edited by:
Dayun Tao,
Yunnan Academy of Agricultural Sciences, China

Reviewed by:
Zhang Yu,
Yunnan Academy of Agricultural Sciences, China
Hee-Jong Koh,
Seoul National University, South Korea

***Correspondence:**
Takahiko Kubo
takubo@agr.kyushu-u.ac.jp

†Present address:
Takahiko Kubo,
Plant Genetics Laboratory, Faculty of Agriculture, Kyushu University, Fukuoka, Japan

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

Received: 31 January 2022
Accepted: 08 February 2022
Published: 08 March 2022

Citation:
Kubo T, Yoshimura A and Kurata N (2022) Loss of OsEAF6, a Subunit of the Histone Acetyltransferase Complex, Causes Hybrid Breakdown in Intersubspecific Rice Crosses. Front. Plant Sci. 13:866404. doi: 10.3389/fpls.2022.866404

Gene duplication plays an important role in genetic diversification, adaptive evolution, and speciation. Understanding the mechanisms and effects of postzygotic isolation genes is important for further studies of speciation and crop breeding. The duplicate recessive genes *hwe1* and *hwe2* cause hybrid breakdown, characterized by poor vegetative growth and reproductive dysgenesis in intersubspecific crosses between *Oryza sativa* ssp. *indica* and *japonica*. Using a map-based cloning strategy, we found that *HWE1* and *HWE2* encode the Esa1-associated factor 6 (EAF6) protein, a component of histone acetyltransferase complexes. The *indica hwe1* and *japonica hwe2* alleles lacked functional *EAF6*, demonstrating that the double recessive homozygote causes hybrid breakdown. Morphological and physiological observations showed that weak plants with double recessive homozygotes had serious morphological defects with a wide range of effects on development and organs, leading to leaves with reduced chlorophyll content, flower and pistil malformation, and anomalies of gametogenesis. These findings suggest that EAF6 plays a pivotal role in the transcriptional regulation of essential genes during the vegetative and reproductive development of rice.

Keywords: hybrid breakdown, histone acetyltransferase, rice, duplicate recessive gene, speciation

INTRODUCTION

In eukaryotic cells, histone acetylation regulates the chromatin structure, affecting gene transcription, DNA replication, and DNA damage repair. Nucleosome acetyltransferase of histone 4 (NuA4), a histone acetyltransferase (HAT) complex, is composed of multiple proteins and preferentially acetylates histones H4 and H2A on the nucleosome. The components of NuA4 are highly conserved in yeast and human (Doyon et al., 2004). Yeast NuA4 consists of 13 subunits, with two independent NuA4 sub-complexes, namely, piccolo-NuA4, composed of Esa1, Epl1, Yng2, and Eaf6, and the TINTIN triad of Eaf5/7/3 (Wang X. et al., 2018). Piccolo-NuA4, which is thought to also exist alone, contains the catalytic subunit protein essential Sas2-related acetyltransferase-1 (Esa1) (Ohba et al., 1999; Boudreault et al., 2003). Esa1 alone can acetylate free histones but cannot acetylate nucleosomal histones (Doyon et al., 2004). This protein also plays a crucial role in cell cycle progression and DNA double-strand break repair (Clarke et al., 1999;

Bird et al., 2002) and is essential for yeast cell viability (Doyon et al., 2004). Another component, Esa1-associated factor 6 (EAF6, known as MEAF6 in mammals), interacts with Piccolo-NuA4 through Yng2 in yeast (Mitchell et al., 2008). Unlike the catalytic subunit Esa1, the yeast *eaf6*Δ mutant is viable without detectable changes, indicating that the NuA4 subunit is not essential for yeast cellular processes (Lafon et al., 2007). However, a study in human showed that a fusion protein of MEAF6 with PHD finger protein 1 generated by chromosomal translocation caused endometrial stromal tumors in human (Panagopoulos et al., 2012). These findings suggest that EAF6 (MEAF6) is important in cell proliferation. Yeast EAF6 is also a component of another HAT complex, nucleosome acetyltransferase of histone 3, which acetylates histone H3. This HAT complex was reported to be involved in transcriptional activation and cell cycle regulation (Lafon et al., 2007). Compared with the number of studies performed to characterize NuA4 components in yeast and mammals, studies in plants are limited. Two components of the NuA4 complex, namely, EAF1 and YAF9, were found to regulate flowering *via* histone H4 acetylation in *Arabidopsis* (Zacharaki et al., 2012; Bieluszewski et al., 2015). Double mutations in HAM1 and HAM2 (*Arabidopsis* ESA1 homologs) induced lethality in diploid plants and haploid gametophytes, suggesting a function in histone acetylation during mitotic cell division of gametogenesis in *Arabidopsis* (Latrasse et al., 2008). The plant EAF6 protein is an uncharacterized potential subunit of plant NuA4; less is known about its biological functions in plant development and growth.

Hybrid breakdown is defined as deleterious characteristics, such as sterility and non-viability, occurring only after F_2 generations of crosses between distantly related species. Similar to other hybrid incompatibility mechanisms, hybrid breakdown contributes to speciation by restricting gene flow between diverging taxa. Although such phenomena are widely observed in numerous animal and plant species (Stebbins, 1958), few hybrid breakdown genes have been identified and characterized at the molecular level. Seminal studies in plants demonstrated that autoimmune responses involving nucleotide-binding site-leucine-rich repeat genes can cause hybrid weakness and breakdown (Bomblies et al., 2007; Alcazar et al., 2010). A further systematic study using a large diallel cross containing more than 6,400 cross combinations in *Arabidopsis* revealed that one hybrid necrosis gene, *Dangerous Mix 2* (*DM2*), plays a central role in the epistatic network involving numerous independent loci related to hybrid necrosis (Chae et al., 2014). Such autoimmune systems cause hybrid incompatibility, including the hybrid breakdown in other plant species (Hannah et al., 2007; Jeuken et al., 2009; Yamamoto et al., 2010; Chen et al., 2014). Previous studies demonstrated that defense systems against biotic stress, including programmed cell death and nucleotide-binding site-leucine-rich repeats, are major and common causes of hybrid weakness and hybrid breakdown in plant species. Although progress has been made recently in identifying the genes involved in hybrid incompatibility, the molecular basis of hybrid breakdown other than the autoimmune response is poorly understood in plants. In many cases, the causal genes remain unknown. Whether other physiological mechanisms underlie hybrid weakness and how different genes contribute to genetic diversification and speciation are also unclear.

Our previous study demonstrated hybrid breakdown characterized by weak growth and complete sterility between *Oryza sativa* ssp. *indica* and *japonica* ($2n = 24$). Genetic analysis has revealed that this weakness is caused by double recessive genes, *hwe1* and *hwe2*, which are localized on rice chromosomes 1 and 12, respectively (Kubo and Yoshimura, 2002). Although the phenomenon and basic genetics of this hybrid breakdown were characterized more than a decade ago, the molecular mechanism is not well-understood. The specific objectives of this study were to isolate *hwe1* and *hwe2*, characterize the abnormal phenotype of the weak plant morphologically and physiologically, and identify the physiological function of the causal genes at the molecular level.

MATERIALS AND METHODS

Plant Materials

The characterization of weak plants and high-resolution mapping was carried out using the backcross population derived from the population previously used for rough mapping of *hwe1* and *hwe2* (Kubo and Yoshimura, 2002). Additionally, we used two other *indica/japonica* populations, namely, BC_2F_2 derived from the Nipponbare/93-11 cross (Kubo et al., 2011) and a newly developed Nipponbare/IR8 F_2 population.

Characterization of Morphological and Physiological Traits

Normal and weak segregants (BC_3F_{6-7}, $n = 10$) were evaluated to determine their seed fertility, column length, and number. The chlorophyll content was examined using the last fully opened leaf blades from each genotype ($n = 5$) at the tillering stage. Seed fertility was evaluated as previously described (Kubo et al., 2016). To evaluate leaf cell viability related to the autoimmune response, leaves from normal and weak plants during vegetative development were collected and stained with trypan blue. For staining, the detached leaves were completely submerged in lactic acid-phenol-trypan blue solution (0.5 mg/ml trypan blue, 25% phenol, 25% lactic acid, and 25% glycerol) and microwaved for 1.0 min in a domestic microwave oven. The tissue was destained by placing the samples in staining solution without trypan blue and overnight incubation. The tissue was transferred to 50% ethanol and observed under a stereomicroscope.

Map-Based Cloning

To perform high-resolution mapping of the *hwe1 and hwe2* loci, seedlings of the segregating populations (approximately 2,387 BC_3F_{5-6} plants for *hwe2* and 383 BC_3F_{5-6} individuals for *hwe1*) were genotyped using polymerase chain reaction (PCR)-based markers, and plants with recombination around the *hwe1* and *hwe2* loci were identified. PCR-based markers, insertion and deletion markers, and simple sequence repeat markers were identified using sequence polymorphism data for

Nipponbare and 93-11 (MSU7.0)[1]. The primer sequences for the DNA markers are listed in **Supplementary Table 1**. For DNA marker genotyping, crude DNA extracts of seedling leaves were prepared using 0.25 M NaOH followed by neutralization with 0.1 M Tris–HCl. These DNA extracts were used in PCR with GoTaq polymerase (Promega, Madison, WI, United States) and the following cycling profile: 94°C for 2 min, followed by 30 cycles of 94°C for 20 s, 50–60°C for 20 s, and 72°C for 30 s.

Gene Cloning and Rice Transformation

Gene cloning and allelic diversity analyses were carried out using PCR analysis of purified DNA from rice varieties and wild accessions prepared using the CTAB method (Murray and Thompson, 1980), followed by sequencing analysis on an ABI Prism 3130xl genetic analyzer (Applied Biosystems, Foster City, CA, United States). For the complementation test, the Asominori genomic DNA fragment (7,023 bp) containing the LOC_Os12g20310 gene with flanking 5′ (2,872 bp) and 3′ (1,487 bp) regions was amplified with KOD-Plus-Neo DNA polymerase (Toyobo, Osaka, Japan) using the primer G20310 (**Supplementary Table 1**). For another candidate, LOC_Os12g20324, the Asominori genomic DNA fragment (7,689 bp) was amplified by PCR using the primer G20324. The amplified fragments were cloned into the pBluescript SK cloning vector and subcloned into the pPZP2H-lac binary vector (Fuse et al., 2001). The cloned genomic fragments and empty vector were transformed into *HWE1* heterozygotes (*Hwe1/hwe1hwe2/hwe2*) *via Agrobacterium tumefaciens*-mediated transformation (Hiei et al., 1994; Nishimura et al., 2006). Complementation was examined based on the phenotype of the selfed progeny (T1 and T2) of the T0 transformant.

Reverse Transcription-PCR Analysis

Total RNA from Asominori plant tissues (i.e., leaf, stem, root, young, and flowering panicles) and NILs were prepared using an RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). cDNA was generated by reverse transcription of 2.0 μg of total RNA using the SuperScript III First-Strand Synthesis System (Life Technologies, Carlsbad, CA, United States). RT-PCR analysis was performed using two primer sets (i.e., eaf6sp and eaf6co) (**Supplementary Table 1**) to discriminate the products from the Nipponbare *HWE1* and *HWE2* loci. The cDNA of *OsAct1* (rice *Actin 1*) was amplified using the Act1 primer and used as a standard control. RT-PCR was performed in a Biometra thermocycler with the following cycling profile: 94°C for 2 min, followed by 32 cycles at 94°C for 20 s, 58°C for 20 s, and 72°C for 30 s.

Histological Experiments

To observe the embryo sac, pre-flowering panicles were collected from normal and weak plants, fixed, and stored in FAA solution (45% ethanol, 5% formalin, and 5% acetic acid). After fixation, the samples were embedded in paraffin (Paraplast Plus; McCormick Scientific, St. Louis, MO, United States), sectioned, and stained with hematoxylin. To observe the morphology of the mature pollen grains, ethanol-fixed pollen grains were stained with 1.0% iodine-potassium iodide (I_2-KI) and observed under a microscope. Male gametogenesis was analyzed using young panicles collected from normal and weak plants at different developmental stages. Panicles were fixed and stored in FAA solution. After fixation, the micropores were extracted from the anther using forceps, and released micropores were stained with hematoxylin solution as described by Kindiger and Beckett (1985).

Histochemical Analysis of Beta-Glucuronidase Expression

An Asominori genome fragment containing 2,872 bp of the upstream region of *OsEAF6* (*LOC_Os12g20310*) was amplified using the primer set G20310pro (**Supplementary Table 1**) and cloned into the *Kpn* I-*Spe* I site of the binary vector pBGH2 (Ito and Kurata, 2008) to drive beta-glucuronidase (GUS) expression. The resulting construct, *ProOsEAF6:GUS*, was transformed into Asominori plants. For GUS staining, tissue samples (i.e., leaf, young spikelets, and stem) were vacuum-infiltrated with staining solution (50 mM Na_3PO_4, pH 7.0, 1.0 mg/ml 5-bromo-4-chloro-3-indolyl-β-glucuronide, and 0.5% Triton X-100) and incubated at 37°C for 16 h. The stained samples were fixed for 10 min in formaldehyde, acetic acid, and 22% ethanol (5:5:90, v/v) and then destained in 100% ethanol until the chlorophyll was removed.

Subcellular Localization Analysis

Transient expression assays using polyethylene glycol-mediated transformation were performed as previously described (Shen et al., 2014). The *35S:OsEAF6-mCherry* construct was prepared by amplifying mCherry from the pmCherry-C1 vector (Takara) and inserting it into the plant binary vector pRI201-ON (Takara). The coding sequence (CDS) of *OsEAF6* (*LOC_Os12g20310*) was cloned into the pCR-Blunt II TOPO vector (Life Technologies) and transferred into the mCherry-pRI201-ON vector. Rice Oc cells were provided by RIKEN BRC through the National Bio-Resource Project of MEXT, Japan. Protoplasts from Oc cells were adjusted to a concentration of 1.0–2.0 $\times$ 10^6 cells/ml; a 0.1-ml aliquot was transfected with 10–20 μg plasmids. After 16 h of incubation at 28°C, transformed cells were observed under an optical/fluorescence microscope (Biozero BZ-8000, Keyence, Osaka, Japan).

Genome Sequence and Haplotype Analysis

The genome sequences of cultivated and wild rice accessions were downloaded from the Gramene database[2]. Genome sequences of the *japonica* variety Nipponbare (IRGSP version 1.0) and *indica* varieties IR8 (IOMAP version 1) and 93-11 (ASM465v1) were used to compare chromosomal structural differences between *indica* and *japonica* around the *HWE1* and *HWE2* regions. The genomic sequences were identified using GenomeMatcher version 2.03 software (Ohtsubo et al., 2008). The genome sequences of the following species were also used:

[1]http://rice.plantbiology.msu.edu/

[2]http://ensembl.gramene.org/Oryza_sativa/Info/Index

Oryza nivara, accession W0106 (version AWHD00000000); *Oryza rufipogon*, W1943 (PRJEB4137); *Oryza barthii*, IRGC105608 (ABRL00000000); *Oryza glaberrima*, IRGC96717 (AGI1.1); *Oryza glumaepatula*, GEN1233_2 (ALNU02000000); *Oryza meridionalis*, W2112 (Oryza_meridionalis_v1.3); *Oryza punctata*, IRGC105690 (AVCL01000000); *Oryza brachyantha*, IRGC101232 (AGAT00000000), and *Leersia perrieri*, IRGC105164 (version 1.4). Gene annotation of Leersia was based on the Gramene database, and the others were based on RiceGAAS (Sakata et al., 2002). Haplotype analysis of the *HWE1* and *HWE2* loci was performed based on published SNP data for cultivated and wild rice compared with the Nipponbare genome. SNP data for cultivars and wild rice were obtained from the International Rice Information System[3] and OryzaGenome[4], respectively.

Phylogenetic Analysis of EAF6 Protein Homologs

A search for EAF6 homologs in different organisms was performed using the BLASTP program (NCBI) website. Protein sequences of 25 plant species and 8 microbe and animal species were obtained from phytozome[5] and the NCBI database (**Supplementary Table 2**). Protein sequences were aligned using ClustalW to construct a phylogenetic tree using the neighbor-joining method (Saitou and Nei, 1987) and MEGA6 software (Tamura et al., 2013). Bootstrap values were calculated with 1,000 replications.

RESULTS

Characteristics of Hybrid Breakdown

A single substitution with *indica* (cv. IR24) chromosome 12 in a *japonica* (Asominori) genetic background causes hybrid breakdown characterized by poor growth and complete sterility (Kubo and Yoshimura, 2002). Simple epistasis with double recessive genes, named *hwe1* and *hwe2*, was sufficient to explain this hybrid breakdown (**Figures 1A–C**). Similarly, genetic analysis showed that epistasis between *hwe1* and *hwe2* caused a hybrid breakdown in other cross combinations (Nipponbare/93-11 and Nipponbare/IR8) (**Supplementary Figures 1, 2**). The weak phenotype was characterized by shorter and smaller culm lengths, partially sheathed panicles, and both male and female sterility (**Figures 1B–F** and **Supplementary Figures 3A,B**). The weak segregant appeared as pale green compared with normal plants at the adult stage, likely due to the lower chlorophyll content in the leaves (approximately half of that in the normal segregant) (**Figure 1F**). Leaves from weak plants during the vegetative growth stage were not stained by trypan blue (**Figures 1G,H**), suggesting that this weak phenotype did not result from an autoimmune response by nucleotide-binding site-leucine-rich repeat or other related molecules, as previously reported (Bomblies et al., 2007; Alcazar et al., 2010). A much more severe phenotype was observed in the reproductive organs. Weak plants produced smaller panicles and yielded fewer spikelets compared with normal plants (**Figures 1B,F**). An abnormal phenotype was also found in the reproductive organs, such as depressed palea or palea-less flowers, degenerated anthers, and abnormal stigma formation (**Supplementary Figures 4A–L**). Microscopy revealed that pollen sterility was attributed to meiotic defects, including abnormal or incomplete cell division (**Supplementary Figures 3C–M**). Although the female gametes were also completely sterile, various ovules swollen by imbibition were observed without fertilization in weak plants (**Supplementary Figure 4M**).

Identification of *HWE1* and *HWE2*

HWE1 and *HWE2* have been roughly mapped to rice chromosomes 12 and 1, respectively (Kubo and Yoshimura, 2002). Using the small-scale mapping population (N = 383), we identified the *HWE1* locus within a 5.4-Mb region close to the centromere of chromosome 12 (**Figure 2A**). As meiotic recombination is repressed around the centromere and pericentromeric regions, we then focused on the partner gene *HWE2*. To isolate *HWE2*, we screened recombinant individuals from a large segregating population (N = 2,387). The result showed that the *HWE2* locus was delimited within a 38.7-kb region between the PCR markers *1c215* and *1c219*, which encoded four predicted genes (i.e., *LOC_Os01g13210*, *LOC_Os01g13229*, *LOC_Os01g13250*, and *LOC_Os01g13260*) (**Figure 2A**). Of these four genes, two genes (i.e., *LOC_Os01g13250* and *LOC_Os01g13260*) shared homology with *LOC_Os12g20310* and *LOC_Os12g20324* in the *HWE1* region, indicating a small segmental duplication between chromosomes 1 and 12 of the Nipponbare genome. In comparative sequence analysis, the DNA sequence of the *indica Hwe2* allele showed much greater similarity to that of the *japonica Hwe1* allele than to that of the *japonica hwe2* allele (**Supplementary Figure 5**). However, there were no segmental blocks corresponding to *LOC_Os12g20310* and *LOC_Os12g20324* on *indica* chromosome 12 (93-11 and IR8 genomes). In the context of double recessive epistasis, these duplicated genes were considered good candidates for *HWE1/2*. The *LOC_Os01g13250* and *LOC_Os12g20310* loci encode *EAF6*, and the adjacent *LOC_Os01g13260* and *LOC_Os12g20324* encode the cyclin-A1 protein. We performed complementation analysis to determine whether one or both of these genes contribute to hybrid breakdown. Due to the complete sterility of weak segregants, we transformed *Hwe1/hwe1* heterozygotes with Asominori genomic DNA containing *LOC_Os12g20310* or *LOC_Os12g20324* and then evaluated the complementation in their selfed progeny (T1 and T2 generations). The resultant transformants with *LOC_Os12g20310* recovered the weak growth phenotype with complete sterility, whereas the other transformants with *LOC_Os12g20324* did not (**Supplementary Figure 6** and **Supplementary Tables 3, 4**). This result indicates that *OsEAF6*, encoded by *HWE1/2*, was responsible for the hybrid breakdown. A full-length cDNA of *OsEAF6* has been previously cloned (GenBank, CT830710). The gene structure of the Nipponbare allele of *LOC_Os01g13250* was predicted

[3]http://oryzasnp.org/iric-portal/index.zul

[4]http://viewer.shigen.info/oryzagenome/mapview/Top.do

[5]https://phytozome-next.jgi.doe.gov

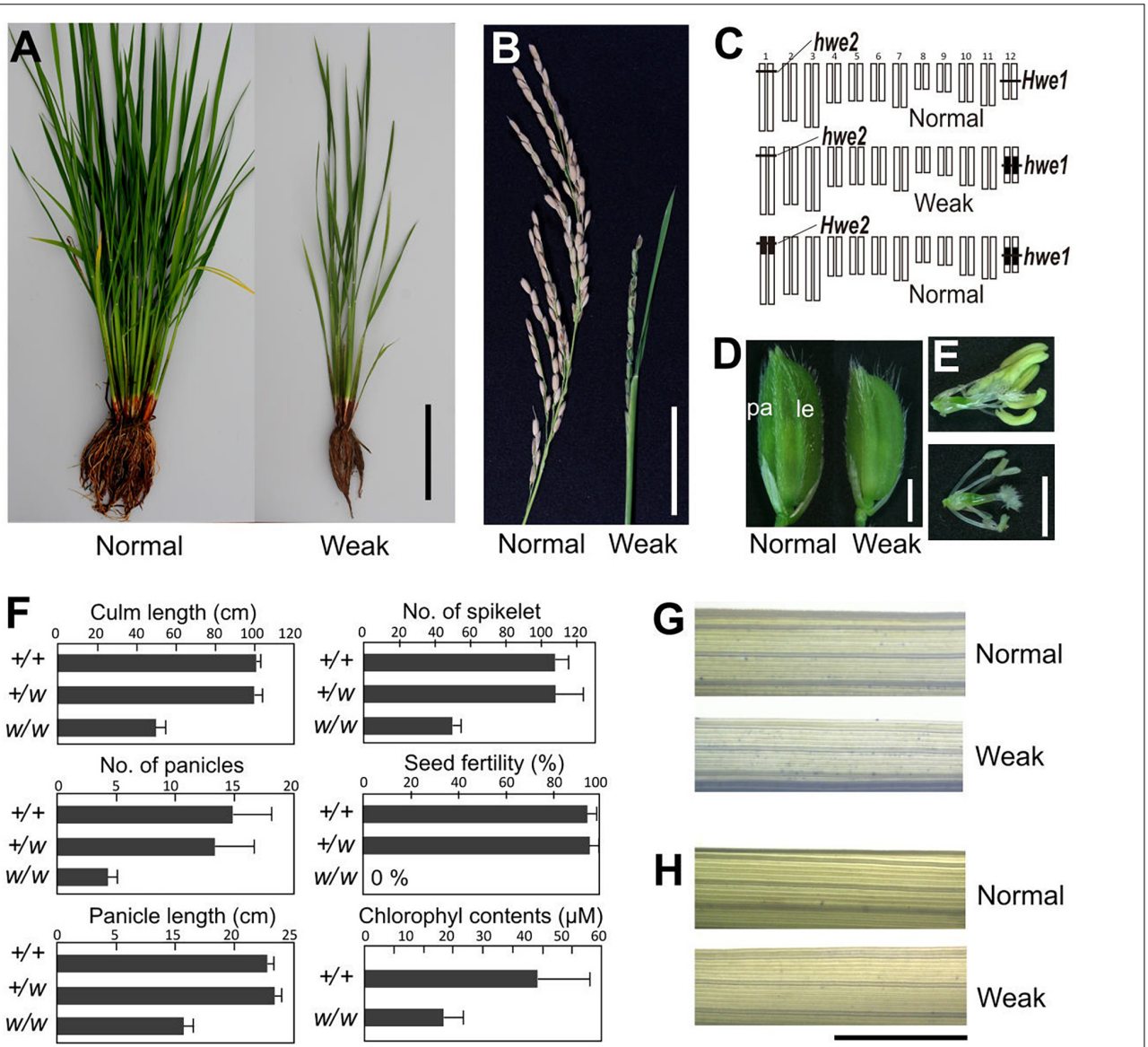

FIGURE 1 | Morphology of weak plants. **(A)** Plant morphology at vegetative stage 3 weeks before heading. **(B)** Morphology of panicle at the mature seed stage. Weak plants exhibited a partially sheathed panicle. **(C)** Graphical genotype of weak plants. White and black bars represent Asominori and IR24 chromosomes, respectively. Only the double recessive homozygote showed the weakness phenotype. **(D)** Spikelets of normal (*left*) and weak plants (*right*). Le, lemma; pa, palea. **(E)** Flower organs of normal (*upper*) and weak plants (*lower*). **(F)** Characterization of weak plants. Column length, number of panicles, number of spikelets per panicle, seed fertility, and leaf chlorophyll contents at 3 weeks before flowering. +/+: *Hwe1/Hwe1 hwe2/hwe2*, +/w: *Hwe1/hwe1, hwe2/hwe2*, and w/w: *hwe1/hwe1 hwe2/hwe2*. *N* = 10 for each genotype excluding the chlorophyll contents (*N* = 5). **(G,H)** Trypan blue staining to assess viability leaf blade cell. Third youngest leaf blades from normal and weak plants at the developmental stage before heading [**(G)**, 90 days after sowing] and after heading [**(H)**, 120 days after sowing) were stained by trypan blue. Neither showed remarkable cell lethality. Scale bar = 10.0 cm in **(A)**, 5.0 cm in **(B)**, 2.0 mm in **(D,E)**, and 5.0 mm in **(G,H)**.

to encode a variant form of *OsEAF6* with extra amino acids (14 extra amino acids) in the C-terminus (**Figures 2B,C** and **Supplementary Figure 7**). We predicted that this variant form (*hwe2-j*) was functionally defective.

Hybrid Breakdown Attributed to Loss of *OsEAF6* Expression

To verify this hypothesis, we examined the expression of *LOC_Os12g20310* and *LOC_Os01g13250* in Asominori and IR24 using RT-PCR with two primer sets, namely, eaf6sp and eaf6co. The first primer set, eaf6sp, was specific to Nipponbare *LOC_Os12g20310* but not to *LOC_Os01g13250*, which encodes the valiant form. The second primer, eaf6co, matched the conserved identical sequence between *LOC_Os12g20310* and *LOC_Os01g13250* and was used to examine the presence of the variant mRNA from *LOC_Os01g13250* (**Figure 3A**). The mRNA from *LOC_Os12g20310* was ubiquitously expressed in all organs examined, except for the root, which showed faint

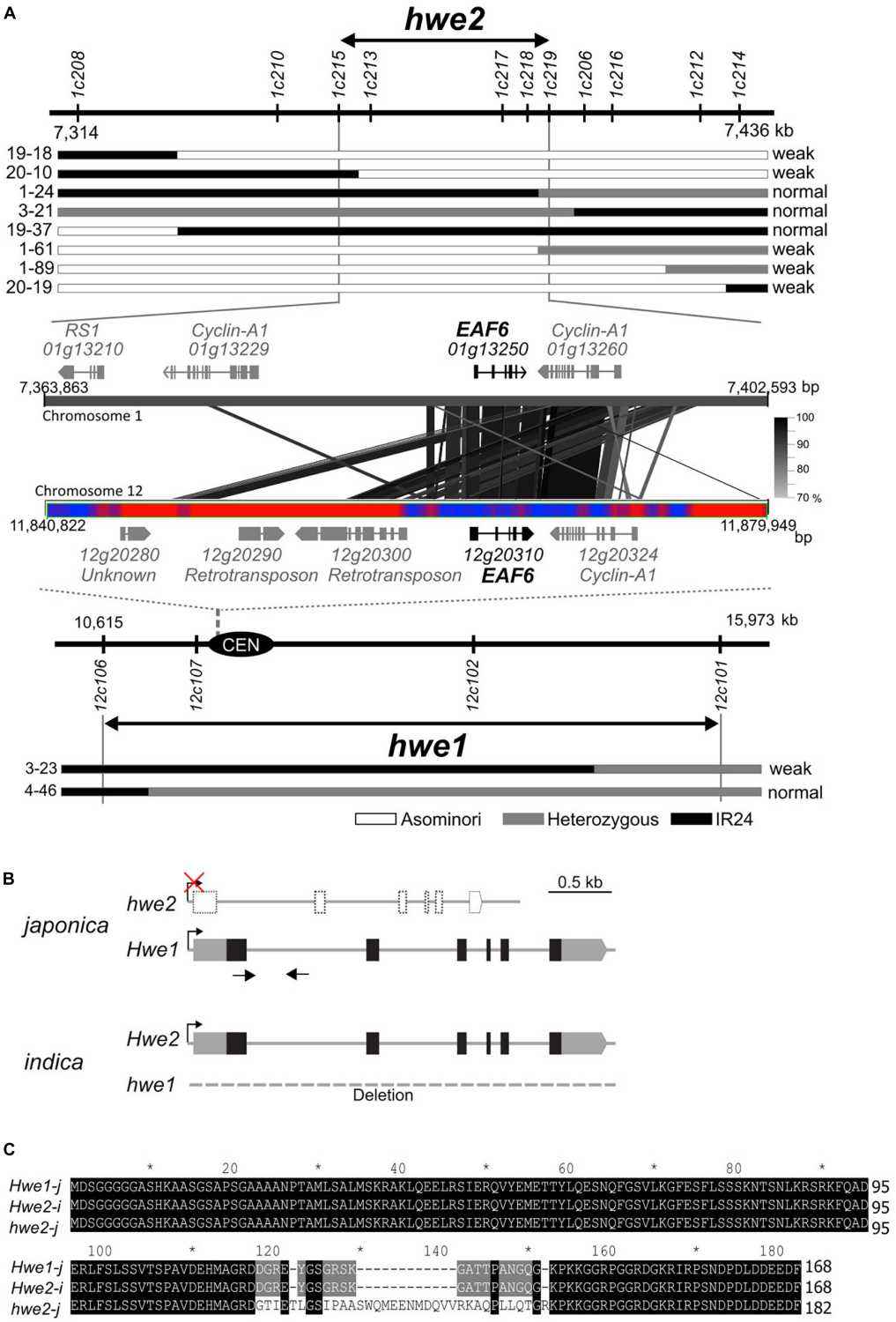

FIGURE 2 | Identification of causal genes for *hwe1* and *hwe2*. **(A)** Map locations of *hwe2* (*upper*) and *hwe1* (*lower*), and MSU7.0 gene annotation in candidate regions. The prefix "LOC_" in locus ID is omitted for convenience. The most informative recombinant genotypes are shown along with the maps. Similarity of DNA sequences between chromosomes 1 and 12 is shown as connected with black and gray lines. The bar with red and blue heat map represents the abundance of repetitive DNA sequences around the centromere (CEN) of chromosome 12. Red represents the repeat region. **(B)** Gene structure of *HWE1* and *HWE2* of *japonica* (Asominori) and *indica* (93-11). Gray and black boxes denote untranslated region (UTR) and coding sequence (CDS) of rice *OsEAF6*, respectively. The CDS sequences of *OsEAF6* were identical between *japonica* and *indica*. **(C)** Multiple alignments of predicted OsEAF6 protein sequences of *japonica* (*Hwe1-j* and *hwe2-j*) and *indica* (*Hwe2-i*).

expression (**Figure 3B**). *OsEAF6* was not expressed in the weak plants, whereas co-introgression with the IR24 segment around the *HWE1* and *HWE2* loci recovered the expression (**Figure 3C**), indicating that active *OsEAF6* mRNA was generated from the *japonica Hwe1-j* and *indica Hwe2-i* alleles. The absence of the mRNA signal from *LOC_Os01g13250* with the eaf6co primer indicated that *japonica hwe2-j* is an *OsEAF6* pseudogene (**Figure 3C**). Thus, these results support that defects were present in functional *OsEAF6* on the recessive *hwe2-j* and *hwe1-i* alleles and indicate that the lack of *OsEAF6* on the *hwe1* and *hwe2* alleles induces hybrid breakdown.

Tissue and Subcellular Localization of OsEAF6

In higher plants, EAF6 proteins are 144–170 amino acids long and highly conserved throughout herbaceous and woody plants (**Supplementary Figure 8** and **Supplementary Table 2**). Their sequences are partially conserved with those of yeast and animal species. Yeast (*Saccharomyces cerevisiae*) EAF6 is a small protein (113 amino acids) and subunit of the NuA4 HAT complex that is involved in transcriptional regulation through nuclear H4 acetylation (Mitchell et al., 2008). To gain insight into the function of OsEAF6, we examined the expression of *OsEAF6* in transgenic rice plants with *ProOsEAF6:GUS* (**Figure 3C**). *GUS* expression was observed in vegetative organs, including the leaves and stems, and in developing spikelets (**Figures 3D–G**). This result is consistent with those of RT-PCR analysis. Subcellular localization analysis using rice Oc cells showed that OsEAF6 protein was present predominantly in the nucleus and, to a lesser extent, in the cytoplasm, whereas the control mCherry plasmid was detectable throughout the cell (**Figures 3H,I**). This result indicates that OsEAF6 functions in the nucleus.

Evolution of *EAF6* in *Oryza* Species

In the rice genome sequencing project, 450 *O. rufipogon* accessions and 3,000 cultivars were sequenced using next-generation sequencing techniques (Huang et al., 2012; Wang W. et al., 2018). Based on published genome sequence data, we investigated the distribution of *hwe1* and *hwe2* alleles in cultivars and their wild relative *O. rufipogon*. The allelic diversity of duplicated *hwe* loci was examined based on eight SNPs at the 5′ and 3′ terminal regions of *OsEAF6*, which can discriminate between alleles in *HWE1* and *HWE2*. The Nipponbare *Hwe1* allele (called Nip-type) is GGAA-ATTT and is common among the *O. sativa-O. rufipogon* complex. The 93-11 *hwe1* null allele (9311-type) appeared to be distributed in the *indica* ecotype (255 accessions) but was minor in *O. rufipogon* (**Figures 4A,B** and **Supplementary Table 5**). We characterized the Nipponbare *hwe2* null allele as TCGC-ATTT (Nip-type) and 9311 *Hwe2* allele as GGAA-GGCC. Most *japonica* subspecies varieties (99.6%, 250/251) and 30–44% of *O. rufipogon* Or-I and Or-III ecotypes carried the Nip-type *hwe2* allele. Since Or-III has been reported as a progenitor of *japonica*, *hwe2* of *japonica* rice may have originated from *Or-III*. Some *O. rufipogon* accessions and *O. sativa* ssp. *indica "aus"* ecotype had two copies of functional *OsEAF6* on chromosomes 1 and 12 (**Supplementary Figure 7**).

To determine the origin and timing of *EAF6* duplication, *EAF6* homologs were investigated in wild *Oryza* species. We first analyzed the AA genome species most closely related to *O. sativa*. A BLAST search of the LOC_Os12g20310 CDS showed hits on chromosome 1 of other AA genome species (**Supplementary Figure 9**). Then, we investigated a local synteny pattern around the *EAF6* locus across distantly related *Oryza* species and five grass species. *Oryza* BB and FF genome species and other grass species (*Leersia*, *Brachypodium*, and *Setaria*) showed large blocks of homologous synteny around the *HWE2* region of Nipponbare chromosome 1 but lacked the whole sequence of *EAF6* on this syntenic block. Instead, these species contained a single copy of *EAF6* on other syntenic chromosomes, which were syntenic to *O. sativa* chromosome 2 (**Figure 4C**). The closest genus *Leersia*, which carried two copies of *EAF6*, showed that the local gene order around the *EAF6* locus on *Leersia* chromosome 5 did not differ among other species, indicating independent duplication events between rice and *Leersia*. The gene order and orientation around the *Cyclin-A1* locus (*LOC_Os01g13260*) were conserved on chromosome 1 of *O. punctata*, *O. brachyantha*, and *Leersia* as a single copy segment. Based on the alignment of other AA genome species, this result suggests that *EAF6* was initially transposed to chromosome 1 from chromosome 2 in the AA genome progenitor, followed by segmental duplication to chromosome 12 (**Supplementary Figure 10**).

DISCUSSION

We demonstrated that hybrid breakdown is caused by *HWE1/2* encoding a rice homolog of the NuA4 HAT complex subunit protein EAF6. The NuA4 HAT complex is an essential transcriptional coactivator involved in gene regulation, cellular processes, and DNA double-strand break repair in eukaryotes. Yeast EAF6 interacts with another catalytic subunit protein, Esa1, *via* Yng2. The functional role of EAF6 protein in plants remains unclear. We found that the loss of the OsEAF6 protein exerted deleterious pleiotropic effects on both vegetative growth and reproductive development in rice. Particularly, it has a broad impact on reproductive development, ranging from inflorescence development to gametogenesis. During the preparation of this study, Zhou et al. (2022) found that the *Arabidopsis eaf6* mutant shows growth inhibition and leaf yellowing and that the NuA4 complex is involved in transcriptional activation, specifically in light-responsive genes. Another research group reported that the loss of *Arabidopsis* Esa1-associated factor 1 (EAF1) inhibited growth and chloroplast development (Bieluszewski et al., 2022). These findings are consistent with our phenotypic observations, such as growth inhibition with reduced chlorophyll content during the vegetative phase in the double homozygote *hwe1/2*. It is suggested that such deleterious pleiotropic phenotypes occurred due to disorders of the universal chromatin state and transcriptional regulation caused by the loss of OsEAF6 protein.

Hypothetical Evolutionary History

Extensive genome sequencing and comparative studies revealed conserved microsynteny (gene order patterns) across different

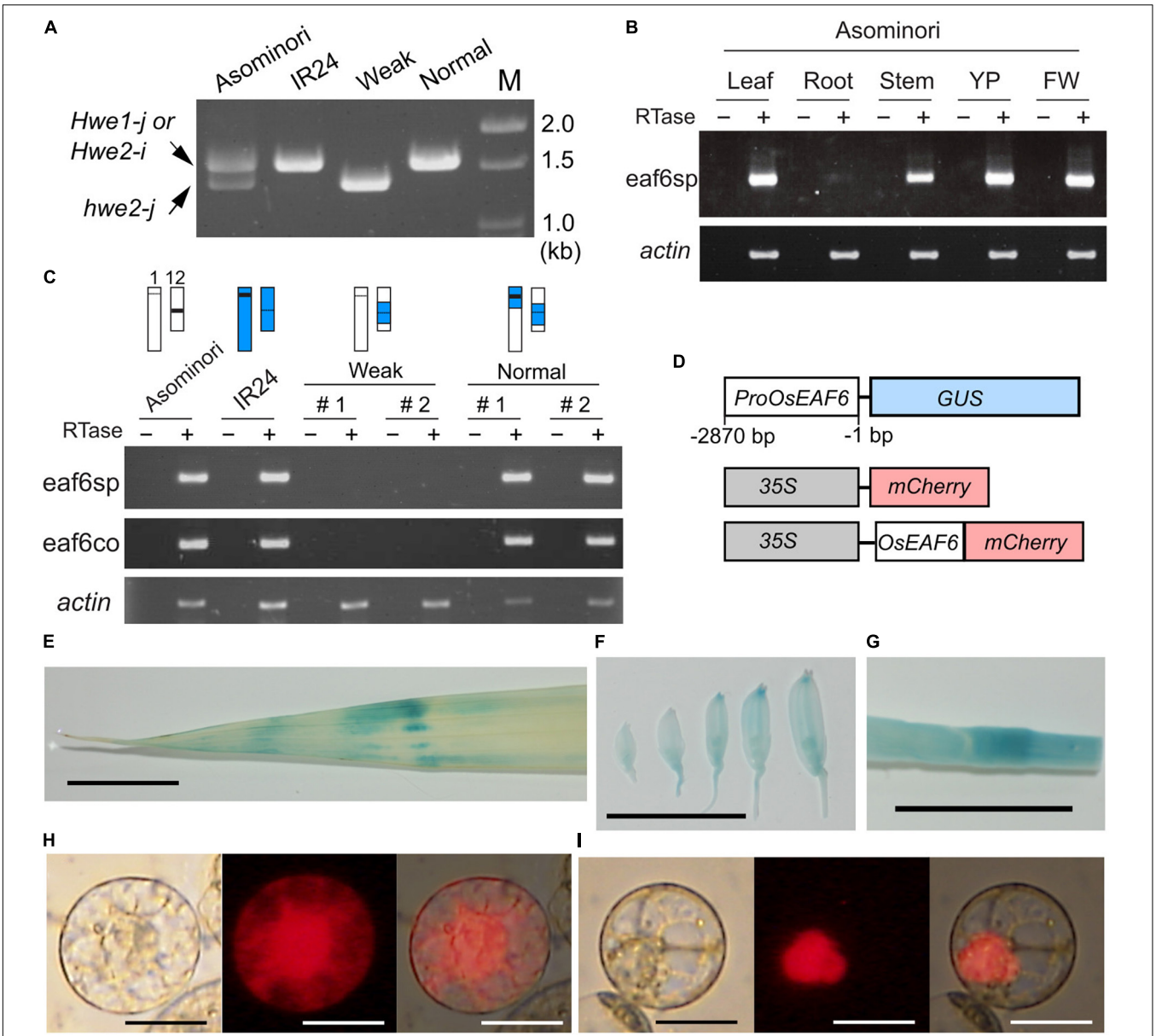

FIGURE 3 | Expression pattern and cellular localization of the OsEAF6. **(A)** Amplified genomic DNA product obtained with the primer eaf6co. The expected sizes of DNA bands were 1,557 bp from *Hwe1-j and Hwe2-i* and 1,352 bp from *hwe2-j.* The primer eaf6co was designed for the identical sequence region between *Hwe1-j* (functional gene) and *hwe2-j* (pseudogene). It is noted that the eaf6co primer could amplify the product from genomic DNA but not from mRNA, as shown in **(C)**. **(B)** mRNA expression of *OsEAF6* in rice tissues detected using RT-PCR analysis. eaf6sp is a primer set specific to functional *OsEAF6* [*japonica Hwe1* (*Hwe1-j*) allele and *indica Hwe2* (*Hwe2-i*) allele] but not the pseudogene on the *hwe2-j* allele. YP, young panicle; FW, flowering panicles **(C)** mRNA expression of *OsEAF6* in the NILs and parents with the two different primer sets by RT-PCR analysis. Genotype of each sample for chromosomes 1 and 12 is shown on the upper part. White and blue bars represent Asominori and IR24 chromosomes, respectively. Two individuals for the weak and normal plants (#1 and 2) are shown. **(D)** Diagram of constructs designed for expression analysis of *OsEAF6*. **(E–G)** GUS expression analysis of *OsEAF6* by using *ProOsEAF6:GUS* expresser in leaf **(E)**, developing flowers **(F)**, and stem **(G)** tissues. **(H,I)** Subcellular localization of OsEAF6 protein fused with mCherry **(I)** and mCherry alone **(H)** in rice protoplast cells. Transient expression of each fluorescent construct was observed. Left, DIC image; middle, mCherry channel image; right, merged image. Scale bar = 1.0 cm in **(E–G)**, and 20 μm in **(H–I)**.

cereal species (Jaiswal et al., 2006). Similar to previously reported microsynteny, a conserved gene order around *EAF6* was observed among the monocot crop species, including the wild rice relatives *Leersia* and sorghum (**Figure 4C**). However, the chromosome position of this synteny block containing *EAF6* was not chromosome 12 or chromosome 1, but rather chromosome 2 in other wild rice species (*O. punctata* and *O. brachyantha*). We hypothesized the evolutionary history of *EAF6* in *Oryza* genomes as follows: (1) *EAF6* has resided on chromosome 2 of primitive *Oryza* species; (2) *EAF6* was transposed to chromosome 1 in an early AA genome progenitor; (3) a segmental duplication occurred and was positioned on

B

Type	SNP sites Chr. 1 1 2 3 4	SNP sites Chr. 1 5 6 7 8	SNP sites Chr. 12 1 2 3 4	SNP sites Chr. 12 5 6 7 8	*O. rufipogon* Or-I	Or-II	Or-III	*O. sativa* *japonica*	*javanica*	*aroma*	*indica*	*aus*	Total
9311	G	G A A	G	G C C	Deletion		0	4	0	1	0	0	254	0	259
NIP	T	C G C	A	T T T	G	G A A	47	2	32	250	355	56	5	1	748
KAS	T	C G A	A	T T T	-	- - -	28	1	8	0	1	5	69	173	285
PA1	G	G A A	A	T T T	-	- - -	6	22	55	0	0	0	0	0	83
JAV	T	C A C	A	T T T	-	- - -	21	0	0	0	46	0	0	0	67
OTH	X	X X X	X	X X X	-	- - -	5	8	12	0	0	0	0	0	25
					Subtotal		107	37	107	251	402	61	328	174	1467
					Total				251					1216	

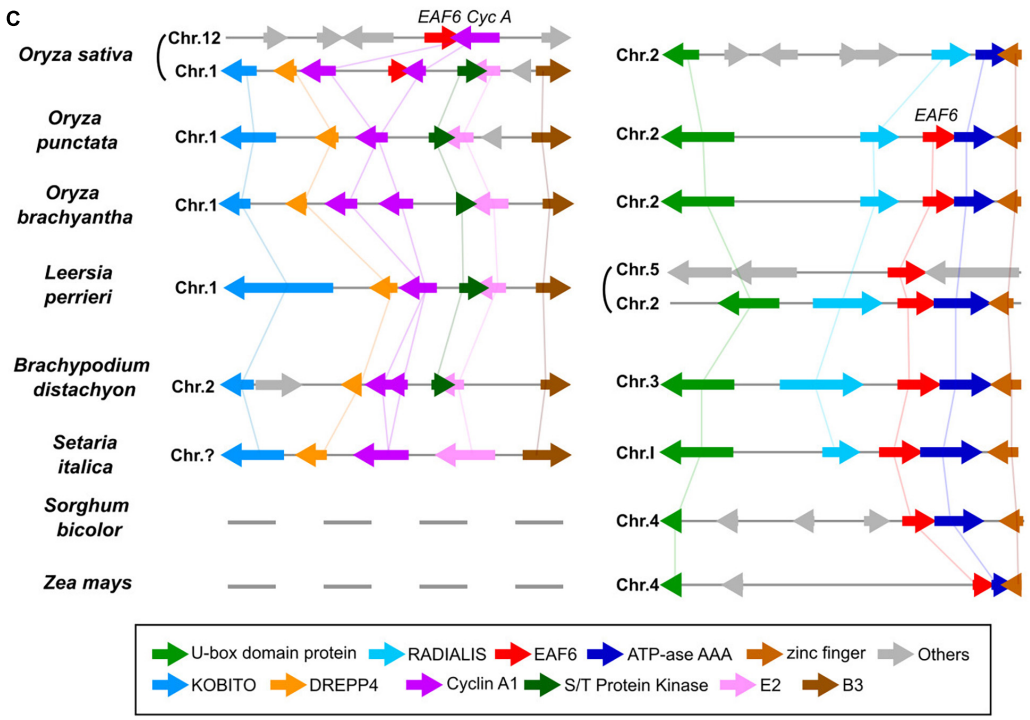

FIGURE 4 | Allele distribution of *OsEAF6* and syntenic gene analysis of *OsEAF6* region. **(A)** Distribution of *hwe1* and *hwe2* alleles in *Oryza sativa–Oryza rufipogon* complex. Positions of eight SNPs (1–8) on the 5′ and 3′ regions of *OsEAF6*. Gray and black boxes denote UTR and CDS of rice *OsEAF6*, respectively. SNP site 1, -309 bp on chr.1 and -304 bp on chr.12; site 2, -299 bp on chr.1 and -294 bp on chr.12; site 3, -205 bp on chr.1 and -206 bp on chr.12; site 4, -157 bp on chr.1 and -158 bp on chr.12; site 5, +2,133 bp on chr.1 and +2,527 bp on chr.12; site 6, +2,494 bp on chr.1 and +2,890 bp on chr.12; site 7, +2,971 bp on chr.1, and site 8, +3,153 bp on chr.12 and +3,555 bp on chr.12. The positions of SNPs (bp) were based on the Nipponbare sequence [+1 refers to **(A)** in the ATG start codon]. **(B)** Distribution of six haplotypes of duplicated *OsEAF6* in *O. rufipogon* and *O. sativa*. We evaluated 251 accessions for *O. rufipogon* and 1,216 accessions for *O. sativa*. The six haplotypes based on eight SNP sites consist of 93-11-type (9311), Nipponbare-type (NIP), Kasalath-type (KAS), PA1-type (PA1), *javanica* type (JAV), and other type (OTH). Hyphens represent identical SNPs with the Nipponbare SNP. **(C)** Conserved syntenic block harboring *EAF6* and adjacent *Cyclin-A1* loci. Gene annotation of the region surrounding the *EAF6* (*right*) and *Cyclin-A1* (*left*) loci in the cereal species. Homologous genes are shown in the same color connected by straight lines.

chromosome 12 in a subpopulation of *O. rufipogon*; and (4) one copy of the gene was lost in a progenitor population of *O. sativa* ssp. *japonica* (**Supplementary Figure 10**). This hypothetical scenario was based on the chromosome synteny and distribution of SNPs discriminating the two *OsEAF6* copies in the cultivars and their close relatives (**Figures 4A,C**). Transposition to chromosome 1 was considered for the following reasons. First, *Cyclin-A1* is not found on the corresponding region

of chromosome 12 in primitive *Oryza* species. Second, the *EAF6* CDS was localized on chromosome 1 according to recent next-generation sequencing analyses of other AA genome species (**Supplementary Figure 7**). Thus, in evolutionary history, *OsEAF6/HWE1* was a copy of *OsEAF6/HWE2* following the transposition from chromosome 2. Other grass species retain microsynteny around the *EAF6* positions. The mechanisms of transposition and duplication of *OsEAF6* remain unclear. Despite the positional differences in *EAF6* in grass species, the protein sequence of EAF6 protein is largely conserved among plant species (**Supplementary Figure 8**), suggesting that it has an essential function in plant development. Therefore, duplicated *EAF6* in other plant genomes may function as a reproductive isolation system.

Functional Role as the Reproductive Isolation System

In some animal studies, DNA-binding proteins, such as OdsH, PRDM9, and Zhr, were identified as causal molecules for hybrid sterility (Maheshwari and Barbash, 2011). These factors are likely associated with the dysfunction of chromatin remodeling in heterozygous hybrid progenies. Thus, abnormal chromatin formation during meiotic cell division in hybrids is a common factor responsible for reproductive isolation. From the perspective of the reproductive isolation mechanism, hybrid breakdown by *hwe1/2* occurred due to the loss of gene function and differed from the disharmonious interactions in the animal cases mentioned above, although the mechanism of action targeting nucleosomes is similar. Since no remarkable changes were observed in the heterozygous state, we did not characterize the detailed phenotype of heterozygous plants. However, heterozygous plants for each single locus of *HWE1* and *HWE2* (i.e., *Hwe1/hwe1 hwe2/hwe2* and *hwe1/hwe1 Hwe2/hwe2*) induced reduced transmission of the recessive alleles (*hwe1* and *hwe2*) in the selfed progeny (Kubo and Yoshimura, 2002). Thus, *hwe1/2* strongly impacts the elimination of the specific genotype around these genes in the hybrid population. Additionally, OsEAF6 may be involved in the haplotype gamete phase in rice. According to previous microarray data and laser capture microdissection of male and female gametes (Hobo et al., 2008; Kubo et al., 2013), *OsEAF6* was substantially and constantly expressed in haploid organs, such as microspores and megaspores (**Supplementary Figure 11**). Furthermore, the involvement of NuA4 in gametogenesis has been previously reported in *Arabidopsis* (Latrasse et al., 2008). Therefore, we believe that OsEAF6 may regulate histone acetylation and transcription levels throughout the rice life-cycle including the diploid and haploid phases. Further studies are required to determine the functions of OsEAF6 as a subunit of the HAT complex in various developmental stages and tissues. This study demonstrated the involvement of EAF6 in plant development and reproductive isolation. These findings will provide a helpful clue to transcriptional regulation by histone acetylation in plant development and also aid to develop an efficient breeding program to overcome reproductive isolation.

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

TK, AY, and NK conceived and designed the experiments. TK performed the experiments, analyzed the data, and wrote the study, with input from AY and NK. All authors read and approved the final manuscript.

FUNDING

This study was partly supported by the JSPS KAKENHI Grant Numbers 18075009 (to TK and NK) and 20K05985 (to TK).

ACKNOWLEDGMENTS

We thank T. Makino, Y. Gonohe, and H. Kondo (National Institute of Genetics, Mishima, Japan) for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.866404/full#supplementary-material

REFERENCES

Alcazar, R., Garcia, A. V., Kronholm, I., de Meaux, J., Koornneef, M., Parker, J. E., et al. (2010). Natural variation at Strubbelig receptor kinase 3 drives immune-triggered incompatibilities between *Arabidopsis thaliana* accessions. *Nat. Genet.* 42, 1135–1139. doi: 10.1038/ng.704

Bieluszewski, T., Galganski, L., Sura, W., Bieluszewska, A., Abram, M., Ludwikow, A., et al. (2015). AtEAF1 is a potential platform protein for *Arabidopsis* NuA4 acetyltransferase complex. *BMC Plant Biol.* 15:75. doi: 10.1186/s12870-015-0461-1

Bieluszewski, T., Sura, W., Dziegielewski, W., Bieluszewska, A., Lachance, C., Kabza, M., et al. (2022). NuA4 and H2A.Z control environmental responses and autotrophic growth in *Arabidopsis*. *Nat. Commun.* 13:277. doi: 10.1038/s41467-021-27882-5

Bird, A. W., Yu, D. Y., Pray-Grant, M. G., Qiu, Q., Harmon, K. E., Megee, P. C., et al. (2002). Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. *Nature* 419, 411–415. doi: 10.1038/nature01035

Bomblies, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J. L., et al. (2007). Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol.* 5:e236. doi: 10.1371/journal.pbio.0050236

Boudreault, A. A., Cronier, D., Selleck, W., Lacoste, N., Utley, R. T., Allard, S., et al. (2003). Yeast enhancer of polycomb defines global Esa1-dependent acetylation of chromatin. *Genes Dev.* 17, 1415–1428. doi: 10.1101/GAD.1056603

Chae, E., Bomblies, K., Kim, S.-T., Karelina, D., Zaidem, M., Ossowski, S., et al. (2014). Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. *Cell* 159, 1341–1351. doi: 10.1016/j.cell.2014.10.049

Chen, C., Chen, H., Lin, Y. S., Shen, J. B., Shan, J. X., Qi, P., et al. (2014). A two-locus interaction causes interspecific hybrid weakness in rice. *Nat. Commun.* 5:3357. doi: 10.1038/ncomms4357

Clarke, A. S., Lowell, J. E., Jacobson, S. J., and Pillus, L. (1999). Esa1p is an essential histone acetyltransferase required for cell cycle progression. *Mol. Cell. Biol.* 19, 2515–2526. doi: 10.1128/MCB.19.4.2515

Doyon, Y., Selleck, W., Lane, W. S., Tan, S., and Côté, J. (2004). Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Mol. Cell. Biol.* 24, 1884–1896. doi: 10.1128/MCB.24.5.1884-1896.2004

Fuse, A., Sasaki, T., and Yano, M. (2001). Ti-plasmid vector useful for functional analysis of rice genes. *Plant Biotehcnol.* 18, 219–222.

Hannah, M. A., Krämer, K. M., Geffroy, V., Kopka, J., Blair, M. W., Erban, A., et al. (2007). Hybrid weakness controlled by the dosage-dependent lethal (DL) gene system in common bean (*Phaseolus vulgaris*) is caused by a shoot-derived inhibitory signal leading to salicylic acid-associated root death. *New Phytol.* 176, 537–549. doi: 10.1111/j.1469-8137.2007.02215.x

Hiei, Y., Ohta, S., Komari, T., and Kumashiro, T. (1994). Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J.* 6, 271–282. doi: 10.1046/j.1365-313x.1994.6020271.x

Hobo, T., Suwabe, K., Aya, K., Suzuki, G., Yano, K., Ishimizu, T., et al. (2008). Various spatiotemporal expression profiles of anther-expressed genes in rice. *Plant Cell Physiol.* 49, 1417–1428. doi: 10.1093/PCP/PCN128

Huang, X., Kurata, N., Wei, X., Wang, Z. X., Wang, A., Zhao, Q., et al. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490, 497–501. doi: 10.1038/nature11532

Ito, Y., and Kurata, N. (2008). Disruption of KNOX gene suppression in leaf by introducing its cDNA in rice. *Plant Sci.* 174, 357–365. doi: 10.1016/j.plantsci.2007.12.007

Jaiswal, P., Ni, J., Yap, I., Ware, D., Spooner, W., Youens-Clark, K., et al. (2006). Gramene: a bird's eye view of cereal genomes. *Nucleic Acids Res.* 34, D717–D723. doi: 10.1093/NAR/GKJ154

Jeuken, M. J. W., Zhang, N. W., McHale, L. K., Pelgrom, K., Den Boer, E., Lindhout, P., et al. (2009). Rin4 causes hybrid necrosis and race-specific resistance in an interspecific lettuce hybrid. *Plant Cell* 21, 3368–3378. doi: 10.1105/TPC.109.070334

Kindiger, B., and Beckett, J. B. (1985). A hematoxylin staining procedure for maize pollen grain chromosomes. *Biotech. Histochem.* 60, 265–269. doi: 10.3109/10520298509113923

Kubo, T., Fujita, M., Takahashi, H., Nakazono, M., Tsutsumi, N., and Kurata, N. (2013). Transcriptome analysis of developing ovules in rice isolated by laser microdissection. *Plant Cell Physiol.* 54, 750–765. doi: 10.1093/PCP/PCT029

Kubo, T., Takashi, T., Ashikari, M., Yoshimura, A., and Kurata, N. (2016). Two tightly linked genes at the hsa1 locus cause both F_1 and F_2 hybrid sterility in rice. *Mol. Plant* 9, 221–232. doi: 10.1016/j.molp.2015.09.014

Kubo, T., and Yoshimura, A. (2002). Genetic basis of hybrid breakdown in a Japonica/Indica cross of rice, *Oryza sativa* L. *Theor. Appl. Genet.* 105, 906–911. doi: 10.1007/S00122-002-1059-1

Kubo, T., Yoshimura, A., and Kurata, N. (2011). Hybrid male sterility in rice is due to epistatic interactions with a pollen killer locus. *Genetics* 189, 1083–1092. doi: 10.1534/genetics.111.132035

Lafon, A., Chang, C. S., Scott, E. M., Jacobson, S. J., and Pillus, L. (2007). MYST opportunities for growth control: yeast genes illuminate human cancer gene functions. *Oncogene* 26, 5373–5384. doi: 10.1038/sj.onc.1210606

Latrasse, D., Benhamed, M., Henry, Y., Domenichini, S., Kim, W., Zhou, D. X., et al. (2008). The MYST histone acetyltransferases are essential for gametophyte development in *Arabidopsis*. *BMC Plant Biol.* 8:121. doi: 10.1186/1471-2229-8-121

Maheshwari, S., and Barbash, D. A. (2011). The genetics of hybrid incompatibilities. *Annu. Rev. Genet.* 45, 331–355. doi: 10.1146/annurev-genet-110410-132514

Mitchell, L., Lambert, J.-P., Gerdes, M., Al-Madhoun, A. S., Skerjanc, I. S., Figeys, D., et al. (2008). Functional dissection of the NuA4 histone acetyltransferase reveals its role as a genetic hub and that Eaf1 Is essential for complex integrity. *Mol. Cell. Biol.* 28, 2244–2256. doi: 10.1128/mcb.01653-07

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4325. doi: 10.1093/nar/8.19.4321

Nishimura, A., Aichi, I., and Matsuoka, M. (2006). A protocol for *Agrobacterium*-mediated transformation in rice. *Nat. Protoc.* 1, 2796–2802. doi: 10.1038/nprot.2006.469

Ohba, R., Steger, D. J., Brownell, J. E., Mizzen, C. A., Cook, R. G., Côté, J., et al. (1999). A novel H2A/H4 nucleosomal histone acetyltransferase in *Tetrahymena thermophila*. *Mol. Cell. Biol.* 19, 2061–2068. doi: 10.1128/MCB.19.3.2061

Ohtsubo, Y., Ikeda-Ohtsubo, W., Nagata, Y., and Tsuda, M. (2008). GenomeMatcher: a graphical user interface for DNA sequence comparison. *BMC Bioinformatics* 9:376. doi: 10.1186/1471-2105-9-376

Panagopoulos, I., Micci, F., Thorsen, J., Gorunova, L., Eibak, A. M., Bjerkehagen, B., et al. (2012). Novel fusion of MYST/Esa1-associated factor 6 and PHF1 in endometrial stromal sarcoma. *PLoS One* 7:e39354. doi: 10.1371/journal.pone.0039354

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/OXFORDJOURNALS.MOLBEV.A040454

Sakata, K., Nagamura, Y., Numa, H., Antonio, B. A., Nagasaki, H., Idonuma, A., et al. (2002). RiceGAAS: an automated annotation system and database for rice genome sequence. *Nucleic Acids Res.* 30, 98–102. doi: 10.1093/nar/30.1.98

Shen, J., Fu, J., Ma, J., Wang, X., Gao, C., Zhuang, C. X., et al. (2014). Isolation, culture, and transient transformation of plant protoplasts. *Curr. Protoc. Cell Biol.* 63, 2.8.1–2.8.17. doi: 10.1002/0471143030.CB0208S63

Stebbins, G. L. (1958). The inviability, weakness, and sterility of interspecific hybrids. *Adv. Genet.* 9, 147–215. doi: 10.1016/S0065-2660(08)60162-5

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30:2725. doi: 10.1093/MOLBEV/MST197

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, 43–49. doi: 10.1038/s41586-018-0063-9

Wang, X., Ahmad, S., Zhang, Z., Côté, J., and Cai, G. (2018). Architecture of the *Saccharomyces cerevisiae* NuA4/TIP60 complex. *Nat. Commun.* 9:1147. doi: 10.1038/s41467-018-03504-5

Yamamoto, E., Takashi, T., Morinaka, Y., Lin, S., Wu, J., Matsumoto, T., et al. (2010). Gain of deleterious function causes an autoimmune response and Bateson-Dobzhansky-Muller incompatibility in rice. *Mol. Genet. Genomics* 283, 305–315. doi: 10.1007/s00438-010-0514-y

Zacharaki, V., Benhamed, M., Poulios, S., Latrasse, D., Papoutsoglou, P., Delarue, M., et al. (2012). The *Arabidopsis* ortholog of the YEATS domain containing protein YAF9a regulates flowering by controlling H4 acetylation levels at the FLC locus. *Plant Sci.* 196, 44–52. doi: 10.1016/j.plantsci.2012.07.010

Zhou, J., Su, X., Zheng, S., Wu, C., Su, Y., Jiang, Z., et al. (2022). The *Arabidopsis* NuA4 histone acetyltransferase complex is required for chlorophyll biosynthesis and photosynthesis. *J. Integr. Plant Biol.* doi: 10.1111/JIPB.13227

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 in Plant Science

ORIGINAL RESEARCH
published: 09 March 2022
doi: 10.3389/fpls.2022.790221

Mapping of QTLs for Yield Traits Using $F_{2:3:4}$ Populations Derived From Two Alien Introgression Lines Reveals *qTGW8.1* as a Consistent QTL for Grain Weight From *Oryza nivara*

Kavitha Beerelli[1,2], Divya Balakrishnan[1], Krishnam Raju Addanki[1,2], Malathi Surapaneni[1], Venkateswara Rao Yadavalli[1] and Sarla Neelamraju[1]**

[1] National Professor Project, ICAR-Indian Institute of Rice Research, Hyderabad, India, [2] Department of Biotechnology, Acharya Nagarjuna University, Guntur, India

OPEN ACCESS

Edited by:
Dayun Tao,
Yunnan Academy of Agricultural Sciences, China

Reviewed by:
Kishor Kumar,
Ramakrishna Mission Vivekananda Educational and Research Institute, India
Wricha Tyagi,
Central Agricultural University, India

***Correspondence:**
Divya Balakrishnan
dbiirr23@gmail.com
Sarla Neelamraju
sarla_neelamraju@yahoo.com

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

Received: 06 October 2021
Accepted: 09 February 2022
Published: 09 March 2022

Citation:
Beerelli K, Balakrishnan D, Addanki KR, Surapaneni M, Rao Yadavalli V and Neelamraju S (2022) Mapping of QTLs for Yield Traits Using $F_{2:3:4}$ Populations Derived From Two Alien Introgression Lines Reveals qTGW8.1 as a Consistent QTL for Grain Weight From Oryza nivara. Front. Plant Sci. 13:790221. doi: 10.3389/fpls.2022.790221

Wild introgressions play a crucial role in crop improvement by transferring important novel alleles and broadening allelic diversity of cultivated germplasm. In this study, two stable backcross alien introgression lines 166s and 14s derived from Swarn/*Oryza nivara* IRGC81848 were used as parents to generate populations to map quantitative trait loci (QTLs) for yield-related traits. Field evaluation of yield-related traits in F_2, F_3, and F_4 population was carried out in normal irrigated conditions during the wet season of 2015 and dry seasons of 2016 and 2018, respectively. Plant height, tiller number, productive tiller number, total dry matter, and harvest index showed a highly significant association to single plant yield in F_2, F_3, and F_4. In all, 21, 30, and 17 QTLs were identified in F_2, $F_{2:3}$, and $F_{2:4}$, respectively, for yield-related traits. QTLs *qPH6.1* with 12.54% phenotypic variance (PV) in F_2, *qPH1.1* with 13.01% PV, *qTN6.1* with 10.08% PV in $F_{2:3}$, and *qTGW6.1* with 15.19% PV in $F_{2:4}$ were identified as major effect QTLs. QTLs *qSPY4.1* and *qSPY6.1* were detected for grain yield in F_2 and $F_{2:3}$ with PV 8.5 and 6.7%, respectively. The trait enhancing alleles of QTLs *qSPY4.1*, *qSPY6.1, qPH1.1*, *qTGW6.1*, *qTGW8.1, qGN4.1,* and *qTDM5.1* were from *O. nivara*. QTLs of the yield contributing traits were found clustered in the same chromosomal region. *qTGW8.1* was identified in a 2.6 Mb region between RM3480 and RM3452 in all three generations with PV 6.1 to 9.8%. This stable and consistent *qTGW8.1* allele from *O. nivara* can be fine mapped for identification of causal genes. From this population, lines $C_2$12, $C_2$124, $C_2$128, and $C_2$143 were identified with significantly higher SPY and $C_2$103, $C_2$116, and $C_2$117 had consistently higher thousand-grain weight values than both the parents and Swarna across the generations and are useful in gene discovery for target traits and further crop improvement.

Keywords: *Oryza*, alien introgression lines, yield QTL, thousand grain weight, wild species

INTRODUCTION

Rice is one of three major food crops across the world especially in the most populated regions, and it provides up to 23% of calories for human consumption (Fisher et al., 2000). A total of 503.5 million tonnes (milled basis) of rice was consumed around the globe during 2017–18 (FAO et al., 2017). Rice breeders have major challenges in increasing the yield potential of the cultivars as there is stagnation due to narrow genetic diversity available in cultivated germplasm. Wild genetic material is a source of important alleles or genes for agronomic traits including yield. Most of the earlier studies using conventional plant breeding methods helped increase of the yield levels of rice by improving the related traits (Brondani et al., 2002). Wild rice species are more diverse in physiological, morphological, and agronomical characteristics than the existing cultivars. Because of the narrow genetic base of the cultivars, there is a need to transfer genes of desirable traits from wild to cultivated rice, and it is an essential strategy in pre-breeding. Wild species have beneficial alleles for yield improvement, but the expression of these alleles is frequently masked due to the presence of other detrimental loci. Yet, several yields enhancing quantitative trait loci (QTLs) have been mapped in the last 25 years from wild species of rice for genetical improvement (Swamy and Sarla, 2008; Gaikwad et al., 2021). The wild rice species, *Oryza nivara,* is the closest wild progenitor of cultivated rice *O. sativa* (Haritha et al., 2018). *O. nivara* accessions showed high genetic diversity in its gene pool with adaptability in different environments (Sarla et al., 2003; Juneja et al., 2006) and is a proven choice to improve the yield levels of cultivars (Vaughan et al., 2003, 2008; Swamy et al., 2014; Ma et al., 2016).

An advanced back cross method is a technique to introduce the favorable alleles from wild into cultivar background (Ma et al., 2016). There are many reports on mapping yield QTLs using wild species (Swamy and Sarla, 2008; Swamy et al., 2011, 2014; Wickneswari et al., 2012; Ma et al., 2016; Bhatia et al., 2018). Back cross inbred lines are useful for mapping QTLs and gene pyramiding (Bhatia et al., 2018; Wang et al., 2018). In addition, QTL pyramiding using introgression lines (ILs) is an effective method in molecular breeding for complex traits (Feng et al., 2018). Back cross introgression lines (BILs) derived from cultivar/wild crosses or alien introgression lines were known for improving yield traits along with different desirable traits, such as quality-related traits and biotic and abiotic stress resistance (Mahmoud et al., 2008; Brar and Singh, 2011; Swamy et al., 2012, 2014). BILs generated by repeated backcrosses are useful in restoring the pollen fertility and eliminating the undesirable trait effects on the cultivar background (Swamy and Sarla, 2008). Several studies used BC_2F_2 populations derived from interspecific crosses to map QTLs. However, when introgression lines that are genetically similar to each other but phenotypically different are used, e.g., near-isogenic lines, the power to map the phenotype increases considerably. Compared to BC_2F_2, these advanced ILs in BC_2F_8 do not simultaneously have the confounding effects of segregation of several QTLs all over the genome or the complex epistatic interactions in the genome. Therefore, two high-yielding stable, fine grain BC_2F_8 Ils 166s and 14s derived from Swarna/*O. nivara* cross, were chosen to map QTLs for yield-related traits. These BC_2F_8 are derived from 166s (IET21938) and 14s (IET2274) which were identified as two fine-grain introgression lines earlier in BC_2F_2 (Swamy et al., 2012) and were shown in BC_2F_{6-7} to be high-yielding in multilocation trials (Haritha et al., 2018) and are also salt tolerant (Ganeshan et al., 2016). The selected parental lines, *viz.,* 166s and 14s with 75.8 and 77.8% recurrent parent genome, respectively, were part of a library of Swarna chromosome segment substitution lines (CSSLs) having chromosome segment substitutions from *O. nivara* (Balakrishnan et al., 2016; Surapaneni et al., 2017) and showed stability in yield levels over the generations. Among several traits to differentiate the two lines, 166s has a larger number of grains, and 14s has higher thousand-grain weight. Both of which are major yield contributing traits (**Supplementary Table 1**).

Previously, Pang et al. (2017) and Feng et al. (2018) reported QTLs in the population, generated by crossing Ils developed from local rice varieties. Similarly, QTLs were reported in the F_2 population derived from a cross between a back cross inbred line (*japonica* × *indica*) and a *japonica* cultivar Z550 (Wang et al., 2018). Various studies reported QTLs for yield traits in rice using primary mapping populations like $F_{2:3}$ (Sabouri et al., 2009; Zhang et al., 2010; Kim et al., 2014; Tian et al., 2015; Biswas et al., 2017; Kumar et al., 2019) and $F_{2:4}$ (Rabiei et al., 2015; Verma et al., 2017; Jeon et al., 2018; Kumar et al., 2019). Crossing of two wild-derived back cross inbred lines generate hybrid lines which can closely resemble the parental lines (Wang et al., 2018) with combination of improved traits. We selected these two high yielding BILs with contrasting yield contributing traits to develop $F_{2:3:4}$ populations with the objective of mapping consistent and precise QTLs for yield-related traits across three generations.

MATERIALS AND METHODS

Plant Material and Field Evaluation

Initially, a set of back cross introgression lines at BC_2F_2 were developed from a cross of Swarna × *O. nivara* (Swamy et al., 2014), and this material was advanced up to BC_2F_8 through single panicle selection of selected individual lines. These lines were screened for three consecutive seasons to study G × E interaction of yield and related traits, and two BC_2F_8 lines, *viz.,* 166s and 14s, were identified as most stable BILs for different yield-related traits (Balakrishnan et al., 2016; Kavitha et al., 2019). BIL 166s had high biomass, total dry matter, number of grains, seedling vigor, and high photosynthetic rate compared to parent Swarna and was also tolerant to aerobic and saline conditions based on our previous studies. BIL 14s had high single plant yield, bulk yield, per-day productivity, harvest index, 1,000 grain weight, and low unfilled grains. These BC_2F_8 BILs, 166s [IET27223] as female parent and 14s [IET 26772] as male parent, were taken as starting material for the experiment and F_1 was generated by crossing them (**Figure 1**). In *Rabi* 2015, F_1 plant was raised and selfed to generate F_2 mapping population. F_2 population was raised in *Kharif 2015* and forwarded to F_3 and F_4 populations in *Rabi* 2016 and *Rabi* 2018, respectively, at the Indian Institute of Rice

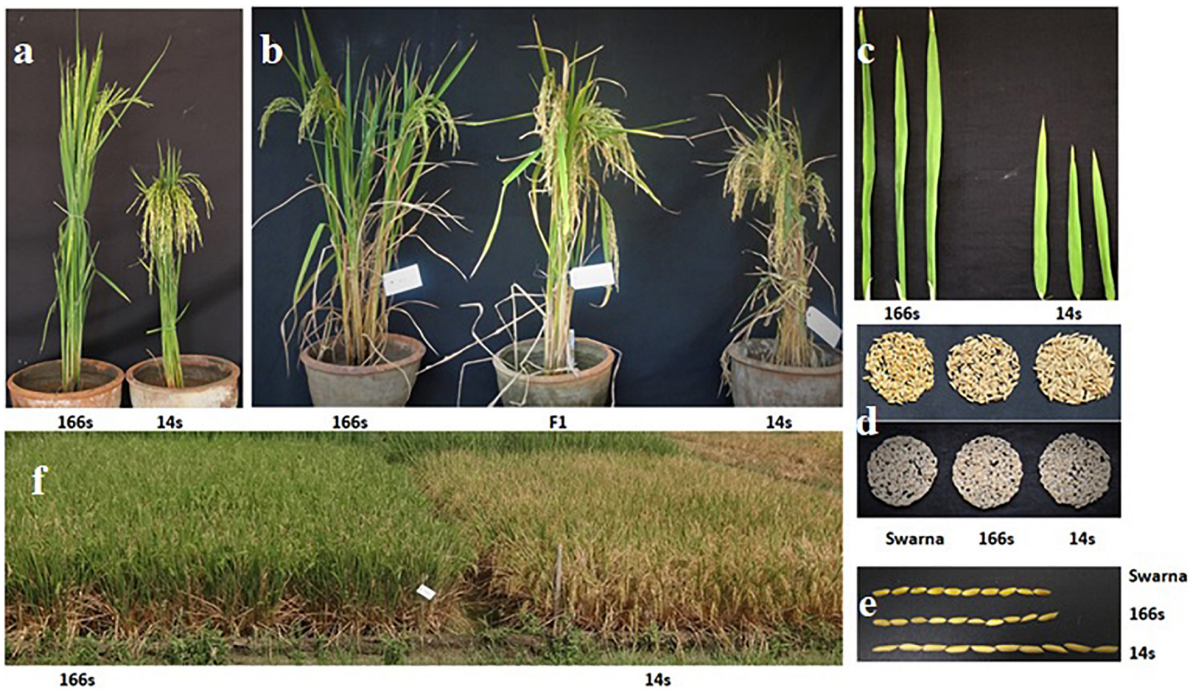

FIGURE 1 | Phenotypic characteristics of parental lines, plant, and grain traits of 166s, 14s, and Swarna. **(a)** Parental lines 166s and 14s, **(b)** 166s, 14s parents with F_1 (166s × 14s hybrid-middle one) plant, **(c)** Flag leaf length of 166s and 14s, **(d,e)** Swarna, 166s and 14s grains with husk and without husk and length of grains, **(f)** Field view of 166s and 14s.

Research (IIRR) field. The farm is located in Hyderabad, India at 17° 19′ N latitude and 78° 29′ E longitude.

Field evaluation of yield-related traits in the F_2 population was carried out in normal irrigated field conditions during *Kharif 2015* using the standard evaluation system of IRRI (SES, IRRI) for 19 phenotypic traits. These were days to initial flowering (DIF), plant height (PH), tiller number (TN), productive tiller number (PTN), flag leaf length (FLL), leaf length (LL), leaf width (LW), culm length (CL), single plant yield (SPY), biomass (BM), total dry matter (TDM), harvest index (HI), panicle length (PL), filled grains (FG), unfilled grains (UFG), grain number (GN), spikelet fertility (SPF), panicle weight (PW), and thousand-grain weight (TGW). In addition, 18 of these traits were also measured in F_3 population, viz., DIF, PH, TN, PTN, CL, SPY, BM, TDM, HI, PL, FG, UFG, GN, SPF, PW, TGW, PDP, and Bulk yield (BY), and 8 traits in F_4 population, viz., PH, TN, PTN, SPY, BM, TDM, HI, and TGW, in randomized with three replications.

DNA Extraction

Fresh leaf samples from 1-month old seedlings were collected from 174 F_2 plants and parents in *Kharif 2015*. DNA was extracted using cetyltrimethyl ammonium bromide (CTAB) mini prep method (Doyle and Doyle, 1990). In this method, 400 μl of CTAB was added to small pieces of leaf in a mini prep plate for grinding. Four hundred microliters of CTAB was added to it and was mixed as well in a 2 ml microcentrifuge tube. To the leaf sample, equal amounts of chloroform:isoamyl alcohol (24:1) was added in each tube while shaking vigorously until it was dissolved. Then tubes were kept for centrifugation at 15,000 rpm for 15 min, and the supernatant was removed gently into a fresh 1.5 ml microcentrifuge tube. An equal amount of isopropanol was added to this supernatant before it was mixed slowly and kept in a freezer at −20°C for about 10 min to 2 days for DNA pellet formation. Afterward, these tubes were centrifuged at 10,000 rpm for 10 min, and the supernatant was removed in this step very slowly to retain the pellet in the tube. Then, 100 μl of 70% of ethanol was added to the pellet and kept for centrifugation at 8,000 rpm for 5 min for pellet cleaning. This step was repeated again to get a clear pellet of DNA before this DNA pellet was dissolved in 100 μl of 1× TE buffer or distilled water for further use.

Genotyping

Parental polymorphism between 166s and 14s using 830 simple sequence repeats (SSRs) was conducted. Out of this, only 79 showed polymorphism. F_2 population of 174 plants was screened using these 79 primers of which only 64 were clearly segregated in the population. The remaining 15 primers only showed 14s kind of bands. Polymerase chain reaction (PCR) was carried out using 10 μl reaction mixture containing 3 μl of DNA sample (50 ng), 3.8 μl of millipore water, 0.1 μl of dNTPs, 1.2 μl of 25 mM $Mgcl_2$, 1 μl of 10× PCR buffer with Mg, 0.8 μl of SSR primer, and 0.1 μl of taq polymerase in each PCR plate. PCR was conducted with the initial denaturation step maintained for 5 min at 95°C. It was also conducted for each cycle with a denaturation temperature of 95°C for 30 s, annealing temperature of 55°C for 30 s, and an

extension temperature of 72°C for 30 s which was maintained for up to 35 cycles before a final extension of up to 7 min at 72°C. The final PCR product after 35 cycles was maintained at a temperature of 10°C before collecting from PCR. For agarose gel electrophoresis, 3% gels were prepared with wells to load the DNA samples, and electrophoresis was carried out at 180 volts in 0.5× Tris/Borate/EDTA (TBE) buffer. The band pattern was documented in UV-light in the gel documentation unit. Gel scoring of samples along with parents was conducted by noting 'A' (P1) for homozygous 166s allele, 'B' (P2) for homozygous 14s allele, and 'H'(P1P2) for heterozygous, i.e., presence of both the alleles.

Statistical Analyses and Quantitative Trait Loci Mapping

Analysis of variance (ANOVA) was performed using the statistical tool for agricultural research (STAR v2.0.1) software, and association between the traits was estimated with plant breeding tools (PB tools) (Ver. 1.4[1]) using Pearson's product-moment correlation method at the significant levels of $^{*}p = 0.05$–0.001 and $^{**}p \geq 0.001$. QTL mapping involves the construction of a linkage map and QTL analysis. Inclusive component interval mapping (ICIM) QTL mapping (IciMapping v4.1) integrated software (Wang, 2009[2]) was used for both linkage mapping and QTL analysis. Single marker analysis, interval mapping, and composite interval mapping were performed using this software using F_2 genotypic data and F_2, F_3 ($F_{2:3}$), and F_4 ($F_{2:4}$) phenotypic data for QTL mapping. Further genotypic dissection within a major QTL region for TGW was carried out in the extreme phenotypes using 5 low TGW (LTGW) and 5 high TGW (HTGW) F_2 lines with SSR markers within the detected QTL region. The co-segregating markers were then used to genotype 10 LTGW and 10 HGTW F_4 lines.

RESULTS

Phenotyping F_2, F_3, and F_4 Populations

Yield-related traits of populations were evaluated in Kharif (wet season) 2015, Rabi (dry season) 2016, and Rabi (dry season) 2018 for F2, F3, and F4, respectively, using standard evaluation system (SES) of Standard Evaluation System for Rice [SES] (2014). The yield-related traits in F_2 (19 traits), $F_{2:3}$ (18 traits), and in $F_{2:4}$ (8 traits) were measured and significant differences among the individuals or lines were observed in the three populations for each trait. The SPY of lines $C_2$12, $C_2$124, $C_2$128, $C_2$143, and C_2 162 showed a positively significant difference with Swarna and P_2 in F_3 (**Supplementary Tables 2, 3**) and showed higher yield in F_3 and F_4. **Supplementary Table 2** shows the cumulative number of lines that are significantly different than parents for each trait. The details of significantly different traits in individual lines as compared to parents are given in **Supplementary Table 3**. The descriptive statistical data for all these yield-related traits of F_2, F_3, and F_4 populations is given in **Table 1**, and frequency distribution is shown in **Supplementary Figures 1–3**, respectively. Box plots for 8 traits in F_2, F_3, and F_4 and 16 traits in F_2 and F_3 are shown in **Supplementary Figures 4, 5**, respectively.

Trait Correlation

Correlation analysis among all yield-related traits in F_2 and F_3 was evaluated using Pearson's product-moment correlation and probability values of $^{*}p = 0.05$–0.001 as significant and $^{**}p \geq 0.001$ as a highly significant correlation between traits. Results are shown in **Table 2**. Some traits consistently showed significant association in F_2, $F_{2:3}$, and $F_{2:4}$. The traits which showed highly significant positive correlations in F_2 were as follows: FLL with PH, CL, SPY, BM, HI, PL, FG, UFG, GN, PW, and PDP; LL with PH, TN, PTN, CL, SPY, BM, TDM, HI, PL, FG, UFG, GN, PW, TGW, PDP, and FLL; and LW with PH, TN, PTN, CL, SPY, BM, TDM, PL, FG, UFG, GN, PW, PDP, FLL, LL, and LW. A highly significant negative association was observed for FLL, LL, and LW with DIF.

In F_3, highly significant positive correlations were observed for BY with PH, CL, SPY, BM, TDM, HI, and PDP. PH, TN, PTN, CL, BM, TDM, HI, and PDP showed significant and positive association to single plant yield, and significant negative association was observed for TGW, SPF, and HI with UFG in both F_2 and $F_{2:3}$. In $F_{2:4}$, a highly significant positive association was observed between PTN and TN; SPY with PH, TN, and PTN; BM and TDM with both TN and PTN; TDM with SPY and BM; HI with SPY; and TGW with PH. There was also a highly significant negative association of HI with BM and TDM. A significant positive association was identified for PH with TDM and HI and TGW with SPY.

Traits PH, TN, PTN, TDM, and HI showed a highly significant positive association with single plant yield in all three generations of F_2, F_3, and F_4. TGW showed a highly significant association with SPY in F_2, and a significant association was observed between TGW and SPY in F_3 and F_4. The traits PH, TN, PTN, and TDM consistently showed a highly significant positive association to single plant yield in all three generations. Highly significant positive correlations were observed for the traits SPY with PH, TN, PTN, CL, BM, TDM, HI, and PDP in F_2 and $F_{2:3}$. The traits which showed a highly significant negative correlation in F_2 and $F_{2:3}$ were HI with DIF and UFG; SPF with DIF and UFG. A significant negative association between TGW with UFG was also observed, in addition to the highly significant negative association in both $F_{2:3}$ and $F_{2:4}$ between traits HI with BM. A significant positive association was observed between traits TGW with SPY.

Genotyping

The F_2 population consisting of 174 lines was genotyped using 79 polymorphic markers. Among these polymorphic SSR markers, segregation distortion was observed for nearly 20 markers. In this study, heterozygous bands ranged from 14.3 (RM1189) to 69.5% (RM430) in the whole population. The 166s (female P1) parental type bands ranged between 1.1% (RM202) and 47.1% (RM1189), while the 14s (male P2) parent type bands ranged from 9.7 (RM430) to 89% (RM3708). The maximum percentage of missing/null alleles was observed for the marker RM4996 (35.6%), followed by RM16649 with 28.1%,

[1]http://bbi.irri.org/products

[2]http://www.isbreeding.net

TABLE 1 | Descriptive statistics for the yield-related traits in F_2, F_3, and F_4 population of 166s × 14s.

Trait	Generation	Minimum	Maximum	Mean	Median	Range	Variance	Standard Deviation	Critical value	Skewness	Kurtosis
DIF	F_2	88	135	109.25	108	47	48.24	6.95	6.36	0.43	0.83
	F_3	88	129	98.87	98	41	42.26	6.5	6.58	1.26	3.01
	F_4	–	–	–	–	–	–	–	–	–	–
PH	F_2	50	100	77.04	77	50	87.52	9.36	12.14	0.03	0.11
	F_3	59.33	96	75.19	75.33	36.67	42.86	6.55	8.71	0.25	0.39
	F_4	68.67	98.33	84.42	84.00	29.66	38.48	6.20	7.35	0.04	−0.19
TN	F_2	2	40	18.25	17	38	64.2	8.01	43.91	0.54	−0.25
	F_3	7.67	28.33	14.46	14	20.66	13.63	3.69	25.52	0.65	0.47
	F_4	6.33	20.25	12.60	12.33	13.92	6.14	2.48	19.67	0.67	0.60
PTN	F_2	2	38	16.17	15	36	43.93	6.63	41	0.6	0.34
	F_3	7.33	26	14.06	13.33	18.67	13.22	3.64	25.86	0.68	0.29
	F_4	6.33	20.00	12.59	12.33	13.67	6.00	2.45	19.46	0.60	0.39
CL	F_2	34	72.83	51.7	51	38.83	47.67	6.9	13.35	0.46	0.45
	F_3	39.33	72	53.41	53	32.67	38.28	6.19	11.59	0.44	0.32
	F_4	–	–	–	–	–	–	–	–	–	–
SPY	F_2	1.6	75.1	25.61	23.4	73.5	205.22	14.33	55.93	0.73	0.49
	F_3	1.8	32.2	16.34	16.9	30.4	38.48	6.2	37.95	0.01	−0.32
	F_4	6.23	42.73	22.99	22.73	36.50	38.31	6.19	26.92	0.36	0.39
BM	F_2	2.78	85.8	27.77	25.4	83.02	213.82	14.62	52.66	0.82	0.88
	F_3	2.4	49.88	13.44	12.1	47.48	44.58	6.68	49.7	1.77	5.56
	F_4	5.53	52.50	25.68	25.03	46.97	157.79	12.56	48.92	0.19	−1.24
TDM	F_2	4.38	138.8	53.18	48.4	134.42	745.7	27.31	51.35	0.61	0.11
	F_3	8.35	68.88	29.81	28.53	60.53	122.43	11.06	37.11	0.75	1.15
	F_4	20.30	92.47	48.47	47.57	72.17	208.57	14.44	29.79	0.34	−0.37
HI	F_2	3.89	69.57	47.31	47.82	65.68	74.83	8.65	18.28	−1.21	3.36
	F_3	13.89	77.51	55.29	57.33	63.62	132.78	11.52	20.84	−0.97	1.2
	F_4	13.94	78.64	49.98	48.34	64.70	204.40	14.30	28.60	0.04	−0.92
PL	F_2	15.37	25.98	21.22	21.28	10.61	3.44	1.85	8.74	−0.24	0.39
	F_3	17.57	24.3	21.16	21.13	6.73	1.89	1.37	6.49	−0.08	−0.27
	F_4	–	–	–	–	–	–	–	–	–	–
FG	F_2	20	191.2	108.96	111	171.2	932.7	30.54	28.03	−0.08	0.07
	F_3	46.33	160	104.82	103.67	113.67	463.17	21.52	20.53	0.07	−0.19
	F_4	–	–	–	–	–	–	–	–	–	–
UFG	F_2	2.4	162.4	25.68	21.6	160	321.61	17.93	69.82	3.13	19.08
	F_3	2	81.67	19.74	15.5	79.67	198.55	14.09	71.4	1.66	3.34
	F_4	–	–	–	–	–	–	–	–	–	–
GN	F_2	38.67	254	134.47	130.8	215.33	1325.93	36.41	27.08	0.2	0.19
	F_3	70.33	192.5	124.38	120	122.17	620.48	24.91	20.03	0.53	−0.14
	F_4	–	–	–	–	–	–	–	–	–	–
SPF	F_2	36.06	98.4	81.13	83.06	62.34	104.45	10.22	12.6	−1.36	2.71
	F_3	37.77	98.2	84.63	87.29	60.43	97.36	9.87	11.66	−1.8	4.31
	F_4	–	–	–	–	–	–	–	–	–	–
PW	F_2	0.57	3.45	2.12	2.13	2.88	0.32	0.57	26.73	−0.27	0.21
	F_3	1.05	3.26	1.92	1.9	2.21	0.15	0.38	20.01	0.36	0.49
	F_4	–	–	–	–	–	–	–	–	–	–
TGW	F_2	10.2	23.8	19.44	19.6	13.6	4.61	2.15	11.05	−0.62	1.21
	F_3	7.9	23.2	17.53	17.3	15.3	4.27	2.07	11.8	−0.25	2.24
	F_4	14.20	25.33	20.44	20.50	11.13	3.40	1.85	9.03	0.01	0.20
PDP	F_2	0.01	0.55	0.19	0.17	0.54	0.01	0.11	57.75	0.79	0.56
	F_3	0.01	0.25	0.13	0.13	0.24	0	0.05	38.16	0.02	−0.32
	F_4	–	–	–	–	–	–	–	–	–	–
BY	F_2	–		–	–	–	–	–	–	–	–
	F_3	0.01	0.25	0.13	0.13	0.615	0.01295	0.1133	30.87	0.3036	0.0525

(Continued)

TABLE 1 | (Continued)

Trait	Generation	Minimum	Maximum	Mean	Median	Range	Variance	Standard Deviation	Critical value	Skewness	Kurtosis
	F_4	–	–	–	–	–	–	–	–	–	–
FLL	F_2	14	39.6	26.89	27	25.6	18.39	4.29	15.95	0.06	0.25
	F_3	–	–	–	–	–	–	–	–	–	–
	F_4	–	–	–	–	–	–	–	–	–	–
LL	F_2	21	47	32.99	32	26	24.26	4.92	14.93	0.33	0.02
	F_3	–	–	–	–	–	–	–	–	–	–
	F_4	–	–	–	–	–	–	–	–	–	–
LW	F_2	0.6	1.6	1.14	1.1	1	0.04	0.2	17.93	−0.29	0.1
	F_3	–	–	–	–	–	–	–	–	–	–
	F_4	–	–	–	–	–	–	–	–	–	–

DIF, days to initial flowering; PH, plant height; TN, tiller number; PTN, productive tiller number; CL, culm length; SPY, single plant yield; BM, biomass; TDM, total dry matter, HI, harvest index; PL, panicle length; FG, filled grains; UFG, unfilled grains; GN, grain number; SPF, spikelet fertility; PW, panicle weight; TGW, thousand-grain weight, PDP, per day productivity; BY, bulk yield; FLL, flag leaf length; LL, leaf length; LW, leaf width.

and RM1189 with 25.8%. The percentage alleles similar to 166s ranged from 1.15 to 47.13% with an average of 20.89%, while allelic similarity to 14s in the population was 9.77 to 58.62% with an average of 25.74%. Heterozygous bands in the F_2 population ranged from 14.37 to 69.54% with an average of 43.28%, and non-amplified/missing/null alleles were 0.57–35.63% (**Supplementary Table 4**). The parental lines of the mapping population were part of the set of CSSLs. 166s had 12.1% and 14s had 17.2% chromosomal segments from *O. nivara* based on genotyping using 111 SSR markers and 75.8 and 77.8% of Swarna alleles, respectively, with remaining heterozygous or unamplified bands. These BC_2F_8 parental lines are fixed sib lines, which explain the similarity between them.

Quantitative Trait Loci Mapping

F_2 linkage map was constructed using ICIM v4.1 software. There were 33 QTLs mapped using single marker analysis in F_2, $F_{2:3}$, and $F_{2:4}$ (**Table 3**). Eleven QTLs were mapped in F_2 with phenotypic variance (PV) between 5.17 and 9.71% with a maximum logarithm of odds (LOD) value of 5.04. Thousand-grain weight QTL showed the maximum LOD. In $F_{2:3}$, 17 QTLs were identified with PV between 4.37 and 12.49% with a maximum LOD value of 3.85. Four QTLs were mapped in $F_{2:4}$ with PV from 6.52 to 14.38%. The maximum phenotypic variance explained (PVE) was shown by QTL for thousand-grain weight or plant height in each generation. In F_2, using inclusive composite interval, mapping 21 QTLs were identified on chromosomes 1, 2, 3, 4, 5, 6, 7, 8, 9, and 11 for 19 yield-related traits, with PV ranging from 4.03 to 12.54%. In $F_{2:3}$, 30 QTLs were identified on chromosomes 1, 3, 4, 5, 6, 7, 8, 10, and 12 for 18 yield-related traits, with PV ranging from 1.32 to 13.01%. In $F_{2:4}$, 17 QTLs were identified on chromosomes 2, 3, 5, 6, 7, 8, and 12 for 8 yield-related traits, with PV ranging from 2.28 to 15.19%. In F_2, $F_{2:3}$, and $F_{2:4}$ (**Figure 2**) PV of QTLs ranged from 1.32 to 15.19%.

Pleiotropic Region for Quantitative Trait Loci

Many QTLs were clustered in the same region. There were six such regions found on chromosomes 1, 3, 4, 5, 6, and 9 in F_2 mapping population, on chromosome 1, 4, 6, 10, and 12 in $F_{2:3}$ mapping, and on chromosome 2, 5, 7, and 12 in $F_{2:4}$ mapping population. In F_2, three loci with the same marker interval responsible for multiple traits of yield QTLs were *qPTN4.1*, *qSPY4.1*, and *qTDM4.1* for productive tiller number, single plant yield, total dry matter mapped at RM16649- RM8213 on chromosome 4, and two QTLs *qPH1.1*, *qUFG1.1* identified in the region between RM1220 and RM3746. Two QTLs for filled grain *qFG3.1* and panicle weight *qPW3.1* mapped between RM4996 and RM7, and two QTLs *qUFG5.1* and *qTDM5.1* mapped at interval RM574- RM18614. *qPH6.1* and *qSPY6.1* were also identified between the same locus between RM19291 and RM217 on chromosome 6. Two QTLs, *qLL9.1* and *qLW9.1,* for leaf length and leaf width were located between RM23861 and RM1189 on chromosome 9.

In $F_{2:3}$, the QTLs for *qPH6.1, qCL6.1, qSPY6.1, qTDM6.1, qTGW6.1,* and *qPDP6.1 are* located at RM19291-RM217 marker interval were detected. Four QTLs, *qDIF4.2*, *qBM4.2*, *qTDM4.2*, and *qGN4.1,* were located between RM6997 and RM6909. Three QTLs, *qDIF1.1, qPH1.2,* and *qCL1.1,* were located in the same marker interval RM1220 - RM3746. QTLs *qDIF4.1*, *qBM4.1*, and *qTDM4.1* were located in the same locus between RM8213 and RM6997, while the locus between RM271 and RM269 harbored *qDIF10.1, qDIF10.2,* and *qPDP10.1* QTLs. Two QTLs for *qPH1.1* and *qUFG1.1* between RM495-RM246 and *qBM12.1* and *qDIF12.1* were located in the same marker interval of RM3331-RM7315.

In $F_{2:4}$, the QTLs for biomass and harvest index were colocated at marker interval regions of RM7485-RM279, RM153-RM574, and RM21539-RM22156, RM3331-RM7315 on chromosomes 2, 5, 5, 7, and 12, respectively. Harvest index (HI) QTL was associated with the last cluster on chromosome 12. The regions strongly associated with more than one trait in both F_2, $F_{2:3}$ include the region between RM19291 and RM217 which showed a cluster of QTLs for PH and SPY in F_2 and PH, CL, SPY, TDM, TGW, and PDP in F_3. The other region, RM1220-RM3746, showed a cluster of QTLs for PH and UFG in F_2 and DIF, PH, and CL in $F_{2:3}$.

Consistent Quantitative Trait Loci Across Generations

In all three populations, the QTL *qTGW8.1* for thousand-grain weight was consistently mapped at the same marker interval

TABLE 2 | Correlation coefficients among yield-related traits of F_2, F_3, and F_4 population of 166s × 14s.

	DIF	PH	TN	PTN	CL	SPY	BM	TDM	HI	PL	FG	UFG	GN	SPF	PW	TGW	PDP	FLL	LL	LW
F_2 DIF	*1***																			
F_3 DIF	*1***																			
F_2 PH	−0.32**																			
F_3 PH	0.18**																			
F_4 PH	–	*1***																		
F_2 TN	−0.18**	*0.29***																		
F_3 TN	0.01	*0.23***																		
F_4 TN	–	−0.02	*1***																	
F_2 PTN	−0.29**	*0.32***	*0.92***																	
F_3 PTN	0.02	*0.24***	*0.98***																	
F_4 PTN	–	−0.01	*0.98***	*1***																
F_2 CL	−0.21**	*0.80***	*0.22***	*0.26***																
F_3 CL	0.19**	*0.91***	*0.26***	*0.25***																
F_2 SPY	−0.33**	*0.47***	*0.74***	*0.79***	*0.36***															
F_3 SPY	0.07	*0.41***	*0.26***	*0.28***	*0.42***															
F_4 SPY	–	*0.28***	*0.39***	*0.39***	–	*1***														
F_2 BM	−0.27**	*0.53***	*0.76***	*0.76***	*0.43***	*0.78***														
F_3 BM	0.38**	*0.35***	*0.25***	*0.24***	*0.39***	*0.47***														
F_4 BM	–	0.02	*0.19***	*0.18***	–	0.08	*1***													
F_2 TDM	*−0.31***	*0.52***	*0.80***	*0.82***	*0.40***	*0.94***	*0.94***													
F_3 TDM	*0.28***	*0.44***	*0.30***	*0.30***	*0.48***	*0.84***	*0.87***													
F_4 TDM	–	0.12*	*0.34***	*0.32***	–	*0.51***	*0.89***	*1***												
F_2 HI	**−0.20****	0.10*	0.09	0.19**	0.06	*0.46***	−0.09	0.19**												
F_3 HI	**−0.31****	0.07	−0.03	0	0.05	*0.40***	−0.52**	−0.09												
F_4 HI	–	0.11*	0.00	0.02	–	*0.35***	*-0.87***	*-0.60***	*1***											
F_2 PL	−0.42**	*0.75***	0.27**	0.30**	*0.60***	0.49**	0.47**	0.50**	0.19**											
F_3 PL	−0.06	*0.29***	−0.09	−0.08	*0.23***	0.11*	0.12*	0.13*	−0.02											
F_2 FG	−0.28**	*0.62***	0.38**	0.43**	0.50**	0.60**	0.53**	0.59**	0.27**	*0.72***										
F_3 FG	0	*0.20***	−0.05	−0.06	0.16*	0.11*	0.11*	0.13*	0.04	*0.37***										
F_2 UFG	0	0.35**	0.02	−0.01	0.27**	−0.01	0.16*	0.08	**−0.27****	0.33**	0.07									
F_3 UFG	0.20**	0.06	−0.10*	−0.10*	−0.01	−0.07	0.14*	0.05	**−0.19****	0.08	−0.02									
F_2 GN	−0.22**	*0.69***	0.33**	0.36**	0.54**	0.50**	0.52**	0.54**	0.09	*0.77***	*0.87***	*0.55***								
F_3 GN	0.12*	*0.19***	−0.10*	−0.10*	0.11*	0.05	0.15*	0.12*	−0.06	*0.36***	*0.84***	*0.53***								
F_2 SPF	**−0.18****	−0.06	0.17**	0.21**	−0.07	0.28**	0.10*	0.20**	*0.41***	0.05	*0.39***	**−0.82****	−0.08							
F_3 SPF	**−0.18****	0	0.07	0.07	0.05	0.09	−0.12*	−0.03	*0.21***	0.03	*0.30***	**−0.94****	−0.26**							
F_2 PW	−0.39**	*0.69***	0.36**	0.41**	*0.55***	0.61**	0.53**	0.59**	0.31**	*0.80***	*0.90***	0.12*	*0.81***	*0.33***						
F_3 PW	0.02	*0.22***	−0.10*	−0.10*	*0.17***	0.14*	0.11*	0.14*	0.08	*0.36***	*0.72***	0	*0.61***	*0.21***						
F_2 TGW	**−0.31****	*0.25***	0.08	0.14*	*0.17***	*0.24***	*0.18***	*0.23***	*0.22***	*0.25***	0.15*	**−0.11***	0.07	*0.23***	*0.31***					
F_3 TGW	0.01	0.06	0.06	0.07	0.08	0.14*	0.05	0.10*	0.07	−0.07	−0.14*	**−0.14***	**−0.20****	0.09	0					
F_4 TGW	–	*0.18***	−0.06	−0.03	–	0.12*	−0.08	0.00	0.10*	–	–	–	–	–	–	*1***				
F_2 PDP	**−0.40****	*0.47***	*0.72***	*0.78***	*0.36***	*1.00***	*0.77***	*0.94***	*0.46***	*0.50***	*0.59***	−0.01	*0.49***	*0.28***	*0.61***	*0.25***				
F_3 PDP	−0.07	*0.39***	*0.27***	*0.28***	*0.39***	*0.99***	*0.42***	*0.80***	*0.44***	0.12*	0.11*	−0.10*	0.04	0.11*	0.14*	0.13*				
F_3 BY	0.10*	*0.25***	0.01	−0.01	*0.29***	*0.30***	*0.16***	*0.27***	*0.19***	0.09	0.06	−0.05	0.01	0.06	0.08	0.07	*0.28***			
F_2 FLL	**−0.39****	*0.64***	0.07	0.15*	*0.51***	*0.32***	*0.30***	*0.31**	*0.18***	*0.55***	*0.50***	*0.29***	*0.55***	−0.04	*0.53***	0.12*	*0.33***	*1***		
F_2 LL	**−0.34****	*0.66***	*0.17***	*0.24***	*0.57***	*0.41***	*0.42***	*0.42***	*0.18***	*0.58***	*0.55***	*0.27***	*0.58***	0.01	*0.58***	*0.22***	*0.42***	*0.72***	*1***	
F_2 LW	**−0.20****	*0.63***	*0.29***	*0.23***	*0.46***	*0.43***	*0.48***	*0.48***	0.10*	*0.51***	*0.50***	*0.27***	*0.56***	0	*0.54***	0.16*	*0.43***	*0.51***	*0.60***	*1***

P = 0.05–0.001, significant lines.

***P ≥ 0.001, highly significant lines.*

Highly significant positive values in italics, highly significant negative values in bold.

F_2, F_3, and F_4-Generations.

DIF, days to initial flowering; PH, plant height; TN, tiller number; PTN, productive tiller number; CL, culm length; SPY, single plant yield; BM, biomass; TDM, total dry matter, HI, harvest index; PL, panicle length; FG, filled grains; UFG, unfilled grains; GN, grain number; SPF, spikelet fertility; PW, panicle weight; TGW, thousand-grain weight, PDP, per day productivity; BY, bulk yield; FLL, flag leaf length; LL, leaf length; LW, leaf width. "–" Observations not taken for one of correlated trait.

TABLE 3 | Additive and dominance effects of quantitative trait loci (QTLs) identified for yield-related traits in F_2, $F_{2:3}$, and $F_{2:4}$ populations of 166s × 14s using ICIM.

SL. No.	QTL	Chromosome	Generation	Left marker	Right marker	LOD	PVE (%)	Additive	Dominance
1	*qLW1.1*	1	F_2	RM8004	RM1220	3.15	6.06	0.11	−0.04
2	*qPH1.1*	1	F_2	RM1220	RM3746	2.54	8.18	2.96	−1.92
3	*qUFG1.1*	1	F_2	RM1220	RM3746	3.59	2.78	5.92	−5.91
4	*qDIF1.1*	1	F_3	RM1220	RM3746	4.2	1.94	2.58	0.38
5	*qPH1.1*	1	F_3	RM495	RM246	3.51	13.01	4.64	−2.45
6	*qPH1.2*	1	F_3	RM1220	RM3746	4.69	6.63	2.7	0.08
7	*qCL1.1*	1	F_3	RM1220	RM3746	2.72	8.93	1.73	1.21
8	*qUFG1.1*	1	F_3	RM495	RM246	8.8	1.63	20.41	−22.33
9	*qHI2.1*	2	F_2	RM3515	RM13599	3.71	4.12	7.27	8.92
10	*qBM2.1*	2	F_4	RM7485	RM279	7.14	2.93	−0.7	20.59
11	*qHI2.1*	2	F_4	RM7485	RM279	3.24	3.62	−0.47	−22.25
12	*qHI2.2*	2	F_4	RM279	RM13616	2.84	3.52	−1.59	−21.53
13	*qFG3.1*	3	F_2	RM4996	RM7	2.53	9.65	−8.87	11.28
14	*qPW3.1*	3	F_2	RM4996	RM7	3.05	9.06	−0.19	0.21
15	*qUFG3.1*	3	F_2	RM232	RM6759	10.04	8.26	4.25	−19.43
16	*qTN3.1*	3	F_3	RM4996	RM7	3	8.35	−0.25	1.75
17	*qUFG3.1*	3	F_3	RM232	RM6759	5.19	1.71	13.53	−14.82
18	*qTN3.1*	3	F_4	RM7	RM232	2.53	6.56	−0.9	1.22
19	*qPL4.1*	4	F_2	RM317	RM17377	3.16	8.17	0.54	0.62
20	*qPTN4.1*	4	F_2	RM16649	RM8213	2.52	5.2	−3.75	−1.65
21	*qSPY4.1*	4	F_2	RM16649	RM8213	2.6	8.5	−7.26	−4.34
22	*qTDM4.1*	4	F_2	RM16649	RM8213	2.65	8.97	−14.53	−7.29
23	*qDIF4.1*	4	F_3	RM8213	RM6997	2.67	5.06	−3.83	−3.41
24	*qDIF4.2*	4	F_3	RM6997	RM6909	4.47	6.98	−4.82	−4.19
25	*qBM4.1*	4	F_3	RM8213	RM6997	7.68	4.29	−4.25	−4.95
26	*qBM4.2*	4	F_3	RM6997	RM6909	6.41	4.22	−4.24	−5.17
27	*qTDM4.1*	4	F_3	RM8213	RM6997	5	8.54	−6.84	−7.01
28	*qTDM4.2*	4	F_3	RM6997	RM6909	4.1	8.45	−7.03	−6.72
29	*qPL4.1*	4	F_3	RM6909	RM317	2.56	6.74	0.46	0.09
30	*qGN4.1*	4	F_3	RM6997	RM6909	3.59	3.63	15.35	−27.04
31	*qTDM5.1*	5	F_2	RM574	RM18614	2.61	4.03	−11.01	8.36
32	*qUFG5.1*	5	F_2	RM574	RM18614	3.39	5.18	10.04	−9.86
33	*qUFG5.1*	5	F_3	RM430	RM3664	5.25	1.83	−13.16	−13.73
34	*qBM5.1*	5	F_4	RM153	RM574	10.15	3.23	−4.47	−22.17
35	*qBM5.2*	5	F_4	RM430	RM3664	8.27	2.94	−1.43	20.49
36	*qHI5.1*	5	F_4	RM153	RM574	4	3.76	2.22	22.88
37	*qHI5.2*	5	F_4	RM430	RM3664	4.24	3.69	2.35	−22.06
38	*qFLL6.1*	6	F_2	RM217	RM402	2.87	7.5	1.57	0.73
39	*qPH6.1*	6	F_2	RM19291	RM217	3.75	12.54	−1.46	−4.93
40	*qSPY6.1*	6	F_2	RM19291	RM217	2.65	6.71	−3.94	−7.16
41	*qPH6.1*	6	F_3	RM19291	RM217	4.52	7.07	−2.34	−1.38
42	*qTN6.1*	6	F_3	RM276	RM204	3.59	10.08	0.11	−1.91
43	*qCL6.1*	6	F_3	RM19291	RM217	3.09	9.81	−1.48	−1.55
44	*qSPY6.1*	6	F_3	RM19291	RM217	2.95	6.51	−0.82	−3.93
45	*qTDM6.1*	6	F_3	RM19291	RM217	2.86	2.36	−1.59	−5.29
46	*qTGW6.1*	6	F_3	RM19291	RM217	2.51	9.33	0.78	−1.54
47	*qPDP6.1*	6	F_3	RM19291	RM217	3.06	6.39	−0.01	−0.04
48	*qTGW6.1*	6	F_4	RM19291	RM217	3.52	15.19	0.56	−2.06
49	*qDIF7.1*	7	F_2	RM21539	RM22156	3.22	8.48	2.81	0.45
50	*qDIF7.1*	7	F_3	RM21539	RM22156	4.84	8.65	5.47	−3.89
51	*qBM7.1*	7	F_4	RM21539	RM22156	6.51	2.92	1.36	20.48
52	*qHI7.1*	7	F_4	RM21539	RM22156	3.39	3.55	−1.91	−21.6
53	*qTGW8.1*	8	F_2	RM3480	RM3452	4.27	9.86	−1.53	0.82

(Continued)

TABLE 3 | (Continued)

SL. No.	QTL	Chromosome	Generation	Left marker	Right marker	LOD	PVE (%)	Additive	Dominance
54	*qTGW8.1*	8	F_3	RM3480	RM3452	5.43	7.48	−1.21	0.03
55	*qTGW8.1*	8	F_4	RM3480	RM3452	5.87	6.11	−0.98	−0.08
56	*qLL9.1*	9	F_2	RM23861	RM1189	2.65	6.91	0.87	−3.91
57	*qLW9.1*	9	F_2	RM23861	RM1189	3.14	8.9	0.06	−0.15
58	*qDIF10.1*	10	F_3	RM271	RM269	4.41	2	2.52	−0.44
59	*qDIF10.2*	10	F_3	RM271	RM269	2.71	1.32	−1.39	2.42
60	*qPDP10.1*	10	F_3	RM271	RM269	2.82	7.1	0.03	0.02
61	*qGN11.1*	11	F_2	RM27154	RM206	2.57	5.61	19.7	−15.02
62	*qDIF12.1*	12	F_3	RM3331	RM7315	2.77	7.64	4.76	−5.87
63	*qBM12.1*	12	F_3	RM3331	RM7315	3.96	4.09	4.4	−5.52
64	*qBM12.1*	12	F_4	RM3331	RM7315	6.18	2.97	2.4	−20.75
65	*qBM12.1*	12	F_4	RM3331	RM7315	7.63	2.98	−2.27	20.87
66	*qBM12.2*	12	F_4	RM7315	RM3747	5.4	2.92	−2.26	20.3
67	*qTDM12.1*	12	F_4	RM3331	RM7315	2.72	2.28	4.29	−21.8
68	*qHI12.1*	12	F_4	RM3331	RM7315	3.38	3.6	0.59	−22.19

QTLs named by trait abbreviation and chromosome number, LOD, logarithm of odds; PVE%, phenotypic variance explained by the QTL.
DIF, days to initial flowering; PH, plant height; PTN, productive tiller number; SPY, single plant yield; TDM, total dry matter; HI, harvest index; PL, panicle length; FG, filled grains; UFG, unfilled grains; GN, grain number; PW, panicle weight; TGW, thousand-grain weight; PDP, per day productivity; BY, bulk yield; FLL, flag leaf length; LL, leaf length; LW, leaf width. −ve values indicate additive and dominance effect from male parent (14s).
Consistent QTL in 3 generations are shown in bold.
>20 dominance values was observed with 14s allele for 10 QTLs (5 QTLs for HI, 2 QTLs for BM, 1 QTL each for TDM, UFG, and GN).
>20 dominance values was observed with166s allele for 5 QTLs (1 QTL for HI and 4 QTLs for BM).

RM3480- RM3452 on chromosome 8 with LOD ranging from 4.2 to 5.8 and PV ranging from 6.1 to 9.8% (**Figure 3**). QTLs detected in only F_2 and $F_{2:3}$ were for PH and SPY between RM19291 and RM217 on chromosome 6. Three QTLs, *qDIF7.1, qBM12.1,* and *qTGW6.1,* were detected in only $F_{2:3}$ and $F_{2:4}$. *qTN3.1* was mapped in $F_{2:3}$ and $F_{2:4}$, but with different marker intervals. QTL *qTGW8.1* was selected for further genomic dissection as it was consistently identified in F_2, F_3, and F_4 generations at the same 2.6 Mb region. SSR markers located between these flanking markers were selected to enrich this 2.6 Mb region and used for further genotyping. Five lines each for high and low TGW were selected in F_2 based on mean TGW and if the same high or low TGW trait was present in at least two of the three generations. First, F_2 leaf DNA samples of five lines, each with high and low TGW, were separately used for genotyping with eight SSR markers within QTL. These samples showed clear polymorphism between high and low TGW lines. Marker trait association was detected between TGW and the locus RM 502- RM3480 with a PVE of 47% at LOD 1.49. Later, DNA samples of 10 high and 10 low TGW F_4 lines were taken and genotyped using two co-segregating markers for validation (**Figure 3**). The high (20.1 to 25.3 g) and low (15.2 to 18.1 g) TGW lines were significantly different from both parents 166s (17.2 g) and 14s (21.6 g) as well as from Swarna (14.2 g). Genotyping of contrasting lines with 8 SSRs showed that two markers RM23407 and RM23447 showed clear polymorphism between high and low TGW lines in F_2. However, only RM23447 showed polymorphism between extreme TGW phenotypes of F_4. In F_2, among five low TGW lines, the percentage of 166s type alleles was higher than 14s alleles, while the five high TGW lines showed more of 14s type alleles than 166s type. Considering all alleles of the 5 high TGW lines, 50% alleles were of 14s type in C2 117, and 37.5% alleles were of 14s type in the other four lines (C2 103, C2 126, C2 128, and C2 152). Three lines (C2 103, C2 126, and C2 128) had 25% heterozygous alleles and two lines C2 117, and C2 152 had only 12.5% heterozygous alleles. The statistically significant lines identified in our study, compared to both the parents and their common cultivar parent Swarna, were further used for small-scale fine mapping of *qTGW8.1* to detect causative genes for grain weight improvement (**Figure 3**). Five lines, each with high and low TGW, were selected in F_2 based on their extreme placement in the frequency distribution curve and if the same high or low TGW trait were present in the same extremes. Likewise, at least two of the three generations were genotyped using SSR markers within the QTL region. Trait marker association using interval mapping detected RM502-RM3480 with PVE of 47% at 1.49 LOD, but no trait association was detected with RM23447 even though 2 groups showed clear co-segregation. Further genotyping in F_4 identified two co-segregating markers, RM23407 and RM23447, covering a 602.4 kbp region {25148785bp to 25751225bp [QTARO database (affrc.go.jp)]} within *qTGW8.1*.

Quantitative Trait Loci Clusters Detected

In this study, among all QTLs detected in F2, F2:3, and F2:4, two specific marker regions were strongly associated with more than one trait. RM19291-RM217 had a cluster of QTLs for PH, SPY in F2, and PH, CL, SPY, TDM, TGW, and PDP in F_3. The other region of RM1220- RM3746 had a cluster of PH, UFG in F_2, and DIF, PH, and CL in $F_{2:3}$. In addition, the QTLs present in each region were correlated except PH with UFG in F2:3 mapping. In $F_{2:4}$, one specific marker interval, RM3331- RM7315, had

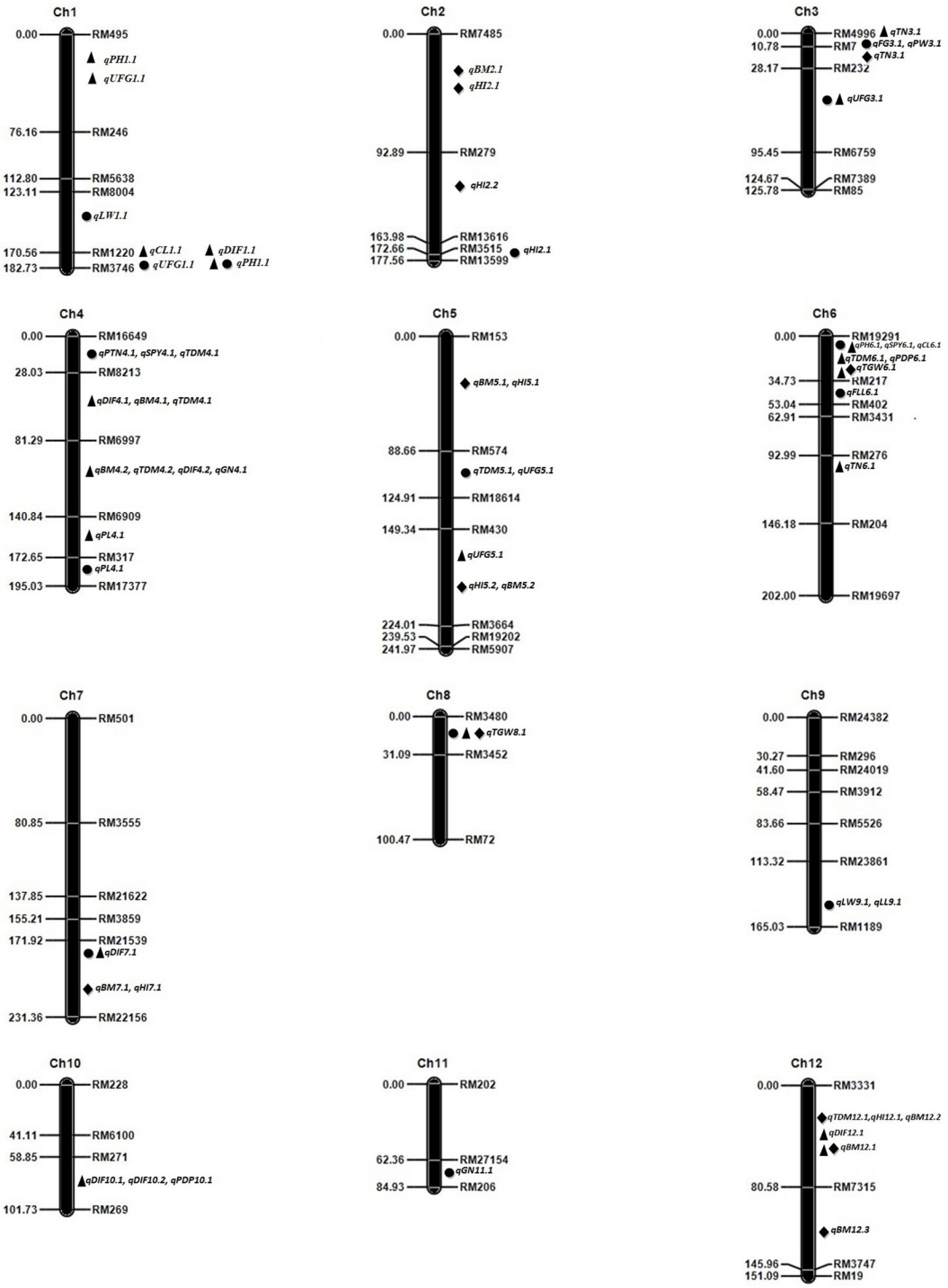

FIGURE 2 | Molecular linkage map of 12 chromosomes with position of Quantitative Train Loci (QTLs) for yield-related traits in $F_{2:3:4}$ generations detected using ICIM. QTLs named by trait abbreviation and chromosome number,● – F_2 generation,▲ – F_3 generation,♦ – F_4 generation. DIF, days to initial flowering; PH, plant height; PTN, productive tiller number; SPY, single plant yield; TDM, total dry matter; HI, harvest index; PL, panicle length; FG, filled grains; UFG, unfilled grains; GN, grain number; PW, panicle weight; TGW, thousand-grain weight; PDP, per day productivity; BY, bulk yield; FLL, flag leaf length; LL, leaf length; LW, leaf width.

a cluster of QTLs for BM (two QTLs), TDM, and HI on chromosome 12. In addition, the region at marker intervals of RM153- RM574 and RM430- RM3664 had a cluster of QTLs for BM and HI on chromosome 5. Lastly, the same QTL cluster for BM and HI were identified on chromosomes 2 and 7 at the regions RM7485-RM379 and RM21539-RM22156, respectively.

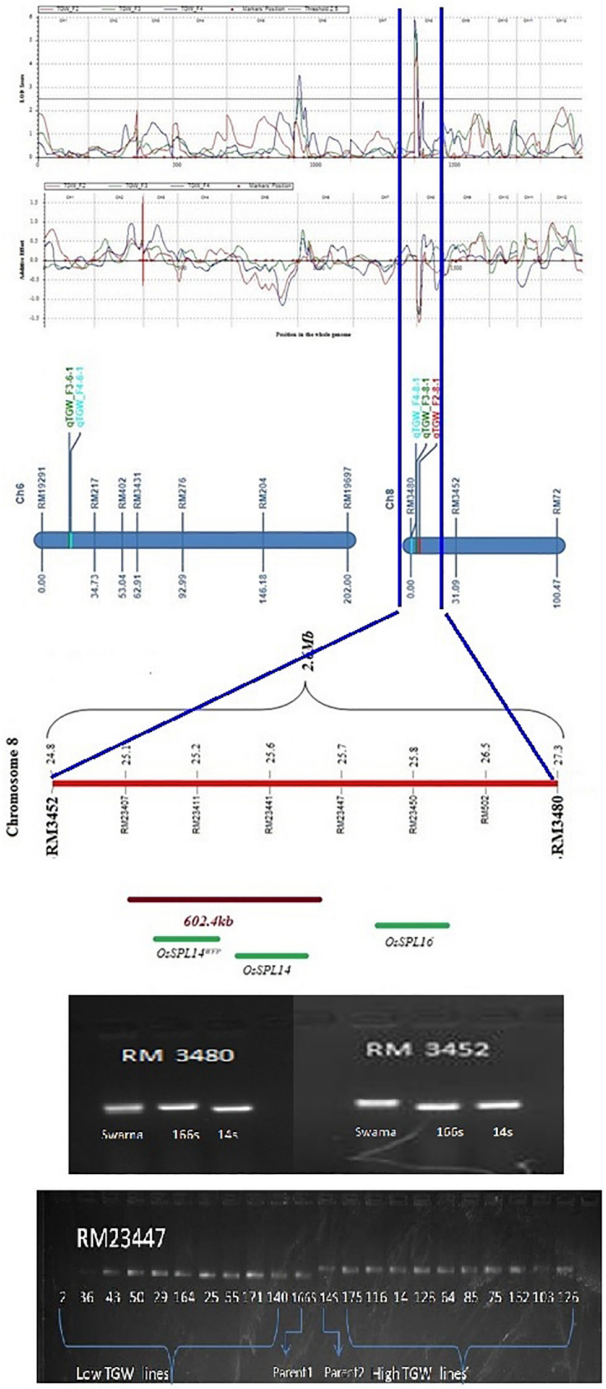

FIGURE 3 | Genetic dissection of *qTGW8.1* QTL from 2.6 Mb region between RM3480-RM3452 and 602.4 Kb region between RM23407 and RM23447 based on genotypic segregation in high and low TGW phenotypes in F_4.

DISCUSSION

Phenotyping the mapping population for 3 generations showed a significant positive association of yield contributing traits PH, TN, PTN, and TDM to single plant yield. Consistent character association revealed the importance of tiller number per plant, productive tillers per plant, and thousand-grain weight as selection criteria for effective yield improvement. Our study showed that F_2 population is more powerful for detecting QTLs of additive effect and can also be used to estimate the degree of dominance for detected QTLs. All the QTLs mapped in F_4 were observed with low additive effects from both the parent alleles. A high dominant effect from male parent alleles for the 9 QTLs *qBM5.1*, *qBM12.1*, *qTDM12.1*, *qHI2.1*, *qHI2.2*, *qHI5.2*, *qHI7.1*, *qHI2.1,* and *qTGW6.1* was also observed. Dominance effects were high for the QTLs mainly because of the superiority of heterotic loci identified with low phenotypic variances. QTLs with the low dominant effects showed high phenotypic variances except for the QTLs FG and UFG in F_2. About 38.1% in F_2 and 30% in $F_{2:3}$ QTLs explained positive dominance effect, and 61.9% QTLs in F_2 and 70% in $F_{2:3}$ explained negative dominance values. The QTL *qTGW6.1* showed additive effects from female parents in F_3 (0.78) and F_4 (0.56), and also showed dominance effects from male parents in both F_3 (−1.54) and F_4 (−2.06). *qTGW8.1* showed a higher additive effect (*O. nivara*) in F_2 than in F_3 and F_4, along with a small dominance effect (0.82) in F_2. The high dominance effects were identified for the QTLs in F_4 (82%), followed by F_2 (14.2%), and F_3 (13.2%).

In this study, significant QTLs were identified in F_2, $F_{2:3}$, and $F_{2:4}$, with 9.0–15.19% PV for several yield-related traits. Trait enhancing allele in major QTLs *qPH1.1*, *qTGW6.1*, and *qTGW8.1* and minor QTLs *qGN4.1* and *qTDM5.1* was from *O. nivara*. Swamy et al. (2011, 2012) identified 40% of *O. nivara* alleles were trait enhancing in QTLs in Swarna/*O. nivara* BC_2F_2 populations. Sarla (2014) reported that two QTLs *yld9.1* (yield) and *nfg9.1* (number of filled grains) were from *O. nivara* in BIL 248S (DRRDhan40), derived from the same cross. Surapaneni et al. (2017) identified 15 QTLs in a 94 BILs mapping population using 111 SSRs and, of these, 26% QTLs had trait enhancing alleles from *O. nivara*. For the yield QTLs *qSPY4.1* and *qSPY6.1,* trait enhancing allele was from Swarna. The common QTLs *qPH1.1* and *qTGW8.1* at RM1220-RM3746 and RM3480-RM3452, respectively, in F_2 and $F_{2:3}$ shared the same locus. In these QTLs, PV value was reduced (*qPH1.1*-6.63% and *qTGW8.1*-7.48%) and LOD values increased (*qPH1.1*-4.69 and *qTGW8.1*-5.43) in $F_{2:3}$ compared to F_2. This reduction of the phenotypic variance of the traits in advanced generations may be due to a reduction in heterozygosity and stabilization of introgression lines. Another common QTL for single plant yield *qSPY6.1* at the locus between RM19291 and RM217 in F_2 and $F_{2:3}$ showed no significant change in LOD (2.65 in F_2, 2.95 in F_3) and PV (6.71% in F_2, 6.51% in F_3) values across the generations.

Major Yield Quantitative Trait Loci Identified in This Study

Among the eight major QTLs, six *qPH1.1, qPH6.1*, *qTGW6.1*, *qFG3.1*, *qTN6.1,* and *qCL6.1* were previously reported in the population derived from cultivated rice varieties for the same trait. The other 2 QTLs, *qTGW8.1* and *qPW3.1,* were reported in wild introgression lines in different genomic regions. *qPH1.1*

was identified with the same flanking marker, RM246, by Zhang et al. (2010). *qPH1.1* was also reported previously from *O. nivara* (Swamy et al., 2014; Surapaneni et al., 2017; Haritha et al., 2018; Balakrishnan et al., 2020). *qTGW6.1* was reported previously by Zhu et al. (2019) in the *O. sativa* recombinant inbred line (RIL) population [Teqing/IRBB lines (TI), Zhenshan97/Milyang46 (ZM) and Xieqingzao/Milyang46 (XM)] within the marker region of our study (**Table 4**).

The QTL, *qSPY4.1,* with PV% 8.5 and LOD of 2.6, which was detected in this study along with Xing et al. (2014), Ma et al. (2016), and Haritha et al. (2018), reported grain yield QTLs on chromosome 4. Fan et al. (2019) also reported *qGYPP4* in the BIL population derived from *Oryza longistaminata.* In our study, 8 QTLs for biomass were identified on chromosomes 2, 4, 5, 7, and 12 with LOD values ranging from 3.96 to 10.15 with a maximum PV of 4.29%. Zhang et al. (2017) mapped the BM QTL on chromosome 12 in $F_{2:3}$ and RILs population. Six QTLs, *qHI2.1, qHI2.2*, *qHI5.1*, *qHI5.2*, *qHI7.1,* and *qHI12.1,* were mapped with a maximum LOD of 3.7 and PV of 4.12% for *qHI2.1* QTL. In a previous study of Sabouri et al. (2009), they identified *qHI2* with 21.35% of phenotypic variance. Grain number of panicle is one of the most important traits to contribute to yield, with *qGN4.1* QTL identified as the highest additive effect QTL and maximum LOD value of 3.5. This QTL was previously reported by Xing et al. (2014), Kim et al. (2017), and Deng et al. (2017). Singh et al. (2018) also reported the same *qGN4.1* QTL in bi-parental RIL population.

The QTL *qFG3.1* was found with 9.6% of phenotypic variance. Feng et al. (2015) and Zhang et al. (2018) reported QTLs for filled grain numbers as *qFGP3a* and *qFGN3,* respectively. QTLs for unfilled grains per panicle were identified on chromosomes 1, 3, and 5 with LOD value of 10.0. Five QTLs were observed for UFG on chromosomes 1, 3, and 5 with different flanking markers in both the populations except for *qUFG3.1* (Rabiei et al., 2015), with 8.2% of phenotypic variance in F_2 population. Two QTLs, *qTN3.1* and *qTN6.1,* for tiller number and *qPTN4.1* for productive tiller number, were identified. Lim et al. (2014) also reported the same QTL on chromosome 6 with LOD of 3.79. In the present study, *qPL4.1* was identified with different flanking markers in F_2 and $F_{2:3}$ with 8.1% of the phenotypic variance on chromosome 4 with above 2.5 LOD value. It was mapped by Lim et al. (2014) and Kim et al. (2017) in *indica-japonica* recombinant inbred line population. Xing et al. (2014) and Zhang et al. (2018) reported QTL for productive tiller number on chromosome 4. Zhang et al. (2018) reported *qPL4* QTL in the background of IL population derived from cross between ZGX1 (high-quality *indica* elite variety) × IR75862 (high iron and zinc *japonica* variety).

Among the QTLs detected, the plant height QTL, *qPH1.1,* was identified in $F_{2:3}$ with 13% of phenotypic variance value and with LOD value of 3.5. The same flanking marker RM246 for plant height QTL was previously reported by Zhang et al. (2010) in F_2 population derived from cross between PA64s (*indica*) × Nipponbare (*japonica*) and Rabiei et al. (2015) in $F_{2:4}$ population derived by crossing two *indica* rice varieties (Sepidrood × Gharib). Sabouri et al. (2009) also identified the same plant height QTL in F_2 and $F_{2:3}$ populations derived by crossing two genetically divergent *indica* type high-yielding rice varieties. Qiao et al. (2016) reported *qPH1.1* and *qPH1.2* with 7.01 and 9.05% PV, respectively, in the 198 CSSLs population derived from cross between *indica* var. 9311 and wild species *O. rufipogon* as donor parent. QTL *qPH1.1,* in BILs/ILs population previously derived from Swarna/*O. nivara* but with different flanking markers in chromosome 1, was reported by Swamy et al. (2014), Surapaneni et al. (2017), Haritha et al. (2018), and Balakrishnan et al. (2020). Another plant height QTL was mapped in F_2 population with 3.7 LOD with a phenotypic variance of 12.5% on chromosome 6. *qPH6.1* QTL was previously reported in the F_2 mapping population (Zhang et al., 2010) and in $F_{2:7}$ recombinant inbred line population derived from a cross between Xiaobaijingzi (upland rice) and Kongyu 131 (Xing et al., 2014). Lim et al. (2014) also identified plant height QTL on chromosome 6 in $F_{2:7}$ RIL population by crossing *indica* × *japonica* (Milyang23/SNUSG1) using the QGene 4.3.10 software.

qFLL6.1 was identified with phenotypic variance of 7.4% and LOD value of 2.8 in the present study with the flanking region between RM217 and RM402. RM402 was detected in the same region in the study reported by Shen et al. (2011) who identified *qFLL6.1* and *qFLL6.2* in the 4.2 Mb region between flanking markers RM4923 and RM402 on short arm of chromosome 6 in NILs. Shao et al. (2009) reported flag leaf length QTL on chromosome 6. Matsubara et al. (2016) and Yang et al. (2018) reported *qFLL* QTL on chromosome 6 in RIL population and DH population, respectively. Leaf length QTL *qLL9.1* was identified with phenotypic variance of 6.9% with LOD of 2.6, while two QTLs for leaf width, *qLW9.1*, *qLW1.1,* were identified with 11 and 5.7% phenotypic variance on chromosome 9 and 1, respectively.

Seven QTLs for days to initial flowering were identified on chromosomes 1, 4, 7, 10, and 12. The QTL with the highest PV of 8.65% was observed with LOD value of 4.84 on chromosome 7. Mohammadi et al. (2013) reported QTLs for number of days to flowering on chromosomes 4 and 10 in the F_2 population of Sadri/FL478 cross under saline field conditions. The major effect heading date QTLs was previously identified in chromosome 7 by Li et al. (2003) in IR64/Azucena doubled-haploid population. Yano et al. (2000) and Xue et al. (2008) identified *Hd1* on chromosome 6 and *Ghd7* on chromosome 7. Salgotra et al. (2015) reported *Ghd7* gene with associated traits of grains per panicle and plant height along with heading date. These previous reports also confirmed detection of the major QTLs in the same chromosomal regions as identified in our study.

Consistent Quantitative Trait Loci Detected in This Study

qTGW8.1 is a novel QTL consistently identified within the flanking marker region between RM3480 and RM3452 in three populations. *GW8.1* was previously mapped from another wild species *O. rufipogon* (Tian et al., 2006; Kang et al., 2018), and QTL for panicle weight was also reported from *O. nivara* on chromosome 3 but with different flanking markers. The number of traits studied was different across the generations and only 8 traits were phenotyped in all three generations. Therefore,

TABLE 4 | Major QTLs identified in present study compared to previously reported QTLs.

S. No.	Major QTLs identified in present study	Flanking markers	Position (Mb)	Previous reports of QTLs/genes in this region	Parental lines	References
1	*qTGW8.1*	RM3480-RM3452	2.614	*gw8.1* (RM531-RM42)	GW QTL region from *O. rufipogon* wild F101 (BIL)/Hwaseongbyeo	Xie et al., 2006
				gw8.1 (RM23201-RM23208)	GW QTL from *O. rufipogon* wild NIL/NIL	Kang et al., 2018
				qTGW8 (RM6845)	Nipponbare/Z550	Wang et al., 2016
				qTGW8 (SNPs)	PA64s/CSSLs and CSSL/9311	Lin et al., 2020
2	*qTGW6.1*	RM19291-RM217	3.01	*qTGW6.1* (C358)	*japonica* (Nipponbare)/*indica* (Kasalath)	Ishimaru, 2003
				qTGW6.1 (RM276-RM136)	ZGX1/IR75862 IL population	Zhang et al., 2018
				qTGW6.1 (RM402-RM5963)	$F_{2:3}$ developed from *indica* rice lines	Sun et al., 2019
				qTGW6 (RM589-190)	Teqing/IRBB lines (TI), Zhenshan 97/Milyang 46 (ZM) and Xieqingzao/Milyang 46 (XM) RIL population	Zhu et al., 2019
3	*qPH1.1*	RM495-RM246	27.1	*qph1.1* (RM246)	PA64s (*indica*)/Nipponbare (*japonica*)	Zhang et al., 2010
				qPH1 (RM237-RM246)	Sepidrood/Gharib Indica varieties	Rabiei et al., 2015
				qPH1-1 (RM128)	CSSL population (9311/*O. rufipogon*)	Qiao et al., 2016
				qPH1.1 (RM226-RM431)	Swarna/*O. nivara* BIL population	Surapaneni et al., 2017
						Haritha et al., 2018
4	*qPH6.1*	RM19291-RM217	3.01	*qPH-6* (RM162-RM412)	Xiaobaijingzi/Kongyu 131 RIL	Xing et al., 2014
				qph6.1 (N6255-S6279.1)	Milyang23/SNUSG1 $F_{2:7}$ RIL	Lim et al., 2014
5	*qFG3.1*	RM4996-RM7	1.3	*qFGP3a* (RM135-168)	Xieqingzao B/Zhonghui 9308 RIL	Feng et al., 2015
				qFGN3 (RM85-RM227)	Ce258/IR75862 (IL)	Zhang et al., 2018
6	*qTN6.1*	RM276-RM204	3	*qTN6.1* (RM3)	Milyang23/SNUSG1 $F_{2:7}$ RIL	Lim et al., 2014
7	*qPW3.1*	RM4996-RM7	1.3	*qpw3.2* (RM85)	Swarna/*O. nivara* BIL population	Haritha et al., 2018
8	*qCL6.1*	RM19291-RM217	3.01	*qCL-6* (RM4447)	Hinohikari/WSS2//Hinohikari BC_1F_1 population	Sato et al., 2004

we could check the consistency of QTLs of only these 8 traits. Further, we checked the consistency for the selected QTLs above threshold level after 1,000 permutations so QTLs below the threshold level, even in one generation, were not considered consistent even though the QTLs were mapped. As the QTL mapping was carried out in different generations, reduction in segregation, heterozygosity, change in context (background genome), and further stabilization of lines and environmental factors may also have contributed to identification of only few consistent QTLs. Stable QTLs which are consistently detected in different generations are useful in marker assisted selection (MAS) for crop improvement (Guo et al., 2005; Su et al., 2010).

Quantitative Trait Loci for Grain Weight and Further Genetic Dissection

Of all the major QTLs detected, *qTGW8.1* was mapped in all three generations of F_2, $F_{2:3}$, and $F_{2:4}$ at the same marker interval of RM3480-RM3452, covering a 2.6 Mb region with significant PVE ranging from 6.1 to 9.8%. Previous studies showed *qTGW* QTL in the background of *O. nivara* parent on other chromosomes 1, 2, 4, and 5 by Kaladhar et al. (2008), on chromosomes 1, 2, 3, and 4 by Swamy et al. (2014), and on chromosomes 1, 4, and 5 by Haritha et al. (2018). *qTGW8.1,* identified in the present study, is a novel QTL with *O. nivara* contributing to the additive effect. This region is already reported to harbor genes, *viz.*, Wealthy Farmer's Panicle (WFP) encoding $OsSPL14^{WFP}$ (Miura et al., 2010) and Ideal Plant Architecture (IPA1) encoding OsSPL14 (Jiao et al., 2010) and located 875 kbp upstream of already reported QTL *qGW8,* with gene Os08g0531600 (*GW8*/OsSPL16) (Wang et al., 2012). The identified QTL in this study could be a novel allele of these reported genes from *O. nivara* or a novel gene/QTL colocated in this region since there is an overlap with three known genes. However, this requires further genetic dissection and confirmation. The QTL trait enhancing allele is from the male parent 14s. Gel pictures of these markers showed a lower band size at 14s compared to Swarna which is indicative of the presence of a novel allele with deletion at the locus RM3480 triggered by recombination or any unknown factors contributed by *O. nivara* or Swarna genome.

Three meta-QTLs for yield on chromosome 8 (MQTL 8.1, 8.2, 8.3) were detected for various yield and grain-related traits. MQTL8.2 (326 kb) was identified as a recombination hot spot suitable for fine mapping yield-related traits including grain weight (Swamy and Sarla, 2011). It is interesting to note that the only meta-QTL for TGW on chromosome 8 (MQTL-GW4 S3680-RM3689 at 18.25–19.33 Mb), reported recently (Wang et al., 2021), is proximal to our *qTGW8.1*. The identified QTL in this study could be a novel allele of these reported genes from *O. nivara* or a novel gene colocated in this region. As the trait enhancing QTL allele is from male parent 14s and the gel pictures of these markers showed a lower band size at 14s compared to Swarna, it indicates the presence of a novel allele with

possible deletions in 14s at loci RM3480, RM3452, and RM23447 compared to Swarna and can be confirmed with sequencing.

Thousand-grain weight QTL, detected in an advanced back cross between Nipponbare as recipient and Xihui 18 as the donor parent, was linked with RM6845. In addition, it is 169.5 kb away from RM3480, the flanking marker of *qTGW8.1* in the present study. *qGL-8,* a novel QTL for grain length, was identified at 22.70–22.05 Mb region at chromosome 8 using a set of recombinant inbred sister lines and high-yielding hybrid rice variety Nei2You No. 6 (Kang et al., 2021). Three QTLs for grain size-related traits, viz., *qGL8* (RM408–RM3702), *qLWR8,* and *qGT8* (RM3845–RM6948), were identified in chromosome 8 in F_2 and $F_{2:3}$ populations derived from an introgression line 'IL188' of Nipponbare/*Oryza minuta* cross (Feng et al., 2021). Thus, this consistent effect region on chromosome 8 seems important for several grain traits and is worthy of further fine mapping, cloning, and functional analysis, similar to that shown in a recent study using CSSLs (Wang et al., 2021). Most of the previous reports detected grain weight QTL in different genomic region compared to *qTGW8.1* of our study which is located between 24867542 to 27481361bp region [QTARO database (affrc.go.jp)] at marker interval of RM3480-RM3452, covering a 2.6 Mb region in the long arm of chromosome 8.

Another TGW QTL, *qTGW6.1,* was mapped in both $F_{2:3}$ and $F_{2:4}$ with LOD value of 2.51–3.52 ($F_{2:4}$) and PV of 9.33–15.19% ($F_{2:4}$) in the present study. Ishimaru (2003) reported that the *qTGW6.1* in backcross inbred lines was derived from cross between *japonica* (Nipponbare)/*indica* (Kasalath) after over 3 years of testing. *qTGW6.1,* at marker interval of RM402-RM5963, was detected in the IL population of ZGX1 × IR75862 (Zhang et al., 2018). One of the flanking markers of this QTL RM3452 was reported to be linked with grain width in the F_2 population of Kasalath × BG23 (Segami et al., 2016) and was also associated with leaf traits (Yang et al., 2018). Zhu et al. (2019) also identified thousand-grain weight QTLs on chromosomes 6 and 8 in three recombinant inbred line populations of Teqing/IRBB lines (TI), Zhenshan97/Milyang46 (ZM), and Xieqingzao/Milyang46 (XM).

Quantitative Trait Loci Clusters Detected

In this study, several QTL clusters were detected, and some of the clusters contained more than four QTL, signifying a heavily populated region of QTL on the chromosome. Xie et al. (2008) detected a cluster of yield-related QTLs on chromosome 9, in BC_3F_4 population from an interspecific cross of Hwaseong/*O. rufipogon* (IRGC 105491). Jeon et al. (2021) employed map-based cloning, gene sequencing, expression analysis, and transgenic approaches to detect putative genes and the genetic architecture of the cluster. The study using BC_4F_4 NILs of the same population demonstrated that the cluster is regulated by a single pleiotropic gene (*ascorbate peroxidase* gene, APX9). Multiple traits mapping in the same region of a chromosome are beneficial to breeders, as QTL clusters allow breeders to focus their efforts on regions of the genome containing the most QTLs of interest. The identification of gene clusters is useful in marker-assisted selection since the markers delineating these regions can be chosen for selecting the traits of interest. Our study demonstrated that favorable alleles from stable wild introgression lines with minimum linkage drag can be introduced to develop high-yielding lines.

CONCLUSION

Inter-specific crosses are bridging tools to introgress beneficial alleles from wild donor into cultivar background to get improved lines for breeding and cultivation. Alien introgression lines/BILs were used as a crossing material in this study, which is a novel approach to detect QTLs and is further useful for QTL/gene pyramiding. In this study, 21, 30, and 17 QTLs were mapped in F_2, $F_{2:3}$, and $F_{2:4}$, derived from a two BILs for a total of 20, 18, and 8 yield-related traits, respectively. This breeding strategy helped to converge beneficial alleles from both the parents to produce superior lines. Five lines, $C_2$12, $C_2$124, $C_2$128, $C_2$143, and $C_2$162 were significantly higher yielding than Swarna and 14s and carried alleles for *qTGW8.1* from *O. nivara* in either a homozygous or heterozygous condition. The F_2, $F_{2:3}$, and $F_{2:4}$ QTL mapping helped to identify 8 major QTLs, namely, *qPH1.1*, *qPH6.1*, *qTGW6.1*, *qTGW8.1, qTN6.1, qPW3.1, qCL6.1*, and *qFG3.1* for yield-related traits. *qTGW8.1* is found as a potential candidate region of 602.4 kb for genetic dissection. The significantly different lines with major yield QTLs will be utilized in further fine mapping and identification of genes for yield improvement. Populations generated from these BIL/CSSL are useful to detect the candidate gene location for important agronomical traits by further fine mapping. This strategy of intercrossing advanced back cross introgression lines can be employed in other backgrounds with advanced genomics tools for crop improvement programs. The statistically significant lines identified by pair-wise mean comparison with both the parents and their common cultivar parent Swarna is useful in gene discovery for target traits and to mine possibly better alleles from the wild progenitors for further crop improvement.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

SN and DB conceptualized, designed, and supervised the experiments, and contributed to the final revision of the manuscript. KB performed the experiments. KA, MS, and VR supported conducting the field experiments. KB and DB analyzed the data and wrote the manuscript with contributions from SN. All authors contributed to the article and approved the submitted version.

FUNDING

This research was part of Ph.D. work of KB and was conducted in the project (ABR/CI/BT/11) on Mapping quantitative trait loci (QTLs) for yield and related traits using backcross inbred lines from elite × wild crosses of rice (*Oryza sativa* L.) as part of ICAR-National Professor Project (F.No.: Edn/27/4/NP/2012-HRD) (2013–2020) funded by Indian Council of Agricultural Research, New Delhi, India. The support of P. Sudhakar and A. Krishna Satya, Department of Biotechnology, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India, and Director, ICAR-IIRR, Hyderabad is acknowledged.

ACKNOWLEDGMENTS

We thank the reviewers for valuable comments that helped improve the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.790221/full#supplementary-material

REFERENCES

Balakrishnan, D., Subrahmanyam, D., Badri, J., Raju, A. K., Rao, V. Y., Kavitha, B., et al. (2016). Genotype × environment interactions of yield traits in backcross introgression lines derived from *Oryza sativa* cv. Swarna /*Oryza nivara*. *Front. Plant Sci.* 7:1530. doi: 10.3389/fpls.2016.01530

Balakrishnan, D., Surapaneni, M., Yadavalli, V. R., Addanki, K. R., Mesapogu, S., Beerelli, K., et al. (2020). Detecting CSSLs and yield QTLs with additive, epistatic and QTL× environment interaction effects from *Oryza sativa* × *O. nivara* IRGC81832 cross. *Sci. Rep.* 10, 1–17.

Bhatia, D., Wing, R. A., Yu, Y., Chougule, K., Kudrna, D., Lee, S., et al. (2018). Genotyping by sequencing of rice interspecific backcross inbred lines identifies QTLs for grain weight and grain length. *Euphytica* 214:41.

Biswas, P. S., Khatun, H., Das, N., Sarker, M. M., and Anisuzzaman, M. (2017). Mapping and validation of QTLs for cold tolerance at seedling stage in rice from an indica cultivar *Habiganj boro* VI (Hbj.BVI). *3 Biotech* 7:359. doi: 10.1007/s13205-017-0993-1

Brar, D. S., and Singh, K. (2011). "*Oryza*," in *Wild Crop Relatives: Genomic and Breeding Resources: Cereals*, ed. C. Kole (Berlin: Springer), 321–365.

Brondani, C., Rangel, P. H. N., Brondani, R. P. V., and Ferreira, M. E. (2002). QTL mapping and introgression of yield-related traits from *Oryza glumaepatula* to cultivated rice (*Oryza sativa*) using microsatellite markers. *Theor. Appl. Genet.* 104, 1192–1203. doi: 10.1007/s00122-002-0869-5

Deng, Z., Cui, Y., Han, Q., Fang, W., Li, J., and Tian, J. (2017). Discovery of consistent QTLs of wheat spike-related traits under nitrogen treatment at different development stages. *Front. Plant Sci.* 8:2120. doi: 10.3389/fpls.2017.02120

Doyle, J. J., and Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus* 12, 13–15.

Fan, F., Long, W., Liu, M., Yuan, H., Pan, G., Li, N., et al. (2019). Quantitative trait locus mapping of the combining ability for yield-related traits in wild rice *Oryza longistaminata*. *J. Agric. Food Chem.* 67, 8766–8772. doi: 10.1021/acs.jafc.9b02224

FAO, IFAD, UNICEF, WFP, and WHO (2017). *The State of Food Security and Nutrition in the World 2017. Building Resilience for Peace and Food Security*. Rome: Food and Agriculture Organization.

Feng, B., Chen, K., Cui, Y., Wu, Z., Zheng, T., Zhu, Y., et al. (2018). Genetic dissection and simultaneous improvement of drought and low nitrogen tolerances by designed QTL pyramiding in rice. *Front Plant Sci.* 9:306. doi: 10.3389/fpls.2018.00306

Feng, Y., Yuan, X., Wang, Y., Yang, Y., Zhang, M., Yu, H., et al. (2021). Validation of a QTL for grain size and weight using an introgression line from a cross between *Oryza sativa* and *Oryza minuta*. *Rice* 14, 1–12.

Feng, Y., Zhai, R.-R., Lin, Z. C., Cao, L. Y., Wei, X. H., and Cheng, S. H. (2015). Quantitative trait locus analysis for rice yield traits under two nitrogen levels. *Rice Sci.* 22, 108-115.

Fisher, K. S., Barton, J., Khush, G. S., Leung, H., and Cantrell, R. (2000). Genomics and agriculture. Collaborations in rice. *Science* 290, 279–280. doi: 10.1126/science.290.5490.279

Gaikwad, K. B., Singh, N., Kaur, P., Rani, S., Babu, P. H., and Singh, K. (2021). Deployment of wild relatives for genetic improvement in rice (*Oryza sativa*). *Plant Breed.* 140, 23–52. doi: 10.1111/pbr.12875

Ganeshan, P., Ajay, J., Parmar, B., Rao, A. R., Sreenu, K., Mishra, P., et al. (2016). Identification of salt tolerant rice lines among interspecific BILs developed by crossing *Oryza sativa* /*O. rufipogon* and *O. sativa* /*O. nivara*. *Austral. J. Crop Sci.* 10, 220–228.

Guo, W. Z., Zhang, T. Z., Ding, Y. Z., Zhu, Y. C., Shen, X. L., and Zhu, X. F. (2005). Molecular marker assisted selection and pyramiding of two QTLs for fiber strength in upland cotton. *Acta Genet. Sin.* 32, 1275–1285.

Haritha, G., Swamy, B. P. M., Naik, M. L., Jyothi, B., Divya, B., and Sarla, N. (2018). Yield traits and associated marker segregation in elite introgression lines derived from *O. sativa* ×*O. nivara*. *Rice Sci.* 25, 19–31. doi: 10.1016/j.rsci.2017.11.001

Ishimaru, K. (2003). Identification of a locus increasing rice yield and physiological analysis of its function. *Plant Physiol.* 133, 1083–1090. doi: 10.1104/pp.103.027607

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, 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

Jiao, Y., Wang, Y., Xue, D., Wang, J., Yan, M., Liu, G., et al. (2010). Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nat. Genet.* 42, 541–544. doi: 10.1038/ng.591

Juneja, S., Das, A., Joshi, S., Sharma, S., Vikal, Y., Patra, B., et al. (2006). *Oryza nivara* (Sharma et Shastry) the progenitor of *O. sativa* (L.) subspecies indica harbours rich genetic diversity as measured by SSR markers. *Curr. Sci.* 91, 1079–1085.

Kaladhar, K., Swamy, B. P. M., Babu, A. P., Reddy, C. S., and Sarla, N. (2008). Mapping quantitative trait loci for yield traits in BC2F2 population derived from Swarna x *O. nivara* cross. *Rice Genet. Newslett.* 24, 34–36.

Kang, Y., Zhang, M., Zhang, Y., Wu, W., Xue, P., Zhan, X., et al. (2021). Genetic mapping of grain shape associated qtl utilizing recombinant inbred sister lines in high yielding rice (*Oryza sativa* L.). *Agronomy* 11:705. doi: 10.3390/agronomy11040705

Kang, Y. J., Shim, K. C., Lee, H. S., Jeon, Y. A., Kim, S. H., Kang, J. W., et al. (2018). Fine mapping and candidate gene analysis of the quantitative trait locus gw8. 1 associated with grain length in rice. *Genes Genomics* 40, 389–397. doi: 10.1007/s13258-017-0640-6

Kavitha, B., Divya, B., Malathi, S., Krishnamraju, A., Venkateshwara Rao, Y., and Sarla, N. (2019). Evaluation of yield and seedling vigour related traits of Swarna/*Oryza nivara* backcross introgression lines under three environment conditions. *Ecol. Genet. Genomics* 11:100036.

Kim, C.-K., Chu, S.-H., Park, H. Y., Seo, J., Kim, B., Lee, G., et al. (2017). Identification of heterosis QTLs for yield and yield-related traits in indica-japonica recombinant inbred lines of rice (*Oryza sativa* L.). *Plant Breed. Biotechnol.* 5, 371–389. doi: 10.9787/pbb.2017.5.4.371

Kim, D. M., Lee, H. S., Kwon, S. J., Fabreag, M. E., Kang, J. W., Yun, Y. T., et al. (2014). High-density mapping of quantitative trait loci for grain-weight and spikelet number in rice. *Rice* 7, 1–11. doi: 10.1186/s12284-014-0014-5

Kumar, N., Jain, R. K., and Chowdhury, V. K. (2019). Linkage mapping of QTLs for grain minerals (iron and zinc) and physio-morphological traits for

development of mineral rich rice (*Oryza sativa* L.). *Indian J. Biotechnol.* 18, 69–80.

Li, Z. K., Yu, S. B., Lafitte, H. R., Huang, N., Courtois, B., Hittalmani, S., et al. (2003). QTL x environment interactions in rice. I. Heading date and plant height. *Theor. Appl. Genet.* 108, 141–153. doi: 10.1007/s00122-003-1401-2

Lim, J. H., Yang, H. J., Jung, K. H., Yoo, S. C., and Paek, N. C. (2014). Quantitative trait locus mapping and candidate gene analysis for plant architecture traits using whole genome re-sequencing in rice. *Mol. Cells* 37, 149–160. doi: 10.14348/molcells.2014.2336

Lin, T., Zhou, C., Chen, G., Yu, J., Wu, W., Ge, Y., et al. (2020). Heterosis-associated genes confer high yield in super hybrid rice. *Theor. Appl. Genet.* 133, 3287–3297. doi: 10.1007/s00122-020-03669-y

Ma, X., Fu, Y., Zhao, X., Jiang, L., Zhu, Z., Gu, P., et al. (2016). Genomic structure analysis of a set of *Oryza nivara* introgression lines and identification of yield-associated QTLs using whole-genome resequencing. *Sci. Rep.* 6:27425. doi: 10.1038/srep27425

Mahmoud, A. A., Sukumar, S., and Krishnan, H. B. (2008). Interspecific rice hybrid of *Oryza sativa* x *Oryza nivara* reveals a significant increase in seed protein content. *J. Agric. Food Chem.* 56, 476–482. doi: 10.1021/jf071776n

Matsubara, K., Yamamoto, E., Kobayashi, N., Ishii, T., Tanaka, J., Tsunematsu, H., et al. (2016). Improvement of rice biomass yield through QTL-based selection. *PLoS one* 11:e0151830. doi: 10.1371/journal.pone.0151830

Miura, K., Ikeda, M., Matsubara, A., Song, X. J., Ito, M., Asano, K., et al. (2010). OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nat. Genet.* 42, 545–549. doi: 10.1038/ng.592

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, 433–443. doi: 10.1007/s12041-013-0285-4

Pang, Y., Chen, K., Wang, X., Wang, W., Xu, J., Ali, J., et al. (2017). Simultaneous improvement and genetic dissection of salt tolerance of rice (*Oryza sativa* L.) by designed QTL pyramiding. *Front. Plant Sci.* 8:1275. doi: 10.3389/fpls.2017.01275

Qiao, W., Qi, L., Cheng, Z., Su, L., Li, J., Sun, Y., et al. (2016). Development and characterization of chromosome segment substitution lines derived from *Oryza rufipogon* in the genetic background of *O. sativa* spp. indica cultivar 93-11. *BMC Genomics* 17:580. doi: 10.1186/s12864-016-2987-5

Rabiei, B., Kordrostami, M., Sabouri, A., and Sabouri, H. (2015). Identification of QTLs for yield related traits in indica type rice using SSR and AFLP Markers. *Agric. Conspectus Sci.* 80, 91–99.

Sabouri, H., Sabouri, A., and Dadras, A. R. (2009). Genetic dissection of biomass production, harvest index and panicle characteristics in indica-indica crosses of Iranian rice (*Oryza sativa* L.) cultivars. *Austral. J. Crop Sci.* 3, 155–166.

Salgotra, R. K., Gupta, B. B., Bhat, J. A., and Sharma, S. (2015). Genetic diversity and population structure of Basmati rice (*Oryza sativa* L.) germplasm collected from North Western Himalayas using trait linked SSR markers. *PLoS One* 10:e0131858. doi: 10.1371/journal.pone.0131858

Sarla, N. (2014). DRR Dhan40- with yield enhancing QTLs from wild species. *DRR Newsletter* 12:2.

Sarla, N., Bobba, S., and Siddiq, E. A. (2003). ISSR and SSR markers based on AG and GA repeats delineate geographically diverse *Oryza nivara* accessions and reveal rare alleles. *Curr. Sci.* 84, 683–690.

Sato, H., Ideta, I. O., Ando, I., Kunihiro, Y., Hirabayashi, H., Iwano, M., et al. (2004). Mapping QTLs forsheath blight resistance in the rice line WSS2. *Breed. Sci.* 54, 265–271.

Segami, S., Yamamoto, T., Oki, K., Noda, T., Kanamori, H., Sasaki, H., et al. (2016). Detection of novel QTLs regulating grain size in extra-large grain rice (*Oryza sativa* L.) lines. *Rice* 9:34. doi: 10.1186/s12284-016-0109-2

Shao, G. N., Tang, S. Q., Luo, J., Jiao, G. A., Tang, A., and Hu, P. S. (2009). QTL analysis for flag leaf and grain shape and populations construction derived from related residual heterozygous lines in rice. *Plant Mol. Breed.* 7, 16–22.

Shen, B., Yu, W. D., Du, J. H., Fan, Y. Y., Wu, J. R., and Zhuang, J. Y. (2011). Validation and dissection of quantitative trait loci for leaf traits in interval RM4923-RM402 on the short arm of rice chromosome 6. *J. Genet.* 90, 39–44. doi: 10.1007/s12041-011-0019-4

Singh, V. K., Ellur, R. K., Singh, A. K., Nagarajan, M., Singh, B. D., and Singh, N. K. (2018). Effect of qGN4.1 QTL for grain number per panicle in genetic backgrounds of twelve different mega varieties of rice. *Rice* 11:8. doi: 10.1186/s12284-017-0195-9

Standard Evaluation System for Rice [SES] (2014). *IRRI*. Manila: SES.

Su, C., Lu, W., Zhao, T., and Gai, J. (2010). Verification and fine-mapping of QTLs conferring days to flowering in soybean using residual heterozygous lines. *Chinese Sci. Bull.* 55, 499–508. doi: 10.1007/s11434-010-0032-7

Sun, P., Zheng, Y., Li, P., Ye, H., Zhou, H., Gao, G., et al. (2019). Dissection and validation of minor quantitative trait loci (QTLs) conferring grain size and weight in rice. *bioRxiv* [Preprint]. doi: 10.1101/511139

Surapaneni, M., Divya, B., Sukumar, M., Raju, A. K., Rao, Y. V., Tripura Venkata, V. G. N., et al. (2017). Identification of major effect QTLs for agronomic traits and CSSLs in rice from Swarna/*O nivara* derived backcross inbred lines. *Front. Plant Sci.* 8:1027. doi: 10.3389/fpls.2017.01027

Swamy, B. P., Kaladhar, K., Reddy, G. A., Viraktamath, B. C., and Sarla, N. (2014). Mapping and introgression of QTL for yield and related traits in two backcross populations derived from *Oryza sativa* cv. Swarna and two accessions of O.nivara. *J. Genet.* 93, 643–654. doi: 10.1007/s12041-014-0420-x

Swamy, B. P. M., Kaladhar, K., Ramesha, M. S., Viraktamath, B. C., and Sarla, N. (2011). Molecular mapping of QTLs for yield and related traits in *Oryza sativa* cv Swarna x *O. nivara* (IRGC81848) backcross population. *Rice Sci.* 18, 178–186. doi: 10.1016/s1672-6308(11)60025-5

Swamy, B. P. M., Kaladhar, K., Shobha Rani, N., Prasad, G. S. V., Viraktamath, B. C., Reddy, A. G., et al. (2012). QTL analysis for grain quality traits in 2 BC2F2 populations derived from crosses between *Oryza sativa* cv Swarna and 2 accessions of *O.nivara*. *J. Hered.* 103, 442–452. doi: 10.1093/jhered/esr145

Swamy, B. P. M., and Sarla, N. (2008). Yield enhancing quantitative trait loci (QTLs) from wild 19 species. *Biotechnol. Adv.* 26, 106–120.

Swamy, B. P. M., and Sarla, N. (2011). Meta-analysis of yield QTLs derived from inter-specific crosses of rice reveals consensus regions and candidate genes. *Plant Mol. Biol. Rep.* 29, 663–680.

Tian, F., Li, D. J., Fu, Q., Zhu, Z. F., Fu, Y. C., Wang, X. K., et al. (2006). Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. *Theor. Appl. Genet.* 112, 570–580. doi: 10.1007/s00122-005-0165-2

Tian, Y., Zhang, H., Xu, P., Chen, X., Liao, Y., Han, B., et al. (2015). Genetic mapping of a QTL controlling leaf width and grain number in rice. *Euphytica* 20, 1–11. doi: 10.1007/s10681-014-1263-5

Vaughan, D. A., Lu, B. R., and Tomooka, N. (2008). The evolving story of rice evolution. *Plant Sci.* 174, 394–408. doi: 10.1016/j.plantsci.2008.01.016

Vaughan, D. A., Morishima, H., and Kadowaki, K. (2003). Diversity in the *Oryza* genus. *Curr. Opin. Plant Biol.* 6, 139–146.

Verma, R. K., Chetia, S. K., Dey, P. C., Baruah, A. R., and Modi, M. K. (2017). Mapping of QTLs for grain yield and its component traits under drought stress in elite rice variety of assam. *Int. J. Curr. Microbiol. Appl. Sci.* 6, 1443–1455.

Wang, D., Sun, W., Yuan, Z., Sun, Q., Fan, K., Zhang, C., et al. (2021). Identifcation of a novel QTL and candidate gene associated with grain size using chromosome segment substitution lines in rice. *Sci. Rep.* 11:189. doi: 10.1038/s41598-020-80667-6

Wang, H., Zhang, X., Yang, H., Liu, X., Li, H., Yuan, L., et al. (2016). Identifcation of heterotic loci associated with grain yield and its components using two CSSL test populations in maize. *Sci. Rep.* 6:38205. doi: 10.1038/srep38205

Wang, J. (2009). Inclusive composite interval mapping of quantitative trait genes. *Acta Agron. Sin.* 35, 239–245.

Wang, S., Cui, G., Wang, H., Ma, F., Xia, S., Li, Y., et al. (2018). Identification and QTL mapping of Z550, a rice backcrossed inbred line with increased grains per panicle. *J. Integr. Agric.* 17, 60345–60347.

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

Wickneswari, R., Bhuiyan, M. A. R., Sabu, K. K., Lim, L. S., Thomson, M. J., Narimah, M. K., et al. (2012). Identification and validation of quantitative trait loci for agronomic traits in advanced backcross breeding lines derived from *Oryza rufipogon* x *Oryza sativa* cultivar MR219. *Plant Mol. Biol. Rep.* 30, 929–939.

Xie, X., Jin, F., Song, M. H., Suh, J. P., Hwang, H. G., Kim, Y. G., et al. (2008). Fine mapping of a yield-enhancing QTL cluster associated with transgressive variation in an *Oryza sativa*× *O. rufipogon* cross. *Theor. Appl. Genet.* 116, 613–622. doi: 10.1007/s00122-007-0695-x

Xie, X., Song, M. H., Jin, F., Ahn, S. N., Suh, J. P., Hwang, H. G., et al. (2006). Fine mapping of a grain weight quantitative trait locus on rice chromosome 8 using near-isogenic lines derived from a cross between *Oryza sativa* and

Oryza rufipogon. *Theor. Appl. Genet.* 113, 885–894. doi: 10.1007/s00122-006-0348-5

Xing, W., Zhao, H., and Zou, D. (2014). Detection of main-effect and epistatic QTL for yield-related traits in rice under drought stress and normal conditions. *Can. J. Plant Sci.* 94, 633–641.

Xue, W., Xing, Y., Weng, X., Zhao, Y., Tang, W., Wang, L., et al. (2008). Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. *Nat. Genet.* 40, 761–767. doi: 10.1038/ng.143

Yang, J., Luo, D., Yang, B., Frommer, W. B., and Eom, J. S. (2018). SWEET11 and 15 as key players in seed filling in rice. *New Phytol.* 218, 604–615. doi: 10.1111/nph.15004

Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., et al. (2000). Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene constans. *Plant Cell* 12, 2473–2484. doi: 10.1105/tpc.12.12.2473

Zhang, J., Ou, X., Hu, H., Du, B., Lv, W., Yang, L., et al. (2018). Identification of QTLs for yield-related traits using two sets of introgression lines with a common donor parent in rice. *Int. J. Agric. Biol.* 20, 15–24.

Zhang, L., Wang, S., Li, H., Deng, Q., Zheng, A., Li, S., et al. (2010). Effects of missing marker and segregation distortion on QTL mapping in F2 populations. *Theor. Appl. Genet.* 121, 1071–1082. doi: 10.1007/s00122-010-1372-z

Zhang, S., He, X., Zhao, J., Cheng, Y., Xie, Z., Chen, Y., et al. (2017). Identification and validation of a novel major QTL for harvest index in rice (*Oryza sativa* L.). *Rice* 10:44. doi: 10.1186/s12284-017-0183-0

Zhu, Y. J., Sun, Z. C., Niu, X. J., Ying, J. Z., Fan, Y. Y., Mou, T. M., et al. (2019). Dissection of three quantitative trait loci for grain size on the long arm of chromosome 10 in rice (*Oryza sativa* L.). *PeerJ* 7:e6966. doi: 10.7717/peerj.6966

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 in Plant Science

ORIGINAL RESEARCH
published: 11 March 2022
doi: 10.3389/fpls.2022.767635

Reticulate Evolution in AA-Genome Wild Rice in Australia

*Sharmin Hasan[1,2], Agnelo Furtado[1] and Robert Henry[1]**

[1] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Brisbane, QLD, Australia,
[2] Department of Botany, Jagannath University, Dhaka, Bangladesh

The wild rice gene pool, i.e., AA-genome, in Australia is geographically and genetically distinct from that in Asia. Two distinct taxa are found growing together in northern Australia, *Oryza meridionalis* (including annual and perennial forms) and an *Oryza rufipogon* like taxa that have been shown to have a chloroplast genome sequence that is closer to that of *O. meridionalis* than to *O. rufipogon* from Asia. Rare plants of intermediate morphology have been observed in the wild despite a reported reproductive barrier between these two species. We now report the resequencing of plants from 26 populations including both taxa and putative hybrids. A comparison of chloroplast and nuclear genome sequences indicated re-combinations that demonstrated hybridisation in both directions. Individuals with intermediate morphology had high nuclear genome heterozygosity consistent with a hybrid origin. An examination of specific genes (e.g., starch biosynthesis genes) revealed the presence of heterozygotes with alleles from both parents suggesting that some wild plants were early generation hybrids. These plants may have low cross-fertility preserving the continuation of the two distinct species. Repeated backcrossing of these rare hybrids to one parent would explain the plants exhibiting chloroplast capture. These observations suggest that reticulate evolution is continuing in wild *Oryza* populations and may have been a key process in rice evolution and domestication.

Keywords: hybridisation, introgression, chloroplast capture, reproductive barrier, reticulate evolution

OPEN ACCESS

Edited by:
Yohei Koide,
Hokkaido University, Japan

Reviewed by:
Mulatu Geleta,
Swedish University of Agricultural Sciences, Sweden
Peng Qi,
University of Georgia, United States

****Correspondence:***
Robert Henry
robert.henry@uq.edu.au

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

Received: *31 August 2021*
Accepted: *04 February 2022*
Published: *11 March 2022*

Citation:
Hasan S, Furtado A and Henry R (2022) Reticulate Evolution in AA-Genome Wild Rice in Australia. Front. Plant Sci. 13:767635. doi: 10.3389/fpls.2022.767635

INTRODUCTION

Hybridisation and introgression undoubtedly play a crucial role in evolution allowing gene exchange between species resulting in genomic diversification and novel genetic combinations (Stebbins, 1959). Cytonuclear non-concordance, i.e., incongruence between organellar (plastid and mitochondrial) and nuclear genome phylogenies appears to reflect both hybridisation and successive introgression (Rieseberg, 1991; Nge et al., 2021). Recurrent hybridisation results in contrasting gene trees in two ways: (1) if recurrent unidirectional cytoplasmic geneflow from one species can invade another species in the absence of nuclear geneflow (Rieseberg and Wendel, 1993), and (2) if unidirectional nuclear geneflow but not cytoplasmic geneflow is allowed to invade a species (Petit et al., 2001). However, cytoplasmic genome invasion occurs in plants more frequently

than nuclear genome introgression (Rieseberg and Wendel, 1993). Cytoplasmic geneflow results in chloroplast capture, which may occur due to partial cytoplasmic male sterility, nuclear genome incompatibilities, or partial selfing in species (Tsitrone et al., 2003). Contrasting gene trees can be attributed to sampling error, evolutionary convergence, evolutionary rate heterogeneity, lineage sorting, and reticulation (Rieseberg and Soltis, 1991). Chloroplast capture has been used to infer reticulate evolution in many plant species such as *Pisum* (Palmer et al., 1985), Australian cotton (Wendel et al., 1991), sunflower (Rieseberg, 1991), peonies (Sang et al., 1995), *Opuntia* (Griffith, 2003), soybeans (Doyle et al., 2003); kiwifruit (Chat et al., 2004), the wheat tribe (*Elymus*) (Mason-Gamer, 2004); baobabs (Karimi et al., 2019); and *Adenanthos* (Nge et al., 2021). Much effort has been put into re-constructing reticulate evolution by means of investigating putative hybrid origins through multiple molecular techniques (Vriesendorp and Bakker, 2005). Morphological features along with geographical, ecological, biosystematics, and cytological data have been used to infer hybrid origin. However, genetic evidence provides the greatest power of inference in detecting putative interspecific hybrid origin (Griffith, 2003). Genetic evidence can be derived through implying an additivity of molecular markers (Wendel et al., 1991; Griffith, 2003), polymorphic nucleotide additivity at a single position, i.e., ITS additivity (Sang et al., 1995), and incongruence between gene trees (Mason-Gamer, 2004). Next-generation sequencing (NGS) approaches appear to be increasingly effective in addressing the causes of cytonuclear incongruence (Lemmon et al., 2012; Weitemier et al., 2014).

Rice (the genus *Oryza*) belongs to the tribe Oryzeae—a member of the grass subfamily Ehrhartoideae. The genus *Oryza* comprises 11 genome groups of which the AA-genome group (with eight diploid species) is the most diverged group, distributed throughout the world, except for, Antarctica (Vaughan, 1989). The other genome groups are BB (one *Oryza* species), BBCC (three *Oryza* species), CC (three *Oryza* species), CCDD (three *Oryza* species), EE (one *Oryza* species), FF (one *Oryza* species), GG (three *Oryza* species), HHJJ (two *Oryza* species), HHKK (one *Oryza* species), and KKLL (one *Oryza* species) (Brar and Khush, 2018). Vicariance events and long-distance dispersal explain the evolution and divergence of AA genomes globally (Tang et al., 2010). The estimated mean divergence age of the AA genome from other *Oryza* genomes is 2.41 million years ago (Stein et al., 2018). Four *Oryza* species have been recognised in Australia (Henry et al., 2010). Of these, two wild *Oryza* species belong to the AA-genome clade: *Oryza meridionalis* Ng. and *Oryza rufipogon* Griff. The former was originally characterised by annual life history, self-pollination, and short anther length (Ng et al., 1981). More recently, the *O. meridionalis* gene pool has been shown to include both annual and perennial forms (Sotowa et al., 2013; Henry, 2019). *O. rufipogon* is considered a native perennial in Northern Queensland, Australia (Henry et al., 2010). Australian *O. rufipogon* populations have been shown to have a chloroplast genome that is closer to *O. meridionalis* than to *O. rufipogon* from Asia (Brozynska et al., 2014). This indicates that the Australian population might need to be considered as separate taxon (Brozynska et al., 2017).

Reproductive barriers appear to be common in the AA genome *Oryza* species, especially in crosses between *O. sativa* and its wild relatives (including *O. meridionalis*), resulting in hybrid sterility (Li et al., 2007). Hybridisation between *O. meridionalis* and Asian wild rice produced low fertility hybrids (Naredo et al., 1997). Isolation due to pollen sterility was identified in *O. sativa* L. × *O. meridionalis* hybrids, resulting in partial and full abortion of male gametes (Li et al., 2018). Post-zygotic reproductive isolation happens to abort seeds in first filial generation (F1) hybrid offspring derived from hybridisation between *O. meridionalis,* and *O. sativa* (Toyomoto et al., 2019). Furthermore, geographic distributions and differences in flowering times influence intra-and interspecific crossability of *O. meridionalis* in Australia (Juliano et al., 2005).

Cytonuclear discordant in phylogenetic trees (i.e., chloroplast capture) has been observed in these wild rice populations, which were assumed to be due to hybridisation between the *O. meridionalis* and *O. rufipogon* type taxa (Moner et al., 2018). These wild hybrid plants were identified in the field due to their intermediate morphological characteristics with combinations of spike morphology and anther size, suggesting the inheritance of traits from both parental taxa. However, the exact nature of their hybrid origin was not characterised by nuclear genome analysis.

We have now reexamined the origins of these hybrids by re-sequencing and analysing the nuclear genomes of these wild rice populations including wild plants of hybrid appearance.

MATERIALS AND METHODS

Sample Collection

Australian wild rice is widely distributed from the south of Townsville to the northern parts of Cape York Peninsula, Queensland, Australia. Fresh leaves of wild rice from 26 different geographical sites were collected in the two consecutive years 2015 and 2016 by Moner et al. (2018) (**Figure 1**). Only one sample was collected from each site. Among these, 22 samples were collected in 2015 from 22 different sites in the northern parts of Cape York Peninsula while only four samples were collected in 2016 from the south of Townsville. The samples were kept in the dry ice all the way from the sampling sites to the laboratory until stored at −80°C freezer.

DNA Extractions, Sequencing, and Quality Filtering

Fresh leaves were pulverised with tissue lyser (Qiagen, United States). Samples were kept in liquid nitrogen to prevent thawing throughout the pulverisation process. DNA was extracted from pulverised leaf tissue of wild rice samples following the CTAB method (Furtado, 2014). The quality and quantity of the extracted DNA were evaluated with a NanoDrop spectrophotometer (Thermo Fisher Scientific, Delaware, United States) as well as by electrophoresis on a 0.7% agarose gel stained with SYBR safe. The 260/280 nm ratio for 26 samples ranged from 1.11 to 2.06. Nextera DNA Flex

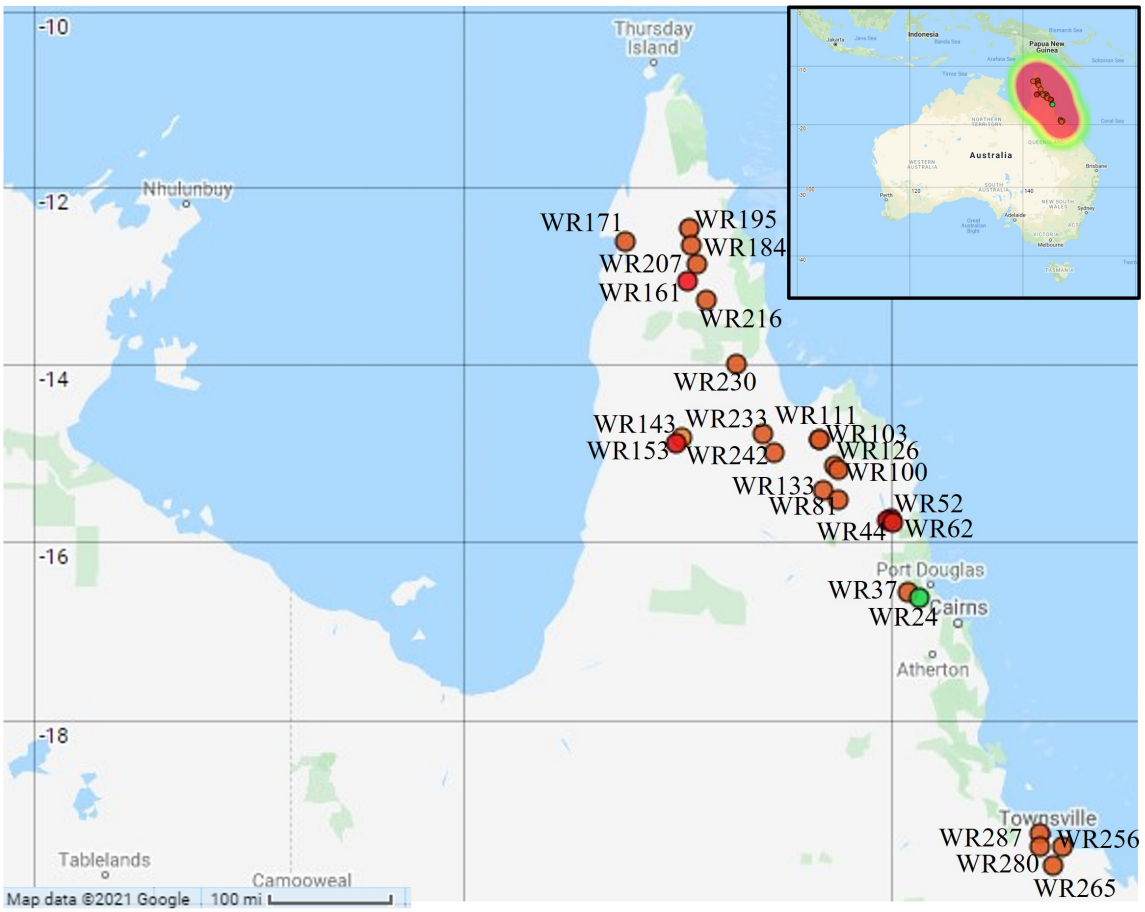

FIGURE 1 | Geographical locations of the collected Australian wild rice samples: orange dot = *Oryza meridionalis*; green dot = *Oryza rufipogon* type taxa; red dot: hybrid. Map created with Map Maker at https://maps.co (Map Maker, 2021).

Libraries were prepared. Nextera DNA Library Preparation Kits were designed in an automation-compatible workflow that introduced tagmentation chemistry, which joined DNA fragmentation and adapter ligation steps into a single 15-min reaction. Then these libraries were sequenced on a NovaSeq 6000 SP and S4 Flow cell along with other samples to produce 2 × 150 paired-end reads with a data yield of 20x whole genome coverage on average. A quality check was applied to the raw data using CLC Genomic Workbench (CLC-GWB) version 20 software (QIAGEN CLC Genomics Workbench 20, Denmark)[1] to identify any sequencing error before performing further analysis. Sequences were trimmed at both 0.05 and 0.01 quality scores to truncate low-quality reads. Based on the percentage of loss of reads and bases, 0.01 trimmed reads were considered for further analysis. The total Illumina raw reads for the 26 samples ranged from 56,349,078 to 104,325,654 bp. The total raw nucleotides ranged between 8,416,673,405 and 15,628,960,467 bp. After trimming, the total 0.01 trimmed reads ranged from 54,631,067 to 98,924,768 and nucleotides from 7,925,007,230 to 14,260,593,039 bp (**Supplementary Table 1**).

[1]http://www.clcbio.com

Illumina Reads Mapping and Genome-Wide Variant Detection

Prior to final mapping, the best mapping parameters were evaluated with the three different alignment quality thresholds: 1.0 length fraction (LF) and 0.95 similarity fraction (SF), 1.0 LF and 0.90 SF and 1.0 LF and 0.85 SF. Illumina paired-end reads were then mapped to the reference genome Os-Nipponbare-Reference-IRGSP-1, which was downloaded at Rice Annotation Project Database (Kawahara et al., 2013) with these three settings. The average mapped read depth (times of reference length), mapped reads (%), mapped bases (%), and total consensus length as a % of reference length were calculated from the output of mapped reads to make a comparison among the three different settings. After comparison, the 1 LF and 0.9 SF settings were found to be stringent enough to consider for final mapping. A total of 26 wild rice Illumina sequences quality trimmed at 0.01 (Phred equivalent > 20, 94–98% of the reads had a phred score greater than 35, average across all bases) was mapped to the Os-Nipponbare-Reference-IRGSP-1, separately using the tool "Map Reads to Reference" in CLC-GWB, with the following

parameters: masking mode = no masking; match score = 1; mismatch cost = 2; cost of insertions and deletions = linear gap cost; insertion cost = 3; deletion cost = 3; length fraction = 1; similarity fraction = 0.9; global alignment = no; auto-detected paired distances = yes; non-specific match handling = map randomly. Mapped reads was subjected to variant analysis using the tool "Basic Variant Analysis" in CLC-GWB, with the following parameters: ploidy = 2; ignore positions with coverage above = 100,000; restrict calling to target regions = not set; ignore broken pairs = no; ignore non-specific matches = no; minimum coverage = 10; minimum count = 3; minimum frequency (%) = 25; neighborhood radius = 5; minimum central quality = 20; minimum neighborhood quality = 15; read direction filter = no; read position filter = no; remove pyro-error variants = no. Total variants that included SNV (single nucleotide variant), insertions, deletions, MNV (multiple nucleotide variants), and replacement were determined. Percentage of SNV and InDels (insertions and deletions) were calculated as: (total number of SNV/InDels × 100)/total number of variants detected (1). Heterozygous SNV positions were determined by SNV loci with two alleles. Heterozygosity across the whole genome was calculated as: (total heterozygous SNP position × 100)/the whole genome size (2). Unique heterozygous SNPs were those that were distinctively located in the reads of a sample. The whole-genome size of *O. rufipogon* type taxa (= Taxon A) and *O. meridionalis* (= Taxon B) was 384.8 and 354.9 Mb, respectively (Brozynska et al., 2017). Other data resources for *O. rufipogon* and *O. meridionalis* are available at (https://oryza-ensembl.gramene.org/Oryza_meridionalis/Info/Index, https://oryza-ensembl.gramene.org/Oryza_rufipogon/Info/Index, http://viewer.shigen.info/oryzagenome21detail/downloads/index.xhtml;jsessionid=16d6389e9b7080a8c4fb7378b753). Sequence data for the Australian *O. rufipogon* like taxon is reported in Brozynska et al. (2017).

Nuclear Geneflow and Genetic Structure Analysis

In further experiment, five natural hybrids (WR44, WR52, WR62, WR153, and WR161) were investigated to test for nuclear introgression from the putative parental populations. Among these, WR44, WR52, and WR153 were reported by Moner et al. (2018). The earlier study also identified two hybrids (WR65 and WR162), which were collected from Lakeland Cook town road and Merluna site 1, respectively. Two hybrids (WR62 and WR161), explored in this study, were also collected from the same sites but not reported as hybrids by Moner et al. (2018). These were collected from the same sites as WR65 and WR162, respectively. To ensure the chloroplast genome of WR62 and WR161 were of hybrid origin showing similarity with the previously published samples WR65 and WR162 (Moner et al., 2018), respectively, 0.01 trimmed data of WR62 and WR161 were mapped to the reference chloroplast genome of *O. meridionalis,* and the reference chloroplast genome of *O. rufipogon* type taxa obtained from Brozynska et al. (2014). Mapping was done following the reference-based mapping workflow explained by Moner et al. (2018) using CLC-GWB. Then the consensus sequences of the chloroplast genome were extracted from these two samples (sequences S1 and S2 provided as **Supplementary Material**). Chloroplast genomes of seven samples; WR62, WR161 from this study, WR65 and WR162 (Moner et al., 2018), two reference chloroplast genomes of *O. meridionalis* and *O. rufipogon* type taxa (Brozynska et al., 2014), and *Oryza officinalis* Wall. ex Watt (Moner et al., 2018) were aligned with Multiple Alignment using Fast Fourier Transform (MAFFT) version 7.407 (Katoh and Standley, 2013) with default parameters in Geneious version 11.1.5 software.[2] The substitution model for the phylogenetic analysis was selected by the jModelTest2 (Darriba et al., 2012) on the Extreme Science and Engineering Discovery Environment (XSEDE). The parameters for best selected model following the Akaike information criterion (AIC) was used in maximum likelihood tree construction: Model = TVM+I+G, partition = 012314, −lnL = 1888839.3939, K = 21, feqA = 0.3068, freqC = 0.1941; feqG = 0.1957; freqT = 0.3034; ti/tv = −; p-inv = 0.897; R(a) = 1.0604; R(b) = 2.5331, R(c) = 0.5698; R(d) = 0.9914; R(e) = 2.5331; R(f) = 1; gamma = 0.587. Phylogenetic reconstruction was performed using PAUP* version 4 software (Swofford, 2003) with maximum likelihood (ML) with heuristic searching, retain group with frequency > 50%, bisection, reconnection branch swapping method and 1,000 bootstrap replicates. *O. officinalis* was selected as an outgroup species in the phylogenetic tree.

Single nucleotide polymorphism (SNP) genotypes of six genes located on chromosomes 4, 6, and 7 from hybrids and putative parental taxa were aligned, with respect to Os-Nipponbare-Reference-IRGSP-1 with CLC-GWB to screen for contributing donor parents in genotypes of hybrid populations. These genes include *seed shattering 4* (*sh4*) on chromosome 4, four starch synthesis related genes on chromosome 6 [*granule bound starch synthase I* (*GBSSI*), *soluble starch synthase-I* (*SSI*), *starch branching enzyme I* (*SBEI*), and *alkali degeneration* (*SSIIa* or *ALK*)] and *granule bound starch synthase II* (*GBSSII*) on chromosome 7. Plausible hybridisation generations were determined based on the level of heterozygosity across the whole genome of these hybrids.

Four taxon (ABBA-BABA) test as a measure of geneflow was performed with Hybrid Check version 1 software (Ward and van Oosterhout, 2016). Four genes (*GBSSI*, *SSI*, *SBEI,* and *ALK)* from two putative parental populations (WR24 and WR81), one hybrid and outgroup species were first aligned with MAFFT version 7.407 (Katoh and Standley, 2013) with default parameters in Geneious version 11.1.5 software to generate an aligned sequence. Then, an aligned sequence was used to perform the four-taxon test with a topology (((P1, P2), P3) A), where a triplet comprises of a set of three aligned sequences from a hybrid (P3) and two putative parental populations (P1 and P2). *O. sativa* spp. *japonica* cv. Nipponbare was considered an outgroup species (A). We simulated a total of 5 four-taxon tests to determine biparental nuclear geneflow between hybrid and two putative parental taxa. We estimated Patterson's *D*-statistic based on the random distribution of ABBA-BABA polymorphisms in a four-taxon test (Martin et al., 2015). The *D*-statistic estimates the deviation from the distribution of ABBA-BABA polymorphisms

[2]https://www.geneious.com

in a four-taxon tree. If $D > 0$, introgression occurs between P2 and P3, and if $D < 0$, introgression occurs between P1 and P3. We performed a block jackknife method using a block length of 1,000 bp to calculate Z score and p-value to test significant deviation from the null expectation ($D = 0$, lineage sorting).

Genome-wide SNP data from 26 wild rice samples were used to perform principal component analysis (PCA) to assess the genetic variation in wild rice populations using plink version 1.9 (Purcell et al., 2007). To ensure the SNP data was not biased with any linkages, linkage disequilibrium (LD) based pruning was done to generate a final neutral SNP data for PCA analysis. LD pruning was performed using the "–indep-pairwise" option by specifying a window of 50 kb, with a step window size of 10 bp and a threshold of linkage, $r^2 = 0.2$ (which recursively eliminates SNPs within the sliding window if r^2 is greater than 0.2). To investigate the relationship of population structure with geographic distance, a subset of genome-wide SNP data (2,987,364 SNPs of 12,247,310 SNPs) from 26 wild rice genotypes (two alleles per loci) were used to perform isolation by distance (IBD) analysis using Mantel test (Mantel, 1967) with the R package adegenet version 2.0.1 (Jombart and Ahmed, 2011). Nei (1972) pairwise genetic distances vs. geographic distances were correlated with the function mantel.randtest using a Monte Carlo simulation with 999 permutations in IBD analysis.

RESULTS

Mapping Outputs and Sequence Variation

Sequence variation from *O. sativa* was substantial among the 26 samples (**Supplementary Table 1**). The average sequence read depth (times reference length) ranged from 20.7 in WR153 to 46.9 in WR103 (**Supplementary Table 1**). While an average mapped sequence read depth (times reference length) varied from 8.88 in WR24 to 46.7 in WR103. The percentage of mapped reads was highest in WR44 (72.0%) and the lowest in WR24 (24.0%). Similarly, the highest and lowest percentage of mapped bases were 71.9 and 23.8% in WR44 and WR24, respectively. The total consensus sequences length ranged from 257,716,289 bp in WR37 to 322,074,565 bp in WR24. The total consensus as a percentage of the reference ranged from 67.4% in WR37 to 83% in WR44. The data for eight samples were also mapped to the lower quality draft genomes (Brozynska et al., 2017) for the two Australian taxa (**Supplementary Table 2**) resulting in somewhat higher genome coverage. WR 24 displayed an *O. rufipogon* type nuclear genome and showed much greater coverage of this reference confirming it was closer to *O. sativa* than the *O. meridionalis* types. Further analysis was based upon the use of the common high-quality *O, sativa* genome to align the reads to explore hybridisation and heterozygosity in the wild populations.

Chloroplast and Nuclear Genome Characterisations

The analysis of phylogenetic relationships based on the whole chloroplast and nuclear genome had suggested that chloroplast capture was present in Australian wild rice populations (Moner et al., 2018). The whole chloroplast and nuclear genome of both *O. rufipogon* type taxa and *O. meridionalis* were compared with the chloroplast and nuclear genome of WR62 and WR161 to determine the origin of these samples that had not been characterised in the earlier study of Moner et al. (2018). The results confirmed that these were of hybrid origin as the chloroplast genome of WR62 was grouped with the *O. meridionalis* type (Brozynska et al., 2014) while the nuclear genome was of the *O. rufipogon* type taxa (**Supplementary Figure 1**). While WR161 contained an *O. meridionalis* type nuclear genome and an *O. rufipogon* type chloroplast genome (Brozynska et al., 2014). This reinforced the presence of bidirectional chloroplast capture in hybridising populations in the Australian wild rice gene pool. Hybrids acquired their chloroplast genome from either the *O. rufipogon* type or *O. meridionalis* during hybridisation. The other three hybrids (WR44, WR52, WR153) contained an *O. rufipogon* type chloroplast genome and an *O. meridionalis* type nuclear genome.

Analysis of Gene Loci

SNP were analysed for six genes on chromosomes 4, 6, and 7 were analysed to assess the genetic contribution of the putative parents to the natural hybrid populations in relation to *O. sativa* spp. *japonica* cv. Nipponbare reference genome. Of the six genes, one gene (*sh4*) on chromosome 4 is associated with seed shattering while the other five genes are starch synthesis-related genes located on chromosome 6 (*GBSSI*, *SSI*, *SBEI*, and *ALK*) and chromosome *7* (*GBSSII*). From SNP data, it was clearly shown that the six genes of WR62 received alleles from both *O. rufipogon* type taxa and *O. meridionalis* parents. *SBEI* and *GBSSII* had heterozygous SNP genotypes with both *O. rufipogon* type taxa and *O. meridionalis* alleles in all of five hybrids (**Table 1** and **Figure 2**). However, the *sh4* alleles were heterozygous for WR44, WR62, and WR161 and homozygous for WR52 and WR153. Two hybrids (WR153 and WR161) also had heterozygous loci for *ALK* and *SSI*, respectively. Sequence alignment images of SNP genotypes of sh4 gene (ID: Os04g0670900-1, position on chromosome 4 from 34,231,186 to 34,233,221 bp), *ALK* gene (ID: Os06g0229800-1, position on chromosome 6 from 6,748,398 to 6,753,302 bp), and *GBSSII* (ID: Os07g0412100-1, position on chromosome 7 from 12,916,883 to 12,924,202 bp) (RAPDB, 2021) showed heterozygous locus as in **Figure 2**. WR44 and WR62 showed heterozygous loci at the 34,231,465 and 34,231,468 base positions (**Figure 2A**) while WR62 and WR44 showed heterozygous loci at 6,751,298 and 12,919,934 base position, respectively (**Figures 2B,C**). In all cases, a heterozygote was formed by receiving an allele from *O. rufipogon* type taxa and the alternative allele from *O. meridionalis*. Hybrids with homozygous loci represented alleles mostly from *O. meridionalis*. The number of total variants across the whole genome varied among hybrids, ranging between 5,484,261 and 8,024,717. W44 and W62 are early generation hybrids as shown by the presence of alleles from both parents while others are backcrossed to recover a nuclear genome similar to the recurrent parent. The highest number of SNPs (6,910,977) was found in WR44, followed by 5,917,274 in WR62, 5,291,708 in WR52, 5,295,508 in WR161 and 4,764,266 in

TABLE 1 | Characterisation of natural hybrids in wild rice populations.

Hybrid	SNP genotypes of six genes on chromosome 4, 6, and 7						Total SNPs	Total heterozygous SNPs	Heterozygosity across the whole genome (%)	Plausible hybrid generations
	sh4 alleles	*GBSSI* alleles	*SSI* alleles	*SBEI* alleles	*ALK alleles*	*GBSSII alleles*				
WR44	AB	BB	BB	AB	BB	AB	6,910,977	2,464,823	0.69	Early (F1/F2)
WR52	BB	BB	BB	AB	BB	AB	5,291,708	1,137,527	0.32	Later
WR62	AB	AB	AB	AB	AB	AB	5,917,274	2,318,878	0.60	Early (F1/F2)
WR153	BB	BB	BB	AB	AB	AB	4,764,266	1,111,890	0.31	Later
WR161	AB	BB	AB	AB	BB	AB	5,295,508	1,189,627	0.34	Later

Plausible hybrid generations are shown based on the level of heterozygosity across the whole genome. AB and BB denote heterozygous and homozygous loci, respectively.

WR153 (**Table 1**). Similarly, the highest number of heterozygous SNPs were found in WR44 (2,464,823) and the lowest number in WR153 (1,111,890). To determine the heterozygosity across the whole genome, the genome size of *O. rufipogon* type taxa was applied to the above-mentioned formula (2) for WR62, which exhibited *O. rufipogon* type nuclear genome while the genome size of *O. meridionalis* was used for the other hybrids. A high level of genetic variation was observed in both WR44 (0.69%) and WR62 (0.6%) across the whole genome followed by WR161 (0.34%), WR52 (0.32%), and WR153 (0.31%) in relation to *O. sativa* spp. *japonica* cv. Nipponbare (**Table 1**). The level of heterozygosity across the whole genome of these hybrids confirmed hybridisation events and led to the conclusion that WR44 and WR62 were early generation hybrids while the other hybrids were much later generations due to recurrent backcrossing with one parent (*O. meridionalis*).

Analysis of Genetic Introgression

A four-taxon test (as *D*-statistic) was performed to test for biallelic mutation patterns (ABBA-BABA) to determine geneflow between the hybrids and two putative perennial parental populations in a topology {[(*O. rufipogon* type taxa, *O. meridionalis*), hybrid], *O. sativa* spp. *japonica* cv. Nipponbare}. The length of a MAFFT aligned sequence was 4,473 bp with a total of four blocks: (1–1,118); (1,119–2,236); (2,237–3,354); and (3,355–4,473). The *D*-statistic value based on the distribution of AABB and BABA on each block showed significant deviation for all four-taxon trees (**Table 2**). The negative *D*-statistic value suggested potential genetic introgression between *O. rufipogon* type taxa (WR24) and hybrid while the positive value suggests geneflow between *O. meridionalis* (WR81) and hybrid. In all cases, *O. sativa* spp. *japonica* cv. Nipponbare was an outgroup. The *Z*-score for all four-taxon tests was significant ($p < 0.001$).

Intra- and Interspecific Variations in the Wild Rice Gene Pool

To estimate intra- and interspecific variations in the wild rice gene pool, variants comprised of five different types: SNV, insertions, deletions, MNV, and replacement were assessed across the whole genome of twenty-one putative parental taxa (including one *O. rufipogon* type taxa genotypes and 20 *O. meridionalis* genotypes). The number of allelic positions where the variants were located was determined. The total number of variants ranged from 1,918,262 in WR24 to 7,310,928 in WR81 (**Table 3**). Overall, 87–90% of the total variants in an individual were SNPs for all samples. While the proportion of InDels (insertions and deletions) was low ranging between 6 and 8%. The number of heterozygous SNP positions varied among wild rice populations. The highest number of heterozygous SNP was found in WR81 (1,365,802) while the lowest number of heterozygous SNP was found in WR24 (632,272). Consequently, the level of heterozygosity across the whole genome was higher in *O. meridionalis* (0.38%), and the least heterozygosity was found in *O. rufipogon* type taxa (0.16%). The SNPs that were uniquely present in individuals were determined to define the genetic relationships between individuals. The rare or unique heterozygous SNP position (%) ranged from 4 to 48%. The highest number of unique heterozygous SNP positions was in *O. rufipogon* type taxa and the lowest number in *O. meridionalis*, indicating *O. rufipogon* type taxa might be subject to more genetic drift compared to *O. meridionalis* populations due to their smaller population size.

Genetic Differentiation of Wild Rice Populations

Genome-wide SNP data (12,247,414 SNPs) of 26 wild rice genotypes were first processed to remove linkage for PCA analysis. A total of 490,655 SNP sites were retained that fell below the linkage threshold. PCA was performed on linkage-pruned sites to determine the genetic groupings of wild rice populations. PCA analysis showed two clear clusters: (1) one group formed with hybrids (WR44, WR62, and WR161) and one putative parental taxon (WR24) and (2) all *O. meridionalis* populations along with two plausible later generations of hybrids (WR52 and WR153) formed another group (**Figure 3**). The first axis (PC1) explained 64.6% of the variation, which separated out the former group. There was a very low level of genetic differentiation observed in the latter group represented by the second axis (PC2) with 22.4% variation.

Isolation by distance analysis based on a Monte-Carlo test with 999 permutations clearly showed that there was no significant relationship ($R^2 = -0.008$, p-value = 0.42) between Nei's pairwise genetic distance and geographic distances (**Figure 4**). A total of 300 pairwise Nei's genetic distance data were plotted against pairwise geographic distance. The observed variation (R^2) was 0.008 with a p-value of 0.42. This indicated genetic

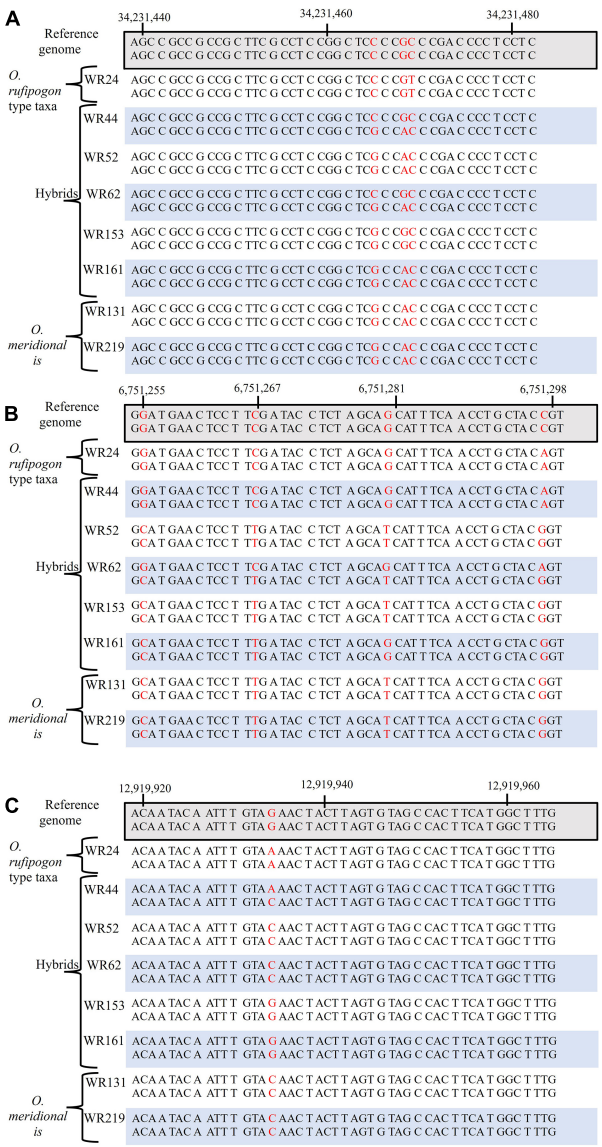

FIGURE 2 | Alignment of single nucleotide polymorphism (SNP) genotypes of *O. rufipogon* type taxa (WR24), hybrids (WR44, WR52, WR62, WR153, and WR161) and *O. meridionalis* (WR133 and WR219), in relation to, *O. sativa* spp. *japonica* cv. Nipponbare as a reference genome: **(A)** *seed shattering 4* (*sh4*) gene on chromosome 4, **(B)** *alkali degeneration* (*ALK*) gene on chromosome 6, and **(C)** *granular bound starch synthase II* (*GBSSII*) on chromosome 7. The red color denotes SNPs.

differentiation was not significantly increased with geographic distance, indicating genetic connectivity in wild rice populations across a wide range.

DISCUSSION

Reticulate Origin of AA Genome Hybrids

This study confirms two types of the chloroplast genome in hybrid populations introgressed from one of their parental taxa. A total of five hybrids were studied of which only one hybrid (WR62) exhibited an *O. meridionalis* type chloroplast genome while the other four hybrids displayed *O. rufipogon* type chloroplast genomes. This indicates that hybridisation in both directions is possible. The previous study of these populations using the whole chloroplast and nuclear genome revealed chloroplast capture results in discordant cytonuclear phylogenies, inferring hybrid origins (Moner et al., 2018). Chloroplast capture can be the consequence of lineage sorting, convergent evolution, and recurrent hybridisation (Tsitrone et al., 2003). Unidirectional geneflow from recurrent hybridisation paves the way to capture either cytoplasm (Takahata and Slatkin, 1984) or nuclear genes (Petit et al., 2001) from one species to another species. However, introgression promotes chloroplast capture rather than nuclear genome capture in sympatric populations as has been detected in many plant species, resulting in discordant cytonuclear phylogenies (Rieseberg and Soltis, 1991). Unlike introgression, chloroplast capture can occur in species through one single hybridisation event and species with the freshly obtained chloroplast genome can be inherited by descendant lineages over time (Nge et al., 2021). In this study, hybridisation or introgressive hybridisation may have allowed chloroplast genome capture among sympatric wild rice populations. It is also known that nuclear genome incompatibilities may stimulate chloroplast capture in a random mating population (Tsitrone et al., 2003). Brozynska et al. (2017) reported that nuclear genomes of two parental taxa were distinguished where *O. rufipogon* type taxa exhibited more sequence similarity with *O. sativa* relative to the other Australian AA genome wild rice species. Whole nuclear genome analysis revealed that the Australian *O. rufipogon* had a nuclear genome that was part of a clade with Asian cultivated rice and its wild progenitors in Asia while *O. meridionalis* was sister to this clade in a nuclear genome phylogenetic tree (Moner et al., 2018). Despite the distinct nuclear genome, the chloroplast genome of the Australian *O. rufipogon* type taxa are more closely related to *O. meridionalis* differing by 53 variants and 36 variants from different accessions of *O. meridionalis*, respectively, while these forms of *O. meridionalis* from different regions (Queensland and Northern Territory) differed from each other by 34 variants (Brozynska et al., 2014). The *O. rufipogon* in Australia was much more divergent from *O rufipogon* in Asia (125 variants) and *O. sativa* spp. Nipponbare cv. *japonica* (125 variants). In the Australian *O. rufipogon*, ancient chloroplast capture has occurred from *O. meridionalis* (Brozynska et al., 2014), however, hybrids showed the two types of chloroplasts (*O. meridionalis* and Australian *O. rufipogon*) due to recent bidirectional chloroplast capture from these parents. The recent introgressive hybridisation demonstrated here may be an explanation of both the apparent chloroplast capture in the Australia *O. rufipogon* populations and the wild hybrids.

Analysis of nuclear genes demonstrated that the hybrids had alleles from both putative parental populations forming a heterozygote in early generations, indicating biparental nuclear geneflow from interspecific hybridisation. However, these loci become homozygous in what appear to be many later generations of hybrids that are fixed for the alleles of the recurrent parent, usually *O. meridionalis*. These observations clearly demonstrate

TABLE 2 | Estimation of geneflow in a triplet species following four-taxon tests with HybridCheck version 1.0 software.

Species in triplet			Outgroup species	Block start	Block end	numABBA	numBABA	*D*-statistic	*Z*-score	Direction of geneflow
P1	P2	P3	A							
WR24	WR81	WR44	Nipponbare	1	1,118	168	10	0.47	0.35*	P3→P1
WR24	WR81	WR52	Nipponbare	1	1,118	133	6	0.38	−0.38*	P3→P1
WR24	WR81	WR62	Nipponbare	1	1,118	66	18	0.06	−1.81*	P3→P1
WR24	WR81	WR153	Nipponbare	1	1,118	73	18	0.03	−2.14*	P3→P1
WR24	WR81	WR161	Nipponbare	1	1,118	55	0	0.33	−0.78*	P3→P1
WR24	WR81	WR44	Nipponbare	1,119	2,236	0	8	−0.28	0.35*	P2→P3
WR24	WR81	WR52	Nipponbare	1,119	2,236	0	13	−0.49	−0.38*	P2→P3
WR24	WR81	WR62	Nipponbare	1,119	2,236	0	27	−0.69	−1.81*	P2→P3
WR24	WR81	WR153	Nipponbare	1,119	2,236	0	15	−0.49	−2.14*	P2→P3
WR24	WR81	WR161	Nipponbare	1,119	2,236	0	14	−0.43	−0.78*	P2→P3
WR24	WR81	WR44	Nipponbare	2,237	3,354	9	19	−0.57	0.35*	P2→P3
WR24	WR81	WR52	Nipponbare	2,237	3,354	9	16	−0.65	−0.38*	P2→P3
WR24	WR81	WR62	Nipponbare	2,237	3,354	9	19	−0.66	−1.81*	P2→P3
WR24	WR81	WR153	Nipponbare	2,237	3,354	9	19	−0.68	−2.14*	P2→P3
WR24	WR81	WR44	Nipponbare	3,355	4,472	19	21	−0.06	0.35*	P2→P3
WR24	WR81	WR52	Nipponbare	3,355	4,472	19	21	−0.12	−0.38*	P2→P3
WR24	WR81	WR62	Nipponbare	3,355	4,472	19	21	−0.16	−1.81*	P2→P3
WR24	WR81	WR153	Nipponbare	3,355	4,472	19	21	−0.31	−2.14*	P2→P3
WR24	WR81	WR161	Nipponbare	3,355	4,472	19	21	−0.14	−0.78*	P2→P3

*Triplets are characterised: P1 = O. rufipogon type taxa; P2 = O. meridionalis; and P3 = Hybrid. O. sativa spp. japonica cv. Nipponbare (A) is an outgroup species. The length of aligned nucleotide is 4,473 bp containing four blocks: (1–1,118); (1,119–2,236); (2,237–3,354); and (3,355–4,473). The biallelic patterns based on the number of ABBA-BABA in a triplet used in the Patterson's D-statistic are shown. P3→P1 denotes geneflow between O. rufipogon type taxa and hybrid while P2→P3 denotes geneflow between O. meridionalis and hybrid. *Z-score represents statistically significant at $p < 0.001$.*

TABLE 3 | Genetic variation across the whole genome of 21 wild rice samples, excluding hybrids.

Species	Samples	Total number of variants	SNP (%)	InDel (%)	Total number of heterozygous SNV position	Heterozygous SNP variation across the whole genome (%)	*Unique heterozygous SNP (%)
O. rufipogon type taxa	WR24	1,918,262	90	6	632,272	0.16	48
O. meridionalis	WR37	3,745,572	87	8	827,689	0.23	6
	WR81	7,310,928	87	8	1,365,802	0.38	8
	WR100	6,007,510	87	8	1,175,173	0.33	4
	WR103	6,256,551	87	8	1,198,555	0.34	5
	WR111	3,899,236	88	7	894,765	0.25	5
	WR126	5,194,396	87	8	1,125,277	0.32	10
	WR133	6,545,660	87	8	1,227,038	0.35	4
	WR143	5,611,030	87	8	1,109,062	0.31	5
	WR171	3,852,279	88	7	966,368	0.27	6
	WR184	5,421,763	87	8	1,074,882	0.30	5
	WR195	5,985,028	87	8	1,168,776	0.33	6
	WR207	6,520,874	87	8	1,210,794	0.34	12
	WR219	6,951,983	87	8	1,312,291	0.37	6
	WR230	6,694,713	87	8	1,261,074	0.36	5
	WR233	3,623,110	88	7	876,065	0.25	4
	WR242	6,849,045	87	8	1,299,754	0.37	6
	WR256	5,991,356	87	8	1,119,055	0.32	6
	WR265	6,625,636	87	8	1,255,312	0.35	5
	WR280	6,621,409	87	8	1,249,681	0.35	5
	WR287	5,722,665	87	8	1,105,062	0.31	5

**Unique heterozygous SNP: SPNs that are uniquely present in one sample but absent in other samples.*

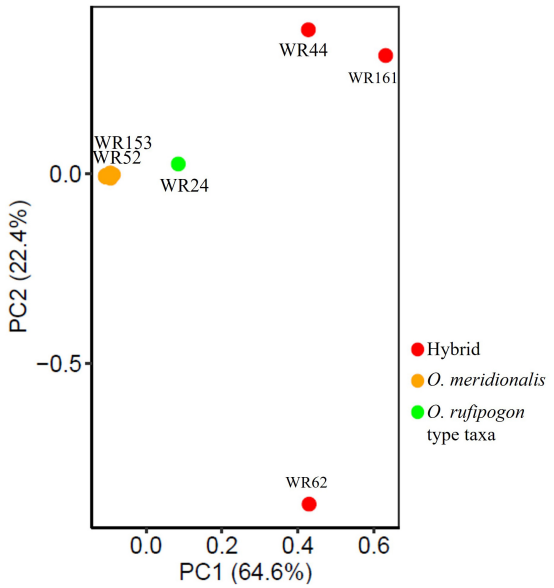

FIGURE 3 | Principal component analysis (PCA) of 26 wild rice populations based on neutral 490,655 SNPs across the whole genome.

an ongoing hybridisation and introgression persisting in the wild rice populations evidenced by the presence of viable hybrids between *O. rufipogon* type taxa and *O. meridionalis* being found at very low frequency. Earlier findings postulated that interspecific crossability between an *O. meridionalis* and a perennial Asian *O rufipogon* was modest but asymmetric (Banaticla-Hilario et al., 2013). As a result, they reported a very low percentage of fertile first filial generation (F1) (i.e., with mean F1 fertility = 17.2%) with perennial habit. However, they posited that a perennial habit enabled the F1 progeny to persist for a longer period than its annual counterpart. Like annual strains, crosses between accessions from the Jpn2 location (known to be *O. meridionalis*- like) and *O. rufipogon* produce low fertility pollen (<10%) as found with putative annual strains (<20–40%) (Sotowa et al., 2013). Intra-specific crossability between *O. meridionalis*, within the same geographical origin, produces fertile hybrids with pollen stainability (96.7%) and panicle fertility (66.9%), however, geographical distance seems to be the reason for the development of reproductive barriers (Juliano et al., 2005). They observed fertile hybrids, which were derived from crosses between Queensland accessions, however, crosses between accessions from different geographical regions produced sterile to partially fertile hybrids. Similarly, very low pollen stainability and seed fertility were reported in intraspecific crosses between *O. meridionalis* populations, from different geographical locations (Naredo et al., 1997).

The positive and negative significant *D*-statistic values from the four-taxon test indicate a nuclear geneflow from both putative parents into hybrids. Early generation hybrids exhibited higher heterozygosity and polymorphic variations than that of the later generations of hybrids. Repeated backcrossing between F1/F2 offspring and *O. meridionalis* may lead to comparatively low heterozygous SNP variation in the later-generation of hybrids. PCA analysis also reveal that early generation hybrids were clustered with *O. rufipogon* type taxa and later generation hybrids (except for WR161) form a group with *O. meridionalis*. Reproductive barriers underpin a termination of hybridisation and thereby, introduce introgression (Rieseberg and Wendel, 1993). An interspecific reproductive barrier gene (*SEED DEVELOPMENT 1*) on chromosome 6 is detected in *O. meridionalis* and *O. sativa* (Toyomoto et al., 2019). A recessive allele of a *SEED DEVELOPMENT 1* gene (*sdv1-m*) in wild rice strain from the Jpn2 site (=*O. meridionalis*) appears to evoke segregation distortion (1:2:0) in a BC_4F_2 progeny. Therefore, offspring carrying homozygous recessive alleles (*sdv1-m)* cause seed abortion at the seed development stage. Toyomoto et al. (2019) made an F1 offspring from a cross between *O. sativa*

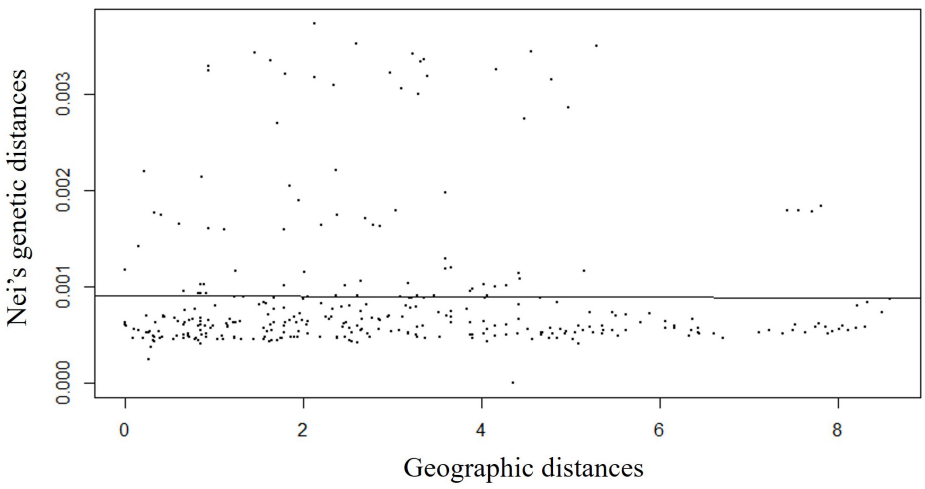

FIGURE 4 | Isolation-by-distance (IBD) plots of Nei's pairwise genetic distance values against the lineal geographic distance of Australian wild rice populations based on 2,987,364 SNPs and 26 sample sites.

and *O. meridionalis*, then backcrossed with *O. sativa* to obtain backcrossing progeny. Despite homozygous recessive alleles, offspring bear seed if they are self-fertilised. *O. rufipogon* from Asia or Australia does not have this abortive gene. Backcrossing with *O. meridionalis* along with selfing may reduce seed abortion in later-generation hybrids.

Phenotypic characteristics of hybrids observed by Moner et al. (2018) seem consistent with the nuclear gene analysis. *O. rufipogon* type taxa are distinguished with open panicles and short awns from *O meridionalis,* which exhibits long awns with closed panicles. However, awn lengths sometimes overlapped (Brozynska et al., 2017; Moner et al., 2018). Early generation hybrids (WR44 and WR62) showed open to partially open panicles orientation like *O. rufipogon* type taxa. While all later generations hybrids exhibit panicles orientation like *O. meridionalis* type closed panicles. Interestingly, both generations of hybrids exhibited an intermediate anther length between two putative parental taxa and awn size close to *O. meridionalis*. Therefore, hybrids exhibit both intermediate and putative parental phenotypic characteristics, depending on which dominant loci of a trait are received from putative parents. Intermediate phenotypic characteristics have been observed in F1 hybrids derived from a cross between *O. rufipogon* and *O. meridionalis* (Naredo et al., 1997). Furthermore, F1 hybrids follow the lifecycle of their parental taxa, i.e., perennial F1 hybrids form when the two parental taxa are perennial in habit and annual and perennial parents produce semi-perennial to perennial offspring (Naredo et al., 1997). In contrast, Banaticla-Hilario et al. (2013) reported F1 offspring derived from a cross between Australasian perennial *O. rufipogon* and an annual *O. meridionalis* appears to be phenotypically closer to annual *O. meridionalis*. While F1 offspring with non-Australian *O. rufipogon* produce phenotypic characteristics more likely to be perennial *O. rufipogon* species. They also reported that interspecific F1 hybrids demonstrate either parental or intermediate phenotypic characteristics, i.e., some characteristics are analogous to their annual parental species and few features are intermediate between the two parents.

Reticulate evolution is a common phenomenon in vascular plants, wherein genetic changes occur through hybrid speciation, introgression, introgressive hybridisation, and horizontal gene transfer (Vriesendorp and Bakker, 2005). Hybrid origins were first proposed in the AA genome wild rice gene pool as an explanation of discordant cytonuclear phylogenies due to chloroplast capture (Moner et al., 2018). However, the disparity in cytonuclear phylogenies alone can never be an indicator of hybrid status (Vriesendorp and Bakker, 2005). Other events include incomplete lineage sorting, i.e., maintenance of ancestral polymorphisms through multiple speciation events (Andreasen and Baldwin, 2003; Goldman et al., 2004), homoplasy, and taxonomic sampling error (Wendel and Doyle, 1998) can cause such discordant patterns. The nuclear gene analysis reported here confirms hybrid events in the wild AA genome rice genepool and confirms that the discordant phylogenies are the result of hybridisation events. The nuclear gene analysis along with chloroplast genome data clearly reveal a reticulate evolution pattern in the wild AA genome rice gene pool.

Oryza rufipogon from other countries shows no evidence of hybridisation with *O. meridionalis* as this species is not found in other countries. No highly divergent AA genome species that *O. rufipogon* could hybridise with is present outside Australia. This analysis is only possible in Australia where large populations of the two divergent species (*O. meridionalis* and the distinct Australian *O. rufipogon* like taxa) were searched to discover extremely rare apparent hybrids based upon their intermediate morphology. This study has confirmed the hybrid identity of these extremely rare plants. The Asian *O. rufipogon* samples that have been sequenced all show the presence of an *O. rufipogon* chloroplast that is distinct from that of *O. meridionalis* and the Australian *O. rufipogon* like taxa (Brozynska et al., 2014; Moner et al., 2018). The two AA genome species in Australia (*O. meridionalis* and the Australian *O. rufipogon* like taxa) remain as distinct populations despite the extremely rare hybridisation events between these two taxa confirmed in this study. The reproductive barrier between them resulting in low pollen fertility requires the backcrossing of the hybrids in subsequent generations.

Australian Wild Rice Populations as a Reservoir for Novel Genetic Resources

This study reveals high levels of heterozygous SNP variation in AA genome wild rice populations, in relation to *O. sativa* ssp. *japonica* cv. Nipponbare. This finding is, in agreement with the result reported by Krishnan et al. (2014). They found that the Australian wild rice genepool contains a high level of genetic variations, in terms of, SNPs across the whole genome, i.e., 2,418,084 SNPs in *O. meridionalis* and 2,564,013 SNPs in *O. rufipogon*, compared to Asian *O. rufipogon* (917,738 SNPs) and cultivated *O. sativa* ssp. *indica* (978,630 SNPs), with respect to *O. sativa* ssp. *japonica* cv. Nipponbare. A molecular study using restriction fragment length polymorphism (RFLP) markers revealed that cultivated rice retained 40% of the wild alleles during rice domestication (Sun et al., 2001). This study reveals the level of heterozygosity is lower in *O. rufipogon* than in *O. meridionalis*. This is also consistent with the previous results reported by Brozynska et al. (2017). They postulated *O. meridionalis* (Taxon B) had a more heterozygous and repetitive genome than *O. rufipogon* (Taxon A), resulting in a higher number of scaffolds in the egnoem of this taxon. Mating systems play an important role in acquiring genetic variations. Perennials exhibit outcrossing mating system, which results in a higher degree of genetic variation than annuals while selfing is a feature of annuals (Oka, 1988). In Australia, *O. meridionalis* is common and widespread while the Australian *O. rufipogon* populations are much smaller. Comparison of the relative diversity of the two species in Australia will require more extensive sampling of the much rarer *O. rufipogon* populations. PCA and Isolation by-distance analysis also reveals a high degree of genetic connectivity among *O. meridionalis* populations across a wide geographic range in Northern Queensland, Australia. The role of animals (especially birds) in dispersing *Oryza* species over large distances is quite evident (Tang et al., 2010; Wambugu et al., 2015).

Natural variations in genes for agronomically important traits are limited in cultivated rice germplasm. In-plant breeding, desirable germplasm is sourced from wild species, landraces, and distant relatives, or for new traits through induced mutation or genetic manipulation (Ganeshan et al., 2010). The Australian wild rice populations include the earliest branching AA genome lineages and, therefore the most genetically and geographically distinct from the other AA genome wild relatives of cultivated rice. Asian domesticated rice was introduced to Australia around 200 years ago and mainly grown far from wild rice populations (Henry et al., 2010) avoiding genetic introgression of domestication genes into the wild populations. Recent studies already reveal some interesting features of the native Australian populations, in terms of disease resistance, grain appearance, and nutritional properties (Kasem et al., 2012, 2014; Tikapunya et al., 2017) suggesting that they could be exploited to extract numerous genes for agronomically important traits for rice improvement.

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: https://www.ncbi.nlm.nih.gov/genbank/, BioProject number PRJNA758754.

AUTHOR CONTRIBUTIONS

RH conceived the project. AF and RH collected field samples. SH, AF, and RH analysed the data. SH wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Cooperative Research Centre for Developing Northern Australia (CRCNA).

ACKNOWLEDGMENTS

We would like to acknowledge the contribution of Ali M. Moner to collecting samples from the field. We also thank the Research Computer Centre at the University of Queensland for their support and advice and the CRCNA for funding.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.767635/full#supplementary-material

Supplementary Figure 1 | The phylogenetic tree of the chloroplast genome of four samples and two reference chloroplast genomes (*O. meridionalis* and *O. rufipogon* type taxa), with *O. officinalis* as the outgroup. Among four samples, WR62 and WR161 (red in colour) were taken from this study while the other two samples WR65 and WR162 (green in colour) were taken from Moner et al. (2018). The tree was generated using the maximum likelihood (ML) method in PAUP* version 4 software with 1,000 bootstrap replicates. The ML bootstrap value (/40) is marked on each node.

REFERENCES

Andreasen, K., and Baldwin, B. G. (2003). Nuclear ribosomal DNA sequence polymorphism and hybridization in checker mallows (*Sidalcea*, Malvaceae). *Mol. Phylogenet. Evol.* 29, 563–581. doi: 10.1016/S1055-7903(03)00136-2

Banaticla-Hilario, M. C. N., McNally, K. L., van den Berg, R. G., and Hamilton, S. N. R. (2013). Crossability patterns within and among *Oryza* series sativae species from Asia and Australia. *Genet. Resour. Crop Evol.* 60, 1899–1914. doi: 10.1007/s10722-013-9965-4

Brar, D. S., and Khush, G. S. (2018). "Wild relatives of rice: a valuable genetic resource for genomics and breeding research," in *The Wild Oryza Genomes*, eds T. K. Mondal and R. J. Henry (Cham: Springer International Publishing), 1–25.

Brozynska, M., Copetti, D., Furtado, A., Wing, R. A., Crayn, D., Fox, G., et al. (2017). Sequencing of Australian wild rice genomes reveals ancestral relationships with domesticated rice. *Plant Biotechnol. J.* 15, 765–774. doi: 10.1111/pbi.12674

Brozynska, M., Omar, E. S., Furtado, A., Crayn, D., Simon, B., Ishikawa, R., et al. (2014). Chloroplast genome of novel rice germplasm identified in Northern Australia. *Trop. Plant Biol.* 7, 111–120. doi: 10.1007/s12042-014-9142-8

Chat, J., Jáuregui, B., Petit, R. J., and Nadot, S. (2004). Reticulate evolution in kiwifruit (*Actinidia*, Actinidiaceae) identified by comparing their maternal and paternal phylogenies. *Am. J. Bot.* 91, 736–747. doi: 10.3732/ajb.91.5.736

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772–772. doi: 10.1038/nmeth.2109

Doyle, J. J., Doyle, J. L., Rauscher, J. T., and Brown, A. H. D. (2003). Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (Glycine subgenus Glycine). *New Phytol.* 161, 121–132. doi: 10.1046/j.1469-8137.2003.00949.x

Furtado, A. (2014). "DNA extraction from vegetative tissue for next-generation sequencing," in *Cereal Genomics: Methods and Protocols*, Vol. 1099, eds R. J. Henry and A. Furtado (Totowa, NJ: Humana Press Inc), 1–5. doi: 10.1007/978-1-62703-715-0_1

Ganeshan, S., Sharma, P., and Chibbar, R. N. (2010). "Functional genomics for crop improvement," in *Molecular Techniques in Crop Improvement*, 2nd Edn, eds S. M. Jain and D. S. Brar (Dordrecht: Springer Netherlands), 63–95. doi: 10.1007/978-90-481-2967-6_3

Goldman, D. H., Jansen, R. K., van den Berg, C., Leitch, I. J., Fay, M. F., and Chase, M. W. (2004). Molecular and cytological examination of *Calopogon* (Orchidaceae, Epidendroideae): circumscription, phylogeny, polyploidy, and possible hybrid speciation. *Am. J. Bot.* 91, 707–723. doi: 10.3732/ajb.91.5.707

Griffith, M. P. (2003). Using molecular evidence to elucidate reticulate evolution in *Opuntia*. *Madroño* 50, 162–169.

Henry, R. J. (2019). Australian wild rice populations: a key resource for global food security. *Front. Plant Sci.* 10:1354. doi: 10.3389/fpls.2019.01354

Henry, R. J., Rice, N., Waters, D. L. E., Kasem, S., Ishikawa, R., Hao, Y., et al. (2010). Australian *Oryza*: utility and conservation. *Rice* 3, 235–241. doi: 10.1007/s12284-009-9034-y

Jombart, T., and Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27, 3070–3071. doi: 10.1093/bioinformatics/btr521

Juliano, A. B., Naredo, M. E. B., Lu, B.-R., and Jackson, M. T. (2005). Genetic differentiation in *Oryza meridionalis* Ng based on molecular and crossability

analyses. *Genet. Resour. Crop Evol.* 52, 435–445. doi: 10.1007/s10722-005-2259-8

Karimi, N., Grover, C. E., Gallagher, J. P., Wendel, J. F., Ané, C., and Baum, D. A. (2019). Reticulate evolution helps explain apparent homoplasy in floral biology and pollination in baobabs (*Adansonia*; Bombacoideae; Malvaceae). *Syst. Biol.* 69, 462–478. doi: 10.1093/sysbio/syz073

Kasem, S., Waters, D. L. E., and Henry, R. J. (2012). Analysis of starch gene diversity in the wild relatives of Oryza sativa. *Trop. Plant Biol.* 5, 286–308. doi: 10.1007/s12042-012-9111-z

Kasem, S., Waters, D. L. E., Ward, R. M., Rice, N. F., and Henry, R. J. (2014). Wild *Oryza* grain physico-chemical properties. *Trop. Plant Biol.* 7, 13–18. doi: 10.1007/s12042-013-9133-1

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kawahara, Y., de la Bastide, M., Hamilton, J. P., Kanamori, H., McCombie, W. R., Ouyang, S., et al. (2013). Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6:4. doi: 10.1186/1939-8433-6-4

Krishnan, S. G., Waters, D. L. E., and Henry, R. J. (2014). Australian wild rice reveals pre-domestication origin of polymorphism deserts in rice genome. *PLoS One* 9:e98843. doi: 10.1371/journal.pone.0098843

Lemmon, A. R., Emme, S. A., and Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Syst. Biol.* 61, 727–744. doi: 10.1093/sysbio/sys049

Li, J. Q., Lu, Y. G., Feng, J. H., Zhao, X. J., and Liu, X. D. (2007). The crossability and F1 hybrid fertility between *Oryza sativa* and other AA genome species. *J. Plant Genet. Resour.* 8, 1–6.

Li, J., Zhou, J., Xu, P., Deng, X., Deng, W., Zhang, Y., et al. (2018). Mapping five novel interspecific hybrid sterility loci between *Oryza sativa* and *Oryza meridionalis*. *Breed. Sci.* 68, 516–523. doi: 10.1270/jsbbs.18001

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27(2 Pt 1), 209–220.

Map Maker (2021). *Map Maker Make & Share Custom Maps*. Available online at: https://maps.co/ (accessed August 12, 2021).

Martin, S. H., Davey, J. W., and Jiggins, C. D. (2015). Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol. Biol. Evol.* 32, 244–257. doi: 10.1093/molbev/msu269

Mason-Gamer, R. J. (2004). Reticulate evolution, introgression, and intertribal gene capture in an allohexaploid grass. *Syst. Biol.* 53, 25–37. doi: 10.1080/10635150490424402

Moner, A. M., Furtado, A., Chivers, I., Fox, G., Crayn, D., and Henry, R. J. (2018). Diversity and evolution of rice progenitors in Australia. *Ecol. Evol.* 8, 4360–4366. doi: 10.1002/ece3.3989

Naredo, M. E. B., Juliano, A. B., Lu, B.-R., and Jackson, M. T. (1997). Hybridization of AA genome rice species from Asia and Australia I. Crosses and development of hybrids. *Genet. Resour. Crop Evol.* 44, 17–23. doi: 10.1023/A:1008670923628

Nei, M. (1972). Genetic distance between populations. *Am. Nat.* 106, 283–292.

Ng, N. Q., Hawkes, J. G., Williams, J. T., and Chang, T. T. (1981). The recognition of a new species of rice (*Oryza*) from Australia. *Bot. J. Linn. Soc.* 82, 327–330. doi: 10.1111/j.1095-8339.1981.tb00964.x

Nge, F. J., Biffin, E., Thiele, K. R., and Waycott, M. (2021). Reticulate evolution, ancient chloroplast haplotypes, and rapid radiation of the Australian plant genus *Adenanthos* (proteaceae). *Front. Ecol. Evol.* 8:616741. doi: 10.3389/fevo.2020.616741

Oka, H. I. (1988). *Origin of Cultivated Rice*. Tokyo: Japan Scientific Societies Press.

Palmer, J. D., Jorgensen, R. A., and Thompson, W. F. (1985). Chloroplast DNA variation and evolution in pisum: patterns of change and phylogenetic analysis. *Genetics* 109, 195–213. doi: 10.1093/genetics/109.1.195

Petit, R. J., Bialozyt, R., Brewer, S., Cheddadi, R., and Combs, B. (2001). "From spatial patterns of genetic diversity to post glacial migration processes in forest trees," in *Integrating Ecology and Evolution in a Spatial Context*, eds J. Silvertown and J. Antonovics (Oxford: Blackwell Science), 295–318.

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, 559–575. doi: 10.1086/519795

RAPDB (2021). *Rice Annotation Project Database*. Available online at: https://rapdb.dna.affrc.go.jp/viewer/gbrowse_details/irgsp1?name=Os06g0229800;feature_id=50299 (accessed June 19, 2021).

Rieseberg, L. H. (1991). Homoploid reticulate evolution in *Helianthus* (Asteraceae): evidence from ribosomal genes. *Am. J. Bot.* 78, 1218–1237. doi: 10.2307/2444926

Rieseberg, L. H., and Soltis, D. E. (1991). Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol. Trends Plants* 5, 65–84.

Rieseberg, L. H., and Wendel, J. F. (1993). "Introgression and its consequences in plants," in *Hybrid Zones and the Evolutionary Process*, ed. R. G. Harrison (Oxford: Oxford University Press), 70.

Sang, T., Crawford, D. J., and Stuessy, T. F. (1995). Documentation of reticulate evolution in Peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA — implications for biogeography and concerted evolution. *Proc. Natl. Acad. Sci. U.S.A.* 92, 6813–6817. doi: 10.1073/pnas.92.15.6813

Sotowa, M., Ootsuka, K., Kobayashi, Y., Hao, Y., Tanaka, K., Ichitani, K., et al. (2013). Molecular relationships between Australian annual wild rice, *Oryza meridionalis*, and two related perennial forms. *Rice* 6:26. doi: 10.1186/1939-8433-6-26

Stebbins, G. L. (1959). The role of hybridization in evolution. *Proc. Am. Philos. Soc.* 103, 231–249.

Stein, J. C., Yu, Y., Copetti, D., Zwickl, D. J., Zhang, L., Zhang, C., et al. (2018). Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. *Nat. Genet.* 50, 285–296. doi: 10.1038/s41588-018-0040-0

Sun, C. Q., Wang, X. K., Li, Z. C., Yoshimura, A., and Iwata, N. (2001). Comparison of the genetic diversity of common wild rice (*Oryza rufipogon* Griff.) and cultivated rice (*O. sativa* L.) using RFLP markers. *Theor. Appl. Genet.* 102, 157–162. doi: 10.1007/s001220051631

Swofford, D. L. (2003). *PAUP4. Phylogenetic Analysis Using Parsimony (4and Other Methods). Version 4. 2003*. Sunderland, MA: Sinauer Associates.

Takahata, N., and Slatkin, M. (1984). Mitochondrial gene flow. *Proc. Natl. Acad. Sci. U.S.A.* 81, 1764–1767. doi: 10.1073/pnas.81.6.1764

Tang, L., Zou, X. H., Achoundong, G., Potgieter, C., Second, G., Zhang, D. Y., et al. (2010). Phylogeny and biogeography of the rice tribe (Oryzeae): evidence from combined analysis of 20 chloroplast fragments. *Mol. Phylogenet. Evol.* 54, 266–277. doi: 10.1016/j.ympev.2009.08.007

Tikapunya, T., Zou, W., Yu, W., Powell, P. O., Fox, G. P., Furtado, A., et al. (2017). Molecular structures and properties of starches of Australian wild rice. *Carbohydr. Polym.* 172, 213–222. doi: 10.1016/j.carbpol.2017.05.046

Toyomoto, D., Uemura, M., Taura, S., Sato, T., Henry, R., Ishikawa, R., et al. (2019). Segregation distortion observed in the progeny of crosses between *Oryza sativa* and *O. meridionalis* caused by abortion during seed development. *Plants (Basel)* 8:398. doi: 10.3390/plants8100398

Tsitrone, A., Kirkpatrick, M., and Levin, D. A. (2003). A model for chloroplast capture. *Evolution* 57, 1776–1782. doi: 10.1111/j.0014-3820.2003.tb00585.x

Vaughan, D. A. (1989). *The Genus Oryza L.: Current Status of Taxonomy*, Vol. 138. Manila: IRRI.

Vriesendorp, B., and Bakker, F. T. (2005). Reconstructing patterns of reticulate evolution in angiosperms: what can we do? *New Trends Plant Syst.* 54, 593–604. doi: 10.2307/25065417

Wambugu, P. W., Brozynska, M., Furtado, A., Waters, D. L., and Henry, R. J. (2015). Relationships of wild and domesticated rices (*Oryza* AA genome species) based upon whole chloroplast genome sequences. *Sci. Rep.* 5:13957. doi: 10.1038/srep13957

Ward, B. J., and van Oosterhout, C. (2016). HYBRIDCHECK: software for the rapid detection, visualization and dating of recombinant regions in genome sequence data. *Mol. Ecol. Resour.* 16, 534–539. doi: 10.1111/1755-0998.12469

Weitemier, K., Straub, S. C. K., Cronn, R. C., Fishbein, M., Schmickl, R., McDonnell, A., et al. (2014). Hyb-Seq: combining target enrichment and

genome skimming for plant phylogenomics. *Appl. Plant Sci.* 2:1400042. doi: 10.3732/apps.1400042

Wendel, J. F., and Doyle, J. J. (1998). "Phylogenetic incongruence: window into genome history and molecular evolution," in *Molecular Systematics of Plants II: DNA Sequencing*, eds D. E. Soltis, P. S. Soltis, and J. J. Doyle (Boston, MA: Springer US), 265–296. doi: 10.1007/978-1-4615-5419-6_10

Wendel, J. F., Stewart, J. M., and Rettig, J. H. (1991). Molecular evidence for homoploid reticulate evolution among Australian species of *Gossypium*. *Evolution* 45, 694–711. doi: 10.1111/j.1558-5646.1991.tb04339.x

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 in Plant Science

PERSPECTIVE
published: 29 March 2022
doi: 10.3389/fpls.2022.857896

The Next Generation of Rice: Inter-Subspecific *Indica-Japonica* Hybrid Rice

*Guiquan Zhang**

Guangdong Provincial Key Laboratory of Plant Molecular Breeding, State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangzhou, China

Rice (*Oryza sativa*) is an important food crop and has two subspecies, *indica* and *japonica*. Since the last century, four generations of rice varieties have been applied to rice production. Semi-dwarf rice, intra-subspecific hybrid rice, and inter-subspecific introgression rice were developed successively by genetic modification based on the first generation of tall rice. Each generation of rice has greater yield potential than the previous generation. Due to the stronger heterosis of *indica-japonica* hybrids, utilization of the inter-subspecific heterosis has long been of interest. However, *indica-japonica* hybrid sterility hinders the utilization of heterosis. In the past decades, *indica-japonica* hybrid sterility has been well understood. It is found that *indica-japonica* hybrid sterility is mainly controlled by six loci, *S5*, *Sa*, *Sb*, *Sc*, *Sd*, and *Se*. The *indica-japonica* hybrid sterility can be overcome by developing *indica*-compatible *japonica* lines (ICJLs) or wide-compatible *indica* lines (WCILs) using genes at the loci. With the understanding of the genetic and molecular basis of *indica-japonica* hybrid sterility and the development of molecular breeding technology, the development of *indica-japonica* hybrid rice has become possible. Recently, great progress has been made in breeding *indica-japonica* hybrid rice. Therefore, the *indica-japonica* hybrid rice will be the next generation of rice. It is expected that the *indica-japonica* hybrid rice will be widely applied in rice production in the near future.

Keywords: heterosis, hybrid sterility, subspecies, rice generation, hybrid rice

OPEN ACCESS

Edited by:
Kenneth Lee McNally, International Rice Research Institute (IRRI), Philippines

Reviewed by:
Deyong Ren, China National Rice Research Institute (CAAS), China

***Correspondence:**
Guiquan Zhang
gqzhang@scau.edu.cn

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

Received: *19 January 2022*
Accepted: *03 March 2022*
Published: *29 March 2022*

Citation:
Zhang G (2022) The Next Generation of Rice: Inter-Subspecific Indica-Japonica Hybrid Rice. Front. Plant Sci. 13:857896. doi: 10.3389/fpls.2022.857896

INTRODUCTION

Rice (*O. sativa*) is the best-known cultivated crop, providing staple food for more than half of the world's population (Fukagawa and Ziska, 2019). The domestication of cultivated rice began about 10,000 years ago. During domestication, the cultivated rice differentiated into two varietal groups. Since the Han dynasty, the Chinese have recognized two rice varietal groups, Hsien (Xian) and Keng (Geng) (Ting, 1949; Wang et al., 2018). In the last century, Kato et al. (1928) divided *O. sativa* into two subspecies, *indica* and *japonica*. Ting (1949, 1957) named the two subspecies as *hsien* and *keng*. Recently, the two subspecies are considered to be *Xian*/Indica (XI) and *Geng*/Japonica (GJ) (Wang et al., 2018). The two subspecies of cultivated rice are two different ecological types, where *indica* rice is suitable for growing in tropical and subtropical areas at low latitudes and low altitudes, while *japonica* rice is suitable for growing in temperate areas at high latitudes or high altitudes (Chang, 1976; Khush, 1997).

Rice is a self-pollinating plant. The heterosis of self-pollinating plants is usually considered to be weak. In the 1970s, *indica* hybrid rice was successfully applied in China. Since the 1980s, *indica* hybrid rice has been widely planted in China, accounting for more than 50% of the rice planting area (Cheng et al., 2007). The *indica* hybrid rice shows strong heterosis, which plays an important role in increasing rice yield (Yuan and Virmani, 1988; Zhang et al., 2021). At the same time, *japonica* hybrid rice has also been successfully developed, and a large number of *japonica* hybrid varieties have been released (Li and Wu, 1991; Zheng et al., 2020). It is believed that inter-subspecific hybrids have stronger heterosis than intra-subspecific hybrids (Fu et al., 2014; Birchler, 2015). Rice inter-subspecific heterosis has long been attempted, but the sterility of inter-subspecific hybrids has hindered the utilization of heterosis (Cheng et al., 2007; Zhang, 2020). Therefore, the key to utilization of inter-subspecific heterosis is to overcome the hybrid sterility.

THE FIVE GENERATIONS OF RICE

Although domestication of the cultivated rice started about 10,000 years ago, purposeful genetic improvement of varieties began in the last century. According to the genetic basis of varieties, the cultivated rice can be divided into five generations (Zhang, 2019, 2020).

The first generation (1G) of rice is tall rice. Since rice was cultivated about 10,000 years ago, the cultivated rice had been tall rice. Tall rice with high stalk was easy to cover the weeds in the field. Before the 1960s, tall rice was suitable for cultivating without chemical fertilizer. Tall rice had many landraces, which were the result of local intuitive selection by farmers for a long time (Zeven, 1998). As the first generation of rice, tall varieties have become an important genetic resource for rice breeding.

The second generation (2G) of rice is semi-dwarf rice. Around the 1950s, farmers began using chemical fertilizer. The yield of tall rice was greatly improved, but with it came the lodging. From lodging tall varieties, semi-dwarf mutants were selected as new varieties. In 1956, for example, two farmers of Guangdong province of China selected a semi-dwarf mutant from the field of lodging tall rice variety Nan-te 16. This semi-dwarf mutant became a new semi-dwarf variety Ai-jiao-nan-te, which was soon widely planted in southern China (Hu, 1965). At the same time, some semi-dwarf germplasm resources were selected to develop new semi-dwarf varieties by hybridization breeding. In 1956, for example, a breeding team of Guangdong Academy of Agricultural Sciences of China selected Ai-zi-zhan as a semi-dwarf parent and developed new semi-dwarf varieties Guang-chang-ai in 1959, Zhen-zhu-ai in 1961, and Guang-liu-ai 4 in 1966 (Guangdong Academy of Agricultural Sciences, 1966). Meanwhile, a semi-dwarf variety with high yield potential, IR8, was released by International Rice Research Institute (IRRI) in 1966 (Peng et al., 1994; Khush, 2001). The semi-dwarf varieties were bred by incorporating a recessive dwarf gene, *sd1*, to reduce plant height (Suh and Hue, 1978; Khush, 2001).

The third generation (3G) of rice is intra-subspecific hybrid rice, including *indica* hybrid rice and *japonica* hybrid rice. In the 1970s, *indica* hybrid rice was developed in China by a group of scientists led by Longping Yuan. Since then, the *indica* hybrid rice has been rapidly applied to production because of its strong heterosis (Yuan and Virmani, 1988). Meanwhile, *japonica* hybrid rice has also been developed and applied in rice production (Shiniyo, 1969; Li and Wu, 1991; Zheng et al., 2020).

The fourth generation (4G) of rice is inter-subspecific introgression rice, including *japonica*-introgressive *indica* rice and *indica*-introgressive *japonica* rice. Since the 1970s, restorer genes have been transferred from *indica* to *japonica* to develop *japonica* restorer lines for *japonica* hybrid rice. In the 1980s, the finding of *S5-n* gene (Ikehashi and Araki, 1986) and the "new plant type" program of IRRI (Peng et al., 2008) promoted the hybridization breeding between *indica* and *japonica* rice. Since then, more and more inter-subspecific introgression varieties have been developed and applied in rice production (Cheng et al., 2007; Ma et al., 2007; Lin et al., 2016).

The fifth generation (5G) of rice is inter-subspecific *indica-japonica* hybrid rice. The next generation rice has heterosis between *indica* and *japonica* subspecies, which will greatly improve the yield potential (Zhang, 2020). It is worth noting that many so-called *indica-japonica* hybrid varieties released are actually inter-subspecific introgression varieties that are 4G rice (Ma et al., 2007; Lin et al., 2016; Zhu et al., 2020).

Since the second half of the last century, rice breeding has developed rapidly. With the utilization of new genetic resources and new breeding techniques, the genetic basis of varieties has changed more and more. The yield potential of each generation of varieties has been greatly improved.

GENETIC BASIS OF *INDICA-JAPONICA* HYBRID STERILITY

Reproductive isolation usually appears in inter-specific and inter-subspecific hybrids of plants. Reproductive isolation can occur at the prezygotic and postzygotic stages. Postzygotic reproductive isolation usually shows hybrid lethality, hybrid necrosis/weakness and hybrid sterility (Baack et al., 2015; Ouyang and Zhang, 2018). In the past decades, about 50 loci related to reproductive isolation of the genus *Oryza* have been identified (Ouyang and Zhang, 2013, 2018; Guo et al., 2016; Li et al., 2020). In *indica-japonica* crosses, reproductive isolation usually shows hybrid sterility. Among the loci for reproductive isolation of *Oryza*, only some of the loci are responsible for the hybrid sterility of *indica-japonica* crosses (Zhang, 2020).

In *indica-japonica* hybrids, female or embryo sac sterility is controlled by the *S5* locus. At the locus, *indica* varieties usually have *S5-i* allele, while *japonica* varieties usually have *S5-j* allele. The interaction of *S5-i* and *S5-j* in *indica-japonica* hybrids causes the abortion of female gametes with the *S5-j* allele (Ikehashi and Araki, 1986). By genetic mapping, the *S5* locus was located on chromosome 6 (Ikehashi and Araki, 1986; Yanagihara et al., 1995). Furthermore, the *S5* gene was cloned and functionally

analyzed (Chen et al., 2008; Yang et al., 2012). In the hybrids of wide *indica-japonica* crosses, female sterility is usually under the control of the *S5* locus (Ikehashi and Araki, 1986; Song et al., 2005).

For male or pollen sterility of *indica-japonica* hybrids, five loci, *Sa*, *Sb*, *Sc*, *Sd*, and *Se*, were identified in wide *indica-japonica* crosses. At the loci, *indica* varieties usually have *S-i* allele, while *japonica* varieties usually have *S-j* allele. In *indica-japonica* hybrids, the interaction between *S-i* and *S-j* at the loci leads to the abortion of male gametes with the *S-j* allele. The male sterility shows two types of abortive pollens. The empty abortive pollen is caused by the *Sa* locus, while the stained abortive pollen is caused by the *Sb*, *Sc*, *Sd*, and *Se* loci. The degree of pollen sterility in *indica-japonica* hybrids depends on the number of heterozygous loci (Zhang and Lu, 1989, 1993, 1996; Zhang et al., 1993, 1994). By molecular mapping, the *Sa*, *Sb*, *Sc*, *Sd*, and *Se* loci were located on chromosomes 1, 5, 3, 1, and 12, respectively (Zhuang et al., 1999, 2002; Zhang and Zhang, 2001; Su and Liu, 2003; Yang et al., 2004; Li et al., 2006, 2008; Zhu et al., 2008). Furthermore, the *Sa* and *Sc* genes have been cloned and functionally analyzed (Long et al., 2008; Shen et al., 2017; Xie et al., 2017). The genes for hybrid sterility *S24*, *S35,* and *S25* (*S36*) are found to be located in the same chromosomal regions as *Sb*, *Sd,* and *Se*, respectively, which may be the same loci (Kubo and Yoshimura, 2001; Wen et al., 2007; Kubo et al., 2008; Zhao et al., 2011). Therefore, the male sterility of *indica-japonica* hybrids is usually under the control of the five loci (Zhang, 2020).

Neutral (n) allele is usually found at the loci for hybrid sterility in plants. When two alleles of a locus interact to cause sterility, there may be a third allele, n allele, at the locus, whose interaction with other two alleles can't cause sterility (Rich, 1966). At *S5* locus, some tropical *japonica* accessions carry *S5-n* allele except *indica* varieties having *S5-i* and *japonica* varieties having *S5-j* (Ikehashi and Araki, 1986). At the *Sa*, *Sb*, *Sc*, *Sd*, and *Se* loci, not only *S-i*, *S-j,* and *S-n* alleles can be divided, but the effects of alleles from different donors also vary quantitatively, resulting in the continuous variation of pollen sterility at a single locus (Zhang et al., 1993). The molecular basis of neutral alleles has been revealed by the cloned genes of *S5* (Chen et al., 2008; Yang et al., 2012), *Sa* (Long et al., 2008; Xie et al., 2017), and *Sc* (Shen et al., 2017).

Summarily, six loci of hybrid sterility are usually found in *indica-japonica* crosses, *S5* for female sterility, and *Sa*, *Sb*, *Sc*, *Sd*, and *Se* for male sterility (**Table 1** and **Figure 1A**). Generally, *indica* varieties have *S-i* allele, *japonica* varieties have *S-j* allele, while some accessions have *S-n* allele at these loci. The genic model of the loci is the one-locus sporo-gametophytic interaction model. In *indica-japonica* hybrids, the allelic interaction of *S-i* and *S-j* causes the abortion of female gametes carrying the *S-j* allele of *S5* locus, and the abortion of male gametes carrying the *S-j* allele of *Sa*, *Sb*, *Sc*, *Sd*, and *Se* loci, resulting in hybrid sterility (**Figure 1B**). In contrast, the interaction of *S-n* with *S-i* or *S-j* can't cause the abortion of any gametes (Ikehashi and Araki, 1986; Zhang, 2020). The understanding of the genetic basis of *indica-japonica* hybrid sterility has laid the foundation for overcoming the hybrid sterility.

STRATEGIES FOR OVERCOMING *INDICA-JAPONICA* HYBRID STERILITY

Based on the genetic basis of *indica-japonica* hybrid sterility, two types of breeding lines can be developed to overcome the hybrid sterility. They are *indica*-compatible *japonica* lines (ICJLs) (Zhang et al., 1994; Zhang and Lu, 1999) and wide-compatibility lines (WCLs). The breeding lines can be used to develop *indica-japonica* hybrid rice without hybrid sterility (Zhang, 2020).

The ICJLs can be developed by transferring *S-i* alleles at the *S5, Sa, Sb*, *Sc*, *Sd,* and *Se* loci from *indica* to *japonica* by backcrossing (**Figure 1C**). For example, a set of Taichung 65 (T65) isogenic F_1-sterile lines (TISLs) having *S-i* alleles at the *Sa, Sb*, *Sc*, *Sd,* and *Se* loci were developed using a set of *indica* varieties as *S-i* donors in the genetic background of T65, a *japonica* variety. Then, the *S-i* alleles of these loci were pyramided together by crossing the TISLs. The hybrid pollen fertility of pyramiding lines with different genotypes at these loci were tested with *indica* and *japonica* testers. The results showed that as the number of *S-i* alleles at these loci in the pyramiding lines increased, the pollen fertility of hybrids with *indica* testers increased, while that of hybrids with *japonica* testers decreased (Zhang and Lu, 1996; Guo et al., 2016). Furthermore, by pyramiding the *S5-n* allele in the pyramiding lines with the *S-i* alleles of the *Sb*, *Sc*, *Sd,* and *Se* loci, several ICJLs with *Sb-i*, *Sc-i*, *Sd-i*, *Se-i,* and *S5-n* alleles in *japonica* genetic background were developed. The ICJLs showed normal or near normal pollen fertility and spikelet fertility in their hybrids with *indica* testers, but serious pollen sterility and spikelet sterility in the hybrids with *japonica* testers. Therefore, the *indica-japonica* hybrid sterility can be overcome in the crosses of ICJLs with *indica* varieties (Guo et al., 2016).

The WCLs can be developed by using the *S-n* alleles of six loci. WCLs with *indica* genetic background are wide-compatible *indica* lines (WCILs), and WCLs with *japonica* genetic background are wide-compatible *japonica* lines (WCJLs). The *indica-japonica* hybrid rice can be developed by using WCILs crossed with *japonica* lines, or by using WCJLs crossed with *indica* lines (**Figure 1D**).

DISCUSSION

Since the 1970s, the intra-subspecific hybrid rice, including *indica* hybrid rice and *japonica* hybrid rice, has been developed. The cytoplasmic male sterility (CMS) system is a three-line system including CMS line, maintainer line and restorer line (Yuan and Virmani, 1988; Chen and Liu, 2014). The photoperiod/thermosensitive genic male-sterility (PTGMS) system is a two-line system including PTGMS line and restorer line (Shi, 1985; Ding et al., 2012; Zhou et al., 2012). The CMS system and the PTGMS system have been widely utilized in hybrid rice production. The application of genic male-sterility (GMS) materials in hybrid rice is the third-generation hybrid rice breeding technology (Deng et al., 2013; Wang and Deng, 2018; Song et al., 2021). A large number of breeding lines have been developed by the three generations of hybrid rice breeding technology. These techniques and breeding lines can

TABLE 1 | The loci for *indica-japonica* hybrid sterility.

Sterility	Locus	Chr.	Molecular mechanism	References		
				Identification	**Molecular mapping**	**Cloning and functional analysis**
Female	*S5*	6	A killer-protector system encoded by three tightly linked genes	Ikehashi and Araki, 1986	Yanagihara et al., 1995; Ji et al., 2005; Qiu et al., 2005	Chen et al., 2008; Yang et al., 2012
Male	*Sa*	1	A two-gene/three component interaction model	Zhang and Lu, 1989, 1993; Zhang et al., 1993, 1994	Zhuang et al., 1999; Su and Liu, 2003	Long et al., 2008; Xie et al., 2017
	Sb	5			Zhuang et al., 2002; Li et al., 2006	
	Sc	3	A model of gene dosage-dependent hybrid male sterility		Zhang and Zhang, 2001; Yang et al., 2004	Shen et al., 2017
	Sd	1			Li et al., 2008	
	Se	12			Zhu et al., 2008	

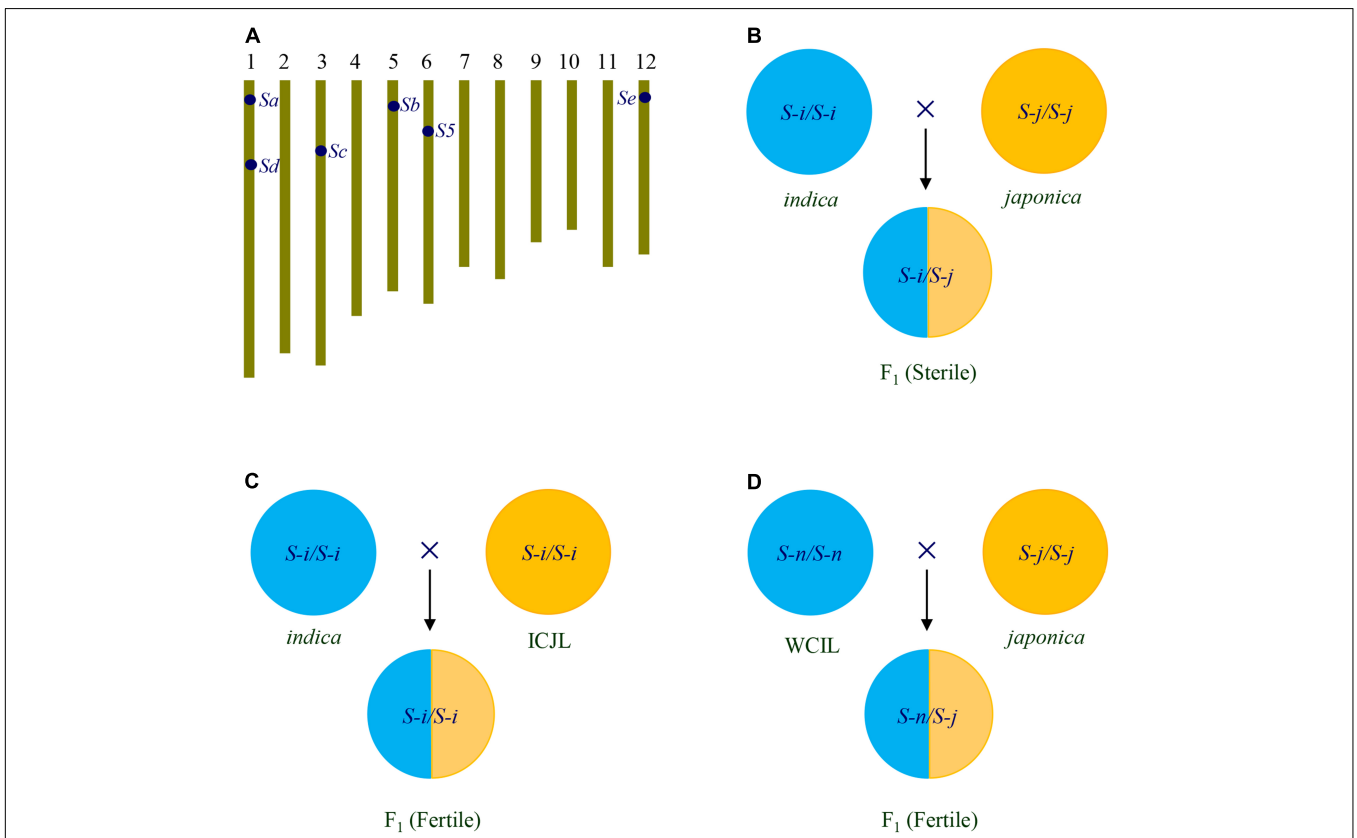

FIGURE 1 | Strategies for overcoming *indica-japonica* hybrid sterility. **(A)** Chromosomal location of the *S5*, *Sa*, *Sb*, *Sc*, *Sd,* and *Se* loci for *indica-japonica* hybrid sterility. **(B)** Hybrid of *indica*/*japonica* is sterile, which is caused by the interaction between *S-i* and *S-j* at the *S5*, *Sa*, *Sb*, *Sc*, *Sd,* and *Se* loci. **(C)** Hybrid of *indica*/ICJL is fertile due to ICJL carrying *S-i* at the *S5*, *Sa*, *Sb*, *Sc*, *Sd,* and *Se* loci. **(D)** Hybrid of WCIL/*japonica* is fertile due to WCIL carrying *S-n* at the *S5*, *Sa*, *Sb*, *Sc*, *Sd,* and *Se* loci. ICJL, *indica*-compatible *japonica* line; WCIL, wide-compatible *indica* line.

be used to develop not only intra-subspecific hybrid rice but also inter-subspecific hybrid rice. For examples, the CMS lines, PTGMS lines, and GMS lines are male sterility lines (MSLs) that can also be used for the breeding of inter-subspecific hybrid rice. The restorer lines of CMS system, PTGMS system and GMS system can be used for the breeding of inter-subspecific hybrid rice after improving their compatibility. Thus, the breeding of intra-subspecific hybrid rice provides available breeding lines for the development of inter-subspecific hybrid rice.

Compared with intra-subspecific hybrid rice and inter-subspecific introgression rice, the development of inter-subspecific hybrid rice will face greater challenges. Firstly, the overcoming of *indica-japonica* hybrid sterility requires to pyramid multiple genes for compatibility. Secondly,

male sterility and fertility restoration should be considered in hybrid rice breeding. Third, there may be some problems caused by the remote genetic backgrounds of the two subspecies (Lin et al., 2016). To meet the challenge, it is necessary to develop molecular breeding techniques. For the past two decades, we have been building a library of chromosome single-segment substitution lines (SSSLs) to construct a target chromosome-segment substitution platform for rice design (Zhang, 2021). The SSSL library was constructed by using forty-three accessions from seven species with AA genome as donors of chromosome segments in the genetic background of Huajingxian 74 (HJX74), an elite *indica* variety in South China. The HJX74-SSSL library consists of 2,360 SSSLs, which collects rich gene resources from donors with genetic diversity (Zhang et al., 2004; Xi et al., 2006; He et al., 2017; Zhao et al., 2019; Zhang, 2021). The HJX74-SSSLs have been used to detect QTLs for complex traits (Zhang et al., 2012; Yang et al., 2016, 2021a,b; Zhou et al., 2017; Tan et al., 2020, 2021, 2022; Pan et al., 2021), to clone genes of agronomic importance and to assess allelic variation (Wang et al., 2008, 2012, 2015; Teng et al., 2012; Sui et al., 2019; Zhang et al., 2020; Gao et al., 2021). Using the HJX74-SSSL library as platform for rice breeding by design, several CMS, maintainer and restorer lines have been developed (Dai et al., 2015, 2016; Luan et al., 2019). These results suggest that the target chromosome-segment substitution is an effective way to rice breeding by design (Zhang, 2021). Recently, the HJX74-SSSL library was used to develop WCILs. Through the restorer gene pyramiding, the WCILs will be developed into wide-compatible *indica* restorer lines (WCIRLs) as restorer lines of *indica-japonica* hybrid rice. Therefore, the target chromosome-segment substitution based on HJX74-SSSL platform provides technical support for the development of *indica-japonica* hybrid rice.

For a century, four generations of rice have provided a large number of elite varieties for rice production. With the changes of genetic basis, each generation of rice has greater yield potential than the previous generation. The application of new generation rice has greatly improved the productivity of modern rice. However, the intra-subspecific hybrid rice can only have intra-subspecific heterosis, and the inter-subspecific introgression rice can only utilize partial inter-subspecific heterosis. In comparison, the inter-subspecific *indica-japonica* hybrid rice can take advantage of complete inter-subspecific heterosis. Therefore, the utilization of heterosis between *indica* and *japonica* subspecies has been expected (Cheng et al., 2007; Zhang, 2020). With the understanding of the genetic and molecular basis of *indica-japonica* hybrid sterility and the development of molecular breeding techniques, it is now possible to develop *indica-japonica* hybrid rice. We are developing *indica-japonica* hybrid rice by crossing WCIRLs developed on the HJX74-SSSL platform with existing *japonica* CMS lines collected from *japonica* planting areas. Many *indica-japonica* hybrid rice combinations have been extensively tested in various rice planting areas of China. Our results showed that *indica-japonica* hybrid rice had stronger heterosis and higher yield potential. Therefore, *indica-japonica* hybrid rice will become the next generation of rice and will be widely applied in rice production in the near future.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

GZ wrote the manuscript independently.

FUNDING

This research was funded by grants from the National Natural Science Foundation of China (91435207 and 91735304) and the Major Program of Transgenic New Variety Breeding of China (2014ZX08009-037B).

REFERENCES

Baack, E., Melo, M. C., Rieseberg, L. H., and Ortiz-Barrientos, D. (2015). The origins of reproductive isolation in plants. *New Phytol.* 207, 968–984. doi: 10.1111/nph.13424

Birchler, J. A. (2015). Heterosis: the genetic basis of hybrid vigour. *Nat. Plants.* 1:15020. doi: 10.1038/nplants.2015.20

Chang, T. T. (1976). The origin, evolution, cultivation, dissemination, and diversification of Asian and African rices. *Euphytica* 25, 425–441. doi: 10.1007/bf00041576

Chen, J., Ding, J., Ouyang, Y., Du, H., Yang, J., Cheng, K., et al. (2008). A triallelic system of *S5* is a major regulator of the reproductive barrier and compatibility of *indica-japonica* hybrids in rice. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11436–11441. doi: 10.1073/pnas.0804761105

Chen, L., and Liu, Y. (2014). Male sterility and fertility restoration in crops. *Annu. Rev. Plant Biol.* 65, 579–606. doi: 10.1146/annurev-arplant-050213-040119

Cheng, S., Cao, L., Zhuang, J., Chen, S., Zhan, X., Fan, Y., et al. (2007). Super hybrid rice breeding in China: Achievements and prospects. *J. Integr. Plant Biol.* 49, 805–810. doi: 10.16288/j.yczz.18-213

Dai, Z., Lu, Q., Luan, X., Cai, J., Zhu, H., Liu, Z., et al. (2015). Development of a platform for breeding by design of CMS lines based on an SSSL library in rice (*Oryza sativa* L.). *Euphytica* 205, 63–72.

Dai, Z., Lu, Q., Luan, X., Ouyang, L., Guo, J., Liang, J., et al. (2016). Development of a platform for breeding by design of CMS restorer lines based on an SSSL library in rice (*Oryza sativa* L.). *Breed. Sci.* 66, 768–775. doi: 10.1270/jsbbs.16044

Deng, X., Wang, H., Tang, X., Zhou, J., Chen, H., He, G., et al. (2013). Hybrid rice breeding welcomes a new era of molecular crop design. *Sci. Sin. Vitae.* 43, 864–868. doi: 10.1360/052013-299

Ding, J., Lu, Q., Ouyang, Y., Mao, H., Zhang, P., Yao, J., et al. (2012). A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2654–2659. doi: 10.1073/pnas.1121374109

Fu, D., Xiao, M., Hayward, A., Fu, Y., Liu, G., Jiang, G., et al. (2014). Utilization of crop heterosis: a review. *Euphytica* 197, 161–173. doi: 10.1007/s10681-014-1103-7

Fukagawa, N. K., and Ziska, L. H. (2019). Rice: Importance for global nutrition. *J. Nutr. Sci. Vitaminol.* 65, S2–S3. doi: 10.3177/jnsv.65.S2

Gao, M., He, Y., Yin, X., Zhong, X., Yan, B., Wu, Y., et al. (2021). Ca^{2+} sensor-mediated ROS scavenging suppresses rice immunity and is exploited by a fungal effector. *Cell* 184, 5391–5404. doi: 10.1016/j.cell.2021.09.009

Guangdong Academy of Agricultural Sciences. (1966). Preliminary summary of rice dwarf breeding in Guangdong Province. *Acta Agron. Sin.* 5, 33–40.

Guo, J., Xu, X., Li, W., Zhu, W., Zhu, H., Liu, Z., et al. (2016). Overcoming inter-subspecific hybrid sterility in rice by developing *indica*-compatible *japonica* lines. *Sci. Rep*. 6:26878. doi: 10.1038/srep26878

He, N., Wu, R., Pan, X., Peng, L., Sun, K., Zou, T., et al. (2017). Development and trait evaluation of chromosome single-segment substitution lines of *O. meridionalis* in the background of *O. sativa*. *Euphytica* 213:281.

Hu, Q. (1965). Experience of Ai-jiao-nan-te in producing 500 kilogram per mu. *Hubei Agri. Sci.* 1965, 24–29.

Ikehashi, H., and Araki, H. (1986). *Genetics of F_1 sterility in remote crosses in rice. In Rice Genetics: Proceedings of the First Rice Genetics Symposium*. Philippines: International Rice Research Institute, 119–130.

Ji, Q., Lu, J., Chao, Q., Gu, M., and Xu, M. (2005). Delimiting a rice wide-compatibility gene *$S5^n$* to a 50 kb region. *Theor. Appl. Genet.* 111, 1495–1503. doi: 10.1007/s00122-005-0078-0

Kato, S., Kosaka, H., and Hara, S. (1928). On the affinity of rice varieties as shown by fertility of hybrid plants. *Sci. Bull. Faculty Agric. Kyushu Univ*. 3, 132–147.

Khush, G. S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant Mol. Biol.* 35, 25–34. doi: 10.1007/978-94-011-5794-0_3

Khush, G. S. (2001). Green revolution: The way forward. *Nat. Rev. Genet*. 2, 815–822. doi: 10.1038/35093585

Kubo, T., Yamagata, Y., Eguchi, M., and Yoshimura, A. (2008). A novel epistatic interaction at two loci causing hybrid male sterility in an inter-subspecific cross of rice (*Oryza sativa* L.). *Genes Genet. Syst*. 83, 443–453. doi: 10.1266/ggs.83.443

Kubo, T., and Yoshimura, A. (2001). Linkage analysis of an F_1 sterility gene in *Japonica/Indica* cross of rice. *Rice Genet. Newslett*. 18, 52–54.

Li, J., Zhou, J., Zhang, Y., Yang, Y., Pu, Q., and Tao, D. (2020). New insights into the nature of interspecific hybrid sterility in rice. *Front. Plant Sci*. 11:555572. doi: 10.3389/fpls.2020.555572

Li, W., Zeng, R., Zhang, Z., Ding, X., and Zhang, G. (2006). Fine mapping of locus *S-b* for F_1 pollen sterility in rice (*Oryza sativa* L.). *Chinese Sci. Bull*. 51, 675–680. doi: 10.1007/s11434-006-0675-6

Li, W., Zeng, R., Zhang, Z., Ding, X., and Zhang, G. (2008). Identification and fine mapping of *S-d*, a new locus conferring the partial pollen sterility of intersubspecific F_1 hybrids in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 116, 915–922. doi: 10.1007/s00122-008-0723-5

Li, Z., and Wu, J. (1991). The present situation and prospect of three-line hybrid *japonica* rice breeding in China. *Hybrid Rice* 15, 13–16.

Lin, J.-R., Song, X.-W., Wu, M.-G., and Cheng, S.-H. (2016). Breeding technology innovation of *indica–japonica* super hybrid rice and varietal breeding. *Sci. Agric. Sin.* 49, 207–218.

Long, Y., Zhao, L., Niu, B., Su, J., Wu, H., Chen, Y., et al. (2008). Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18871–18876. doi: 10.1073/pnas.0810108105

Luan, X., Dai, Z., Yang, W., Tan, Q., Lu, Q., Guo, J., et al. (2019). Breeding by design of CMS lines on the platform of SSSL library in rice. *Mol. Breed* 39:126.

Ma, R. R., Xu, D. H., Wang, X. Y., Yu, S. M., Jin, Q. Y., Ouyang, Y. N., et al. (2007). Heterosis on plant morphology of Yongyou 6, an *indica-japonica* inter-subspecific super high-yielding hybrid rice. *Chinese J. Rice Sci.* 21, 281–286.

Ouyang, Y., and Zhang, Q. (2013). Understanding reproductive isolation based on the rice model. *Annu. Rev. Plant Biol.* 64, 111–135. doi: 10.1146/annurev-arplant-050312-120205

Ouyang, Y., and Zhang, Q. (2018). The molecular and evolutionary basis of reproductive isolation in plants. *J. Genet. Genomics.* 45, 613–620. doi: 10.1016/j.jgg.2018.10.004

Pan, Z., Tan, B., Cao, G., Zheng, R., Liu, M., Zeng, R., et al. (2021). Integrative QTL identification, fine mapping and candidate gene analysis of a major locus *qLTG3a* for seed low-temperature germinability in rice. *Rice* 14:103. doi: 10.1186/s12284-021-00544-2

Peng, S., Khush, G. S., and Cassman, K. G. (1994). "Evolution of the new plant ideotype for increased yield potential," in *Breaking the Yield Barrier: Proceedings of a Workshop on Rice Yield Potential in Favorable Environments*, ed. K. G. Cassman (Philippines: International Rice Research Institulte), 5–20. .

Peng, S., Khush, G. S., Virk, P., Tang, Q., and Zou, Y. (2008). Progress in ideotype breeding to increase rice yield potential. *Field Crops Res*. 108, 32–38. doi: 10.1016/j.fcr.2008.04.001

Qiu, S. Q., Liu, K., Jiang, J. X., Song, X., Xu, C. G., Li, X. H., et al. (2005). Delimitation of the rice wide compatibility gene *$S5^n$* to a 40-kb DNA fragment. *Theor. Appl. Genet*. 111, 1080–1086. doi: 10.1007/s00122-005-0033-0

Rich, C. M. (1966). Abortion of male and female gametes in the tomato determined by allelic interaction. *Genetics* 53, 85–96. doi: 10.1093/genetics/53.1.85

Shen, R., Wang, L., Liu, X., Wu, J., Jin, W., Zhao, X., et al. (2017). Genomic structural variation-mediated allelic suppression causes hybrid male sterility in rice. *Nat. Commun.* 8:1310. doi: 10.1038/s41467-017-01400-y

Shi, M. (1985). The discovery and preliminary studies of the photoperiod-sensitive recessive male sterile rice (*Oryza sativa* L. subsp. *japonica*). *Sci. Agric. Sin*. 2, 44–48.

Shiniyo, C. (1969). Cytoplasmic-genetic male sterility in cultivated rice. *Oryza sativa* L. *J. Genet*. 44, 149–156. doi: 10.1266/jjg.44.149

Song, S., Wang, T., Li, Y., Hu, J., Kan, R., Qiu, M., et al. (2021). A novel strategy for creating a new system of third-generation hybrid rice technology using a cytoplasmic sterility gene and a genic male-sterile gene. *Plant Biotechnol. J.* 19, 251–260. doi: 10.1111/pbi.13457

Song, X., Qiu, S., Xu, C., Li, X., and Zhang, Q. (2005). Genetic dissection of embryo sac fertility, pollen fertility, and their contributions to spikelet fertility of intersubspecific hybrids in rice. *Theor. Appl. Genet.* 110, 205–211. doi: 10.1007/s00122-004-1798-2

Su, J., and Liu, Y. (2003). Fine mapping and cloning of the gene *S-a* for F_1 pollen sterility in cultivated rice (*Oryza sativa* L.). *Mol. Plant Breed.* 1, 757–758.

Suh, H. S., and Hue, M. H. (1978). The segregation mode of plant height in the cross of rice varieties. XI. Linkage analysis of the semi-dwarfness of the rice variety 'Tongil'. *Korean J. Breed.* 10, 1–6. doi: 10.1270/jsbbs1951.10.1

Sui, F., Zhao, D., Zhu, H., Gong, Y., Tang, Z., Huang, X., et al. (2019). Map-based cloning of a new total loss-of-function allele of *OsHMA3* causes high cadmium accumulation in rice grain. *J. Exp. Bot*. 70, 2857–2871. doi: 10.1093/jxb/erz093

Tan, Q., Wang, C., Luan, X., Zheng, L., Ni, Y., Yang, W., et al. (2021). Dissection of closely linked QTLs controlling stigma exsertion rate in rice by substitution mapping. *Theor. Appl. Genet*. 134, 1253–1262. doi: 10.1007/s00122-021-03771-9

Tan, Q., Zhu, H., Liu, H., Ni, Y., Wu, S., Luan, X., et al. (2022). Fine mapping of QTLs for stigma exsertion rate from *Oryza glaberrima* by chromosome segment substitution. *Rice Sci*. 29, 55–66. doi: 10.1016/j.rsci.2021.12.005

Tan, Q., Zou, T., Zheng, M., Ni, Y., Luan, X., Li, X., et al. (2020). Substitution mapping of the major quantitative trait loci controlling stigma exsertion rate from *Oryza glumaepatula*. *Rice* 13:37. doi: 10.1186/s12284-020-00397-1

Teng, B., Zeng, R., Wang, Y., Liu, Z., Zhang, Z., Zhu, H., et al. (2012). Detection of allelic variation at the *Wx* locus with single-segment substitution lines in rice (*Oryza sativa* L.). *Mol. Breed*. 30, 583–595. doi: 10.1007/s11032-011-9647-x

Ting, Y. (1949). Origination of the rice cultivation in China. *J. College Agric. Sun Yat-sen Univ*. 7, 11–24.

Ting, Y. (1957). The origin and evolution of cultivated rice in China. *Acta Bot. Sin*. 8, 243–260.

Wang, E., Wang, J., Zhu, X., Hao, W., Wang, L., Li, O., et al. (2008). Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat. Genet*. 40, 1370–1374. doi: 10.1038/ng.220

Wang, H., and Deng, X. W. (2018). Development of the "third-generation" hybrid rice in China. *Genom. Proteom. Bioinf*. 16, 393–396. doi: 10.1016/j.gpb.2018.12.001

Wang, S., Li, S., Liu, Q., Wu, K., Zhang, J., Wang, S., et al. (2015). The *OsSPL16-GW7* regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nat. Genet.* 47, 949–955. doi: 10.1038/ng.3352

Wang, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., 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, 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, 43–49. doi: 10.1038/s41586-018-0063-9

Wen, J., Zhang, W., Jiang, L., Chen, L., Zhai, H., and Wan, J. (2007). Two novel loci for pollen sterility in hybrids between the weedy strain Ludao and the *japonica*

variety Akihikari of rice (*Oryza sativa* L.). *Theor. Appl. Genet*. 114, 915–925. doi: 10.1007/s00122-006-0489-6

Xi, Z., He, F., Zeng, R., Zhang, Z., Ding, X., Li, W., et al. (2006). Development of a wide population of chromosome single-segment substitution lines in the genetic background of an elite cultivar of rice (*Oryza sativa* L.). *Genome* 49, 476–484. doi: 10.1139/g06-005

Xie, Y., Niu, B., Long, Y., Li, G., Tang, J., Zhang, Y., et al. (2017). Suppression or knockout of *SaF*/*SaM* overcomes the *Sa*-mediated hybrid male sterility in rice. *J. Integr. Plant Biol*. 59, 669–679. doi: 10.1111/jipb.12564

Yanagihara, S., Mccouch, S. R., Ishikawa, K., Ogi, Y., Maruyama, K., and Ikehashi, H. (1995). Molecular analysis of the inheritance of the *S*-5 locus, conferring wide compatibility in *Indica/Japonica* hybrids of rice (*Oryza sativa* L.). *Theor. Appl. Genet*. 90, 182–188. doi: 10.1007/bf00222200

Yang, C., Chen, Z., Zhuang, C., Mei, M., and Liu, Y. (2004). Genetic and physical fine-mapping of the *Sc* locus conferring *indica-japonica* hybrid sterility in rice (*Oryza sativa* L.). *Chinese Sci. Bull*. 49, 1718–1721. doi: 10.1360/04wc0197

Yang, J., Zhao, X., Cheng, K., Du, H., Ouyang, Y., Chen, J., et al. (2012). A killer-protector system regulates both hybrid sterility and segregation distortion in rice. *Science* 337, 1336–1340. doi: 10.1126/science.1223702

Yang, T., Zhang, S., Zhao, J., Liu, Q., Huang, Z., Mao, X., et al. (2016). Identification and pyramiding of QTLs for cold tolerance at the bud bursting and the seedling stages by use of single segment substitution lines in rice (*Oryza sativa* L.). *Mol. Breed*. 36:96.

Yang, W., Liang, J., Hao, Q., Luan, X., Tan, Q., Lin, S., et al. (2021a). Fine mapping of two grain chalkiness QTLs sensitive to high temperature in rice. *Rice* 14:33. doi: 10.1186/s12284-021-00476-x

Yang, W., Xiong, L., Liang, J., Hao, Q., Luan, X., Tan, Q., et al. (2021b). Substitution mapping of two closely linked QTLs controlling grain shape on rice chromosome 8. *Rice* 14:85. doi: 10.1186/s12284-021-00526-4

Yuan, L., and Virmani, S. (1988). "Status of hybrid rice research and development," in *Hybrid rice*, (Manila: Internationa Rice Research Institute), 7–24.

Zeven, A. C. (1998). Landraces: a review of definitions and classifications. *Euphytica* 104, 127–139.

Zhang, G. (2019). Evolution and development of five generations of rice. *J. South China Agri. Univ*. 40, 211–216. doi: 10.1007/s11103-020-01037-4

Zhang, G. (2020). Prospects of utilization of inter-subspecific heterosis between *indica* and *japonica* rice. *J. Integr. Agr*. 19, 1–10. doi: 10.1016/s2095-3119(19)62843-1

Zhang, G. (2021). Target chromosome-segment substitution: A way to breeding by design in rice. *Crop J*. 9, 658–668. doi: 10.1016/j.cj.2021.03.001

Zhang, G., and Lu, Y. (1989). Genetic studies of the hybrid sterility in cultivated rice (Oryza sativa). I. Diallel analysis of the hybrid sterility among isogenic F_1 sterile lines. *Chinese J. Rice Sci*. 3, 97–101.

Zhang, G., and Lu, Y. (1993). Genetic studies of the hybrid sterility in cultivated rice (Oryza sativa). II. A genic model for F_1 pollen sterility. *Acta Genet. Sin*. 20, 222–228.

Zhang, G., and Lu, Y. (1996). "Genetics of F_1 pollen sterility in *Oryza sativa*," in *Rice Genetics III*, ed. G. Khush (Los Banos: International Rice Research Institute), 418–422.

Zhang, G., and Lu, Y. (1999). Breeding of the *indica*-compatible *japonica* lines and their use in the breeding of super-high-yield hybrid rice. *Hybrid Rice* 14, 3–5.

Zhang, G., Lu, Y., Liu, G., Yang, J., and Zhang, H. (1993). Genetic studies on the hybrid sterility in cultivated rice (*Oryza sativa*). III. Allele Differentiation of F_1 pollen sterility in different types of varieties. *Acta Genet. Sin*. 20, 541–551.

Zhang, G., Lu, Y., Zhang, H., Yang, J., and Liu, G. (1994). Genetic studies on the hybrid sterility in cultivated rice (*Oryza sativa*). IV. Genotypes for F_1 pollen sterility. *Acta Genet. Sin.* 21, 34–41.

Zhang, G., Zeng, R., Zhang, Z., Ding, X., Li, W., Liu, G., et al. (2004). The construction of a library of single segment substitution lines in rice (*Oryza sativa* L.). *Rice Genet. Newsl*. 21, 85–87.

Zhang, S., Huang, X., and Han, B. (2021). Understanding the genetic basis of rice heterosis: Advances and prospects. *Crop J*. 9, 688–692. doi: 10.1016/j.cj.2021.03.011

Zhang, S., Zhu, L., Shen, C., Ji, Z., Zhang, H., Zhang, T., et al. (2020). Natural allelic variation in a modulator of auxin homeostasis improves grain yield and nitrogen use efficiency in rice. *Plant Cell.* 33, 566–580. doi: 10.1093/plcell/koaa037

Zhang, Y., Yang, J., Shan, Z., Chen, S., Qiao, W., Zhu, X., et al. (2012). Substitution mapping of QTLs for blast resistance with SSSLs in rice (*Oryza sativa* L.). *Euphytica* 184, 141–150. doi: 10.1007/s10681-011-0601-0

Zhang, Z., and Zhang, G. (2001). Fine mapping of the *S-c* locus and marker-assisted selection using PCR markers in rice. *Acta Agron. Sin*. 27, 704–709.

Zhao, H., Sun, L., Xiong, T., Wang, Z., Liao, Y., Zou, T., et al. (2019). Genetic characterization of the chromosome single-segment substitution lines of *O. glumaepatula* and *O. barthii* and identification of QTLs for yield-related traits. *Mol. Breed*. 39:51.

Zhao, Z., Zhu, S., Zhang, Y., Bian, X., Wang, Y., Jiang, L., et al. (2011). Molecular analysis of an additional case of hybrid sterility in rice (*Oryza sativa* L.). *Planta* 233, 485–494. doi: 10.1007/s00425-010-1313-8

Zheng, W., Ma, Z., Zhao, M., Xiao, M., Zhao, J., Wang, C., et al. (2020). Research and development strategies for hybrid *japonica* rice. *Rice* 13:36. doi: 10.1186/s12284-020-00398-0

Zhou, H., Liu, Q., Li, J., Jiang, D., Zhou, L., Wu, P., et al. (2012). Photoperiod- and thermo-sensitive genic male sterility in rice are caused by a point mutation in a novel noncoding RNA that produces a small RNA. *Cell Res*. 22, 649–660. doi: 10.1038/cr.2012.28

Zhou, Y., Xie, Y., Cai, J., Liu, C., Zhu, H., Jiang, R., et al. (2017). Substitution mapping of QTLs controlling seed dormancy using single segment substitution lines derived from multiple cultivated rice donors in seven cropping seasons. *Theor. Appl. Genet*. 130, 1191–1205. doi: 10.1007/s00122-017-2881-9

Zhu, K., Zhou, Q., Shen, Y., Yan, J., Xu, Y., Wang, Z., et al. (2020). Agronomic and physiological performance of an *indica–japonica* rice variety with a high yield and high nitrogen use efficiency. *Crop Sci.* 60, 1556–1568. doi: 10.1002/csc2.20150

Zhu, W., Li, W., Ding, X., Zhang, Z., Zeng, R., Zhu, H., et al. (2008). Preliminary identification of F_1 pollen sterility gene *S-e* in *Oryza sativa*. *J. South China Agri. Univ*. 29, 1–5. doi: 10.15625/2615-9023/15888

Zhuang, C., Mei, M., Zhang, G., and Lu, Y. (2002). Chromosome mapping of the *S-b* locus for F_1 pollen sterility in cultivated rice (*Oryza sativa* L.) with RAPD markers. *Acta Genet. Sin*. 29, 700–705.

Zhuang, C., Zhang, G., Mei, M., and Lu, Y. (1999). Molecular mapping of the *S-a* locus for F_1 pollen sterility in cultivated rice (*Oryza sativa* L.). *Acta Genet. Sin.* 26, 213–218.

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

frontiers in Plant Science

ORIGINAL RESEARCH
published: 29 March 2022
doi: 10.3389/fpls.2022.866276

Fine Mapping of Two Interacting Loci for Transmission Ratio Distortion in Rice (*Oryza sativa* L.)

Chaopu Zhang[1,2†], Jilin Wang[1,2†], Xiongfeng Xiao[1,2], Dianwen Wang[1,2], Zhiyang Yuan[1,2], Xiaodan Zhang[1,2], Wenqiang Sun[1,2] and Sibin Yu[1,2,3]*

[1] National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China, [2] College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [3] Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan, China

Transmission ratio distortion (TRD) denotes the observed allelic or genotypic frequency deviation from the expected Mendelian segregation ratios in the offspring of a heterozygote. TRD can severely hamper gene flow between and within rice species. Here, we report the fine mapping and characterization of two loci (*TRD4.1* and *TRD4.2*) for TRD using large F_2 segregating populations, which are derived from rice chromosome segment substitution lines, each containing a particular genomic segment introduced from the *japonica* cultivar Nipponbare (NIP) into the *indica* cultivar Zhenshan (ZS97). The two loci exhibited a preferential transmission of ZS97 alleles in the derived progeny. Reciprocal crossing experiments using near-isogenic lines harboring three different alleles at *TRD4.1* suggest that the gene causes male gametic selection. Moreover, the transmission bias of *TRD4.2* was diminished in heterozygotes when they carried homozygous $TRD4.1^{ZS97}$. This indicates an epistatic interaction between these two loci. *TRD4.2* was mapped into a 35-kb region encompassing one candidate gene that is specifically expressed in the reproductive organs in rice. These findings broaden the understanding of the genetic mechanisms of TRD and offer an approach to overcome the barrier of gene flow between the subspecies in rice, thus facilitating rice improvement by introgression breeding.

Keywords: rice, reproductive isolation, transmission ratio distortion, allele frequency, gametic selection, epistatic interaction

OPEN ACCESS

Edited by:
Kazuki Matsubara, Institute of Crop Science (NARO), Japan

Reviewed by:
Sung-Ryul Kim, International Rice Research Institute (IRRI), Philippines
Yoshiyuki Yamagata, Kyushu University, Japan

***Correspondence:**
Sibin Yu
ysb@mail.hzau.edu.cn

† These authors have contributed equally to this work

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

Received: *31 January 2022*
Accepted: *01 March 2022*
Published: *29 March 2022*

Citation:
Zhang C, Wang J, Xiao X, Wang D, Yuan Z, Zhang X, Sun W and Yu S (2022) Fine Mapping of Two Interacting Loci for Transmission Ratio Distortion in Rice (Oryza sativa L.). Front. Plant Sci. 13:866276. doi: 10.3389/fpls.2022.866276

INTRODUCTION

Reproductive isolation is regarded as a driving force in the process of evolution in various species (Ouyang and Zhang, 2013; Baack et al., 2015). The development of reproductive isolation relies on the accumulation of genic incompatibilities, which can lead to non-Mendelian inheritance of alleles and genotypes in the offspring of heterozygotes (or hybrids) (Fishman et al., 2008; Leppälä et al., 2013). Transmission ratio distortion (TRD) is a naturally occurring phenomenon in which one allele is preferentially transmitted to the progeny than the opposite allele in hybrids between species (Jenczewski et al., 1997; Casellas et al., 2012). If one allele at a locus can reduce gametic or zygotic fitness, then the genomic regions linked to it will cause distorted allele or genotype frequencies in heterozygotes (Vogl and Xu, 2000; Xu et al., 2013). TRD is frequently observed and characterized in

intraspecific or interspecific segregating populations from various plant species, such as *Arabidopsis* (Leppälä et al., 2013; Seymour et al., 2019), cotton (Chandnani et al., 2017; Dai et al., 2017), wheat (Kumar et al., 2007), and rice (Koide et al., 2008a,b, 2012; Li et al., 2017, 2019). In particular, TRD is one of the primary origins of reproductive isolation or speciation and can severely hamper the exchange of genes among rice subspecies. Understanding the mechanisms responsible for TRD is important for using agriculturally interesting alleles from rice germplasm.

Transmission ratio distortion occurs before or after fertilization due to various reasons, such as meiotic drive, gametic competition, inbreeding depression, and hybrid incompatibilities (Huang et al., 2013; Fishman and McIntosh, 2019). It can be caused either by incompatible allelic interaction at a single-locus (Bauer et al., 2007; Corbett-Detig et al., 2013) or by two or multi-loci interaction, in which one gene effect is dependent on the presence/absence of other genes (Giesbers et al., 2019). With regard to rice, numerous loci/regions have been identified for TRD (Li et al., 2017, 2019; Zhang et al., 2020a). A few of them have been cloned and characterized in rice. They revealed that a given locus usually consists of multiple tightly linked genes and it is easily affected by the complex genetic background. For example, several well-known killer–protector and toxin–antidote segregation distortion systems that consist of multiple tightly linked genes have been identified. A killer–protector system contains three closely linked genes at the *S5* locus that regulate both hybrid fertility and segregation distortion (Yang et al., 2012). The toxin–antidote system of *qHMS7* contains two tightly linked genes (*ORF2* and *ORF3*), of which *ORF2* encodes a toxic genetic element that can kill the pollen without the protection of *ORF3*, leading to segregation distortion in heterozygotes (Yu et al., 2018). The *S1* locus, constituted of three genes *S1A4, S1TPR,* and *S1A6,* also demonstrates a killer–protector system that can eliminate the gametes carrying the Asian allele (*S1-s*), resulting in a preferential transmission of the African rice *S1* allele to the progeny (Xie et al., 2017, 2019). TRD might occur due to two or multi-loci interactions, such as S27/S28, DPL1/DPL2, and S25/S24 regions (Mizuta et al., 2010; Yamagata et al., 2010; Kubo et al., 2011; Nguyen et al., 2017). Despite these well-investigated examples, understanding the genetic and molecular mechanisms of TRD is still incomplete. Therefore, it is critical to identify more loci and candidate genes for dissecting the genetic and molecular basis of such a complex phenomenon.

Our previous studies, using backcross inbred lines (BIL) derived from an intersubspecific cross of the *japonica* cultivar Nipponbare (NIP) and the *indica* cultivar 9,311 identified and validated a number of genomic regions for TRD in rice (Zhang et al., 2020a). Two (*TRD4.1* and *TRD4.2*) were identified on chromosome 4 in which alleles were significantly skewed toward 9,311 in the BIL. In this study, the main objective is to fine-map these two TRD loci and identify their interaction pattern using several F_2 segregating populations, which were derived from chromosome segment substitution lines, each containing a particular NIP genomic segment with the Zhenshan 97 (ZS97) background in rice. We also used a reciprocal crossing approach to detect whether *TRD4.1* is involved in male or female gametic selection. Our findings will help shed light on the genetic mechanisms underlying TRD to improve the utilization of subspecific introgression breeding in rice.

MATERIALS AND METHODS

Plant Materials

A set of 148 chromosome segment substitution lines (CSSLs) was derived from the inter-subspecific cross between two genome-sequenced rice cultivars, *japonica* Nipponbare (NIP, as the donor) and *indica* Zhenshan97 (ZS97, as the recurrent parent), using a backcross scheme with a marker-assisted selection approach (Sun et al., 2015; Zhang et al., 2020b). A CSSL that harbors two loci/regions of interest was selected to cross with ZS97 to develop an F_1 hybrid. Deviation of the allele and genotype frequencies from Mendelian expectations at a target locus was assessed in segregating the progeny of the hybrid. Furthermore, from the CSSL-derived progenies, several independent near-isognice lines (NILs), each containing only a single introduced heterozygous segment covering either *TRD4.1* or *TRD4.2* in the otherwise uniform background of ZS97, were developed. Six independent segregating populations were derived from relevant NILs to validate TRD loci. To delimit *TRD4.1* and *TRD4.2*, two large segregated populations were generated to select recombinant individuals per locus. The recombinants with heterozygous target regions were self-crossed to generate progeny for phenotyping TRD. In addition, the reported BIL population from the cross of NIP and 9,311, which were genotyped by the genotyping-by-sequencing method (Yuan et al., 2019), was also used for the analysis of epistatic interaction of particular TRD loci.

For reciprocal cross to test gametic selection, a line (named as NZ) carrying *TRD4.1* at a heterozygous state was first obtained by crossing ZS97 with a NIL that harbors the NIP alleles at *TRD4.1*. Then, two reciprocal populations were generated by crossing NZ as male or female parent with a developed NIL (named as MM) that contains introduced Minghui 63 (MH) alleles at *TRD4.1* (Chen et al., 2018). Both NZ and MM have the same genetic background as ZS97.

Uniformly germinated seeds were planted on 96-well plates with the bottoms removed (Li et al., 2015), and the plates were placed in a growth chamber (Dongnan, Ningbo, China) or a greenhouse at 30°C under 16-h light/8-h dark conditions. The 7-day seedlings were used for genotype analysis. Some lines and segregating populations were grown in the experimental field of Huazhong Agricultural University (HAU) in Wuhan (30.48N, 114.2E), China. Each line was planted in a row with 10 individual spacings of 16.7 cm × 26.6 cm for genotyping. The field management was managed according to the local standard practices.

DNA Extraction and Genotype Analysis

Genomic DNA was extracted from young seedling leaves, as described previously (Zhang et al., 2020b). The genome-wide genotyping of CSSLs was conducted using the RICE 6K array (Sun et al., 2015; Chen et al., 2018). For the genotyping of

segregating populations, a number of polymorphic markers, including simple sequence repeat and insertion/deletion markers, were designed (**Supplementary Table 1**) and used for PCR reaction following the described procedure (Panaud et al., 1996).

Transmission Ratio Distortion Analysis

The allele and genotype frequencies were assayed by several polymorphic markers in a segregated population. The statistical Chi-square test (χ^2) was performed using *chisq.test* function of R[1] to determine whether the observed frequencies are distorted from Mendelian segregation ratios in any segregated population. A genomic region with two or more distorted consecutive markers was considered as one with TRD. Also, the TRD analysis in two subsets of the BIL population was performed using a single-marker analysis model in QTL IciMapping (v4.1), as described previously (Zhang et al., 2020a). A significance level of LOD > 5.0 was set as the threshold to declare the presence of a putative TRD effect in a given bin/marker.

Gametic and Zygotic Selection

To determine the cause of TRD (e.g., gametic or zygotic selection) at a given locus, the allele and genotype (NIP and ZS97) frequencies were assayed to test if they fit the Mendelian ratios, respectively. The comparative patterns of the allele and genotype frequencies may infer gametic or zygotic selection in multiple F_2 populations. If the allele frequency does fit the theoretical ratio of 1:1, and the genotype frequency is biased to 1: 2: 1, the TRD may be raised by zygotic factors. In this regard, an additional test is also performed to determine whether the observed heterozygote frequency deviated from the theoretical genotype frequency (0.5) in the F_2 populations (Fishman and McIntosh, 2019). If the allele frequency is distorted from 1:1, and the genotype frequency fits the ratio of 1: 2: 1, the TRD may result from gametic factors. If there is a distortion of both allele and genotype frequencies from the Mendelian ratios, then the TRD is due to both gametic and zygotic factors.

To track male- and female-specific transmission patterns for a given region using a reciprocal crossing test, diagnostic markers that were polymorphic among the parental lines (NIP, ZS97, and MH) were used for genotyping in the reciprocal cross progeny. The pattern of unequal frequencies through gametic selection is illustrated in **Figure 1**. If the F_1 (ZN) is the male parent (or conversely the female), distortion of genotype (NM and ZM) frequencies in the progeny indicates the gametic selection at the F_1-heterozygous locus.

Pollen Fertility Observation

Pollen fertility was examined with the I_2-KI staining method using mixed pollen grains of more than eight florets from the panicles of each plant, as previously described (Li et al., 2017). A microscope field with at least 200 pollen grains was observed. Pollen fertility was scored as the percentage of filled and stained grains in total grains (Kubo et al., 2011; Li et al., 2017).

[1]https://www.r-project.org/

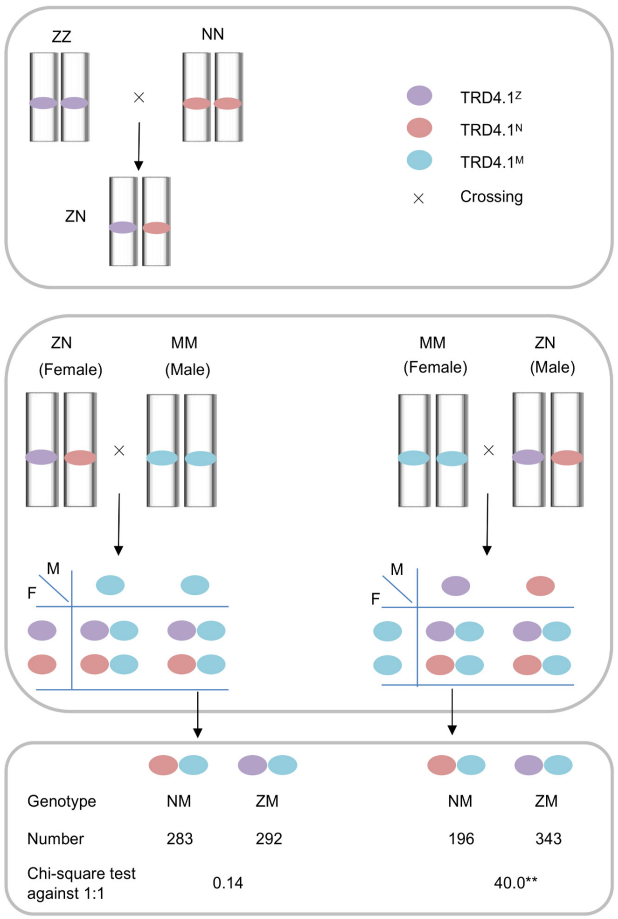

FIGURE 1 | Schematic mode of gametic selection of *TRD4.1* in reciprocal crosses between the substitution lines ZN and MM within the Zhanshan97 background. ZN and MM contain heterozygous and homozygous alleles at *TRD4.1*, respectively. Three alleles N, Z, and M indicate Nipponbare, Zhanshan97, and Minghui63 alleles, respectively.

RESULTS

Detection of Transmission Ratio Distortion Regions on Chromosome 4 Using Chromosome Segment Substitution Lines-Derived Populations

To detect the TRD effect of the introduced NIP segments on chromosome 4, one of the CSSLs that contained the particular segments was crossed with ZS97 to produce a segregating population (name CSSL91-derived population). Graphical genotype analysis showed that CSSL91 carried four introduced NIP segments within the ZS97 background, two on chromosome 4 being targeted, and two on chromosome 2 (**Supplementary Figure 1**). Initially, the population ($n = 191$) was genotyped using four polymorphic markers, each located in the corresponding segment. The allele and genotype frequencies at the two markers (M1 and M2) on chromosome 2 showed normal-Mendelian segregation ratios. However, a non-Mendelian

segregation of the two markers (M3 and M4) on chromosome 4 was observed in both allele and genotype frequencies (**Table 1** and **Supplementary Figure 1**). Furthermore, eight additional consecutive markers (M5 to M12) distributed in the two TRD regions were all distorted significantly from the expected Mendelian ratios in the allele and genotype frequencies. The distortion of these markers was all biased toward ZS97 (**Table 1** and **Supplementary Table 2**), suggesting that the ZS97 alleles at these two regions/loci (named *TRD4.1* and *TRD4.2*) were transmitted to the progeny at a higher frequency than the NIP alleles.

Validation of *Transmission Ratio Distortion 4.1* and *Transmission Ratio Distortion 4.2*

To validate the effect of *TRD4.1*, two lines that carried heterozygous *TRD4.1* along with *TRD4.2* homozygous ZS97 or NIP were obtained and self-crossed to produce corresponding segregating populations. As a cluster of molecular markers showing TRD suggests that the chromosomal region may have one or more genes causing TRD, the representative marker (M7) was used to analyze the TRD effect at TRD4.1. A non-Mendelian segregation of *TRD4.1* was observed in these two populations (**Figure 2** and **Table 2**). Moreover, the ZS97 alleles at *TRD4.1* were preferentially transmitted to the progeny in heterozygotes.

To check the effect of *TRD4.2*, two lines that harbored heterozygous *TRD4.2* and homozygous at *TRD4.1* either with ZS97 or NIP alleles were self-pollinated to produce segregating populations (**Figure 2** and **Table 2**). Consistently, *TRD4.2* assayed by the representative marker M12 exhibited a significant TRD effect with the preferential transmission of the ZS97 alleles in the progeny of heterozygotes harboring homozygous $TRD4.1^{NIP}$. In contrast, normal segregation of *TRD4.2* was observed in the progeny of the heterozygotes when carried with homozygous $TRD4.1^{ZS97}$ (**Table 2**).

An additional segregated population (n = 902) was also generated from the heterozygotes at both *TRD4.1* and *TRD4.2* (**Figure 2**). Frequencies of the nine genotypes assayed by two markers (M7 and M12) linked with *TRD4.1* and *TRD4.2* did not fit the Mendelian segregation ratio (χ^2 = 185.5; $P < 2.2 \times 10^{-16}$), given the two loci are linked (**Supplementary Table 2**). In particular, the number of genotype homozygous $TRD4.1^{ZS97}TRD4.2^{ZS97}$ was more than that of $TRD4.1^{NIP}TRD4.2^{NIP}$, $TRD4.1^{ZS97}TRD4.2^{NIP}$, and $TRD4.1^{NIP}TRD4.2^{ZS97}$ in the population. Collective data confirmed that TRD arose from both *TRD4.1* and *TRD4.2* and a preferential transmission of the ZS97 gametes at the two loci.

Detection of Other Transmission Ratio Distortion Loci Conditions on *Transmission Ratio Distortion 4.1*

To determine whether *TRD4.1* affects other genomic regions on TRD, the BIL population that was previously developed from the cross of NIP and 9,311 (Zhang et al., 2020b) was divided into two subpopulations according to the *TRD4.1* genotype: SubN (n = 75) in which all lines had $TRD4.1^{NIP/NIP}$ and SubJ (n = 325) that had $TRD4.1^{9311/9311}$ assayed by 10 consecutive bin-markers (from 13.47 to 15.87 Mb) at the TRD4.1 region (**Supplementary Table 3**). Four and ten regions were detected in SubN and SubJ, respectively. Among them, three TRD regions (*TR1.3*, *TR8.2*, and *TR12.1*) were common in both subpopulations. The other seven regions were identified only in SubJ. In particular, the *TRD4.2* region was detected only in SubN but not in SubJ (**Supplementary Table 3**). These results revealed the effect of *TRD4.1* on the detection of other TRD loci in the BIL population.

Gametic Selection Leading to Transmission Ratio Distortion

To investigate whether *TRD4.1* and *TRD4.2* are involved in the gametic or zygotic factors, the population segregated at both TRD4.1 and TRD4.2 was analyzed. The results showed that the frequencies of heterozygous genotypes at either *TRD4.1* or *TRD4.2* were approximately 0.5. The observed frequencies of heterozygotes at either *TRD4.1* or *TRD4.2* exhibited no significant derivation from the expected

TABLE 1 | Identification of *TRD4.1* and *TRD4.2* in the CSSL-derived segregating populations.

Loci	Marker	Position (Mb)	Genotype frequency					Allele frequency				Toward[a]
			P1	H	P2	χ^2 (df = 2)	P-value	ZS97	NIP	χ^2 (df = 1)	P-value	
TRD4.1	M5	6.46	71	95	20	28.05	8.1E-07	0.64	0.36	27.97	1.2E-07	ZS97
	M6	8.63	79	89	22	34.96	2.6E-08	0.65	0.35	34.20	5.0E-09	
	M3	11.65	73	95	22	27.37	1.1E-06	0.63	0.37	27.38	1.7E-07	
	M7	14.06	77	94	20	34.07	4.0E-08	0.65	0.35	34.02	5.5E-09	
	M8	14.65	73	97	19	30.90	1.9E-07	0.64	0.36	30.86	2.8E-08	
TRD4.2	M4	19.95	67	93	25	19.07	7.2E-05	0.61	0.39	19.07	1.3E-05	ZS97
	M9	20.09	74	94	23	27.30	1.2E-06	0.63	0.37	27.24	1.8E-07	
	M10	20.17	67	98	18	27.16	1.3E-06	0.63	0.37	26.24	3.0E-07	
	M11	20.20	70	85	26	22.10	1.6E-05	0.62	0.38	21.39	3.7E-06	
	M12	21.63	76	88	27	26.32	1.9E-06	0.63	0.37	25.14	5.3E-07	

[a] *The allele is preferentially transmitted in heterozygotes. P1, P2, and H indicate homozygous ZS97, Nipponbare (NIP), and heterozygous genotypes, respectively.*

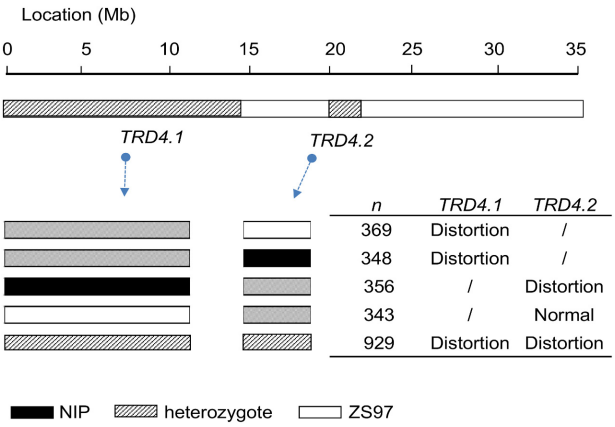

FIGURE 2 | Detection of single and combined effects of *TRD4.1* and *TRD4.2* on transmission ratio distortion using five segregating populations (comprising *n* individuals), which were derived from corresponding chromosome segment substitution line that contains the particular heterozygous region (s), covering either *TRD4.1* or *TRD4.2* or both. The graphic genotype of the two loci in the relevant line is indicated.

frequencies (**Supplementary Table 2**). This non-distortion of heterozygote frequency may indicate a gametophytic barrier. In addition, normal pollen fertility (greater than 95%) of the nine genotypes revealed that TRD was not caused by pollen fertility (**Supplementary Table 4**). These results confirm that the gametic factors are involved in *TRD4.1* and *TRD4.2*.

To investigate whether the *TRD4.1* effect is caused by female or male gametic factors, a reciprocal test cross was made between NILs with a heterozygous (ZN) segment and a MM segment that contains homozygous Minghui 63 alleles at the locus (**Figure 1**). Segregation of the locus was assessed in a F_1 progeny from the crosses. The progenies of the reciprocal crosses (NZ × MM and MM × ZN) were examined for TRD effects. The polymorphic markers M13 and M14 that were tightly linked with *TRD4.1* and which could distinguish the three genotypes (NIP, ZS97, and MH63) were used to classify two progeny genotypes (NM and ZM) of the crosses. The genotype frequency of *TRD4.1* showed normal-Mendelian segregation (no TRD) by the Chi-square test against the expected ratio of 1:1 (NM:ZM = 283:292; $\chi^2 = 0.14$; $P > 0.70$) in the cross of ZN × MM ($n = 575$), in which ZN was used as the female parent. The results suggest that NIP and ZS97 alleles were equally transmitted to the progeny. However, significant distortion of genotype frequencies (NM:ZM = 196:343; $\chi^2 = 40.09$; $P < 2.5 \times 10^{-10}$) was observed in the progeny of the cross of MM × NZ ($n = 539$), in which ZN was used as the male parent (**Figure 1**). In addition, normal pollen fertility (greater than 95%) of ZN was observed. These results indicate that the ZS97 allele was preferentially transmitted to the progeny over the NIP allele. Therefore, the distortion of genotype frequencies at *TRD4.1* in the progeny involved male gamete selection.

Fine Mapping of *Transmission Ratio Distortion 4.1* and *Transmission Ratio Distortion 4.2*

To narrow the region of *TRD4.1*, a mapping population comprising of approximately 6,500 individuals was developed from one plant that harbored only one heterozygous *TRD4.1* region and one homozygous $TRD4.2^{NIP}$ within the ZS97 background. Initially, five recombinant individuals (R1 to R5) were obtained using several markers linked with *TRD4.1*. Then, the recombinants were self-crossed to generate five independent segregating populations for further genotyping (**Figure 3**). Based on frequencies of the genotypes in each recombinant-derived population using representative markers (M7 or M8), the distorted region covering *TRD4.1* was determined in the corresponding recombinant. *TRD4.1* was preliminarily mapped to an approximately 590-kb region between markers M7 and M8. For fine mapping of *TRD4.1*, eight recombinant individuals (R6 to R13) within the interval of M7-M8 were selected and generated eight independent recombinant-derived populations. TRD analysis based on frequency genotypes in recombinant-derived populations delimited *TRD4.1* into a 100-kb region between markers M19 and M8.

Based on the same approach, fine-mapping of *TRD4.2* was conducted on a segregated population (composed of approximately 5,800 individuals) derived from a plant that carried only one heterozygous *TRD4.2* region and one homozygous $TRD4.1^{NIP}$ segment in the ZS97 background. Seventeen recombinant individuals in the *TRD4.2* region flanked by markers M23 and M12 were obtained (**Figures 4A,B**). Based on TRD analysis of genotype frequencies in each of the 17 recombination-derived populations and genotyping with additional markers in the target region, *TRD4.2* was

TABLE 2 | Validation of *TRD4.1* and *TRD4.2* in four segregating populations.

Parental genotype	Representative Marker	Observed genotype frequency[a]						Observed allele frequency[b]				Toward[c]
		ZZ	H	NN	Sum	χ^2 (*df* = 2)	*P*-value	ZS97	NIP	χ^2 (*df* = 1)	*P*-value	
$TRD4.1^{het}/TRD4.2^{ZS}$	M7	128	178	63	369	23.4	8.5E-06	0.59	0.41	22.9	1.7E-06	ZS97
$TRD4.1^{het}/TRD4.2^{NIP}$	M7	122	174	52	348	28.2	7.7E-07	0.60	0.40	28.2	1.1E-07	ZS97
$TRD4.1^{NIP}/TRD4.2^{het}$	M12	136	181	39	356	53.0	3.2E-12	0.64	0.36	52.9	3.6E-13	ZS97
$TRD4.1^{ZS}/TRD4.2^{het}$	M12	81	184	78	343	1.9	3.9E-01	0.50	0.50	0.1	8.2E-01	Equal

[a,b]Genotype and allele frequencies assayed by representative markers tightly linked with TRD4.1 or TRD4.2, respectively. ZZ, NN, and H indicate the homozygous ZS97 (ZS), Nipponbare (NIP), and heterozygous genotype (het), respectively. [c]Toward means preferential allele transmission from heterozygotes. "Equal" indicates that both alleles are transmitted equally to the progeny.

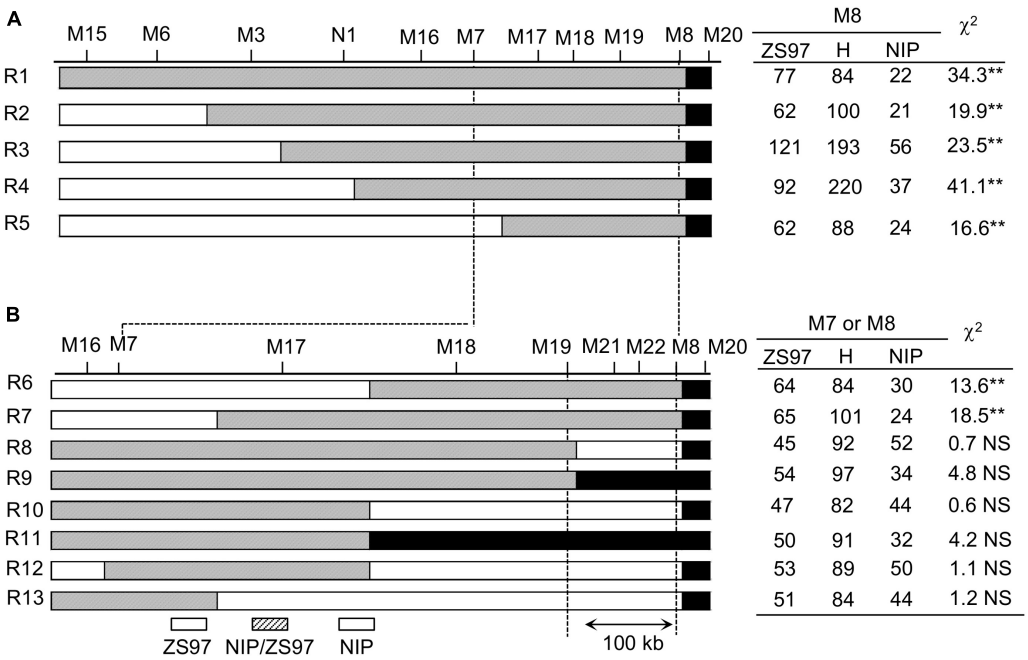

FIGURE 3 | Fine-mapping of TRD4.1. R1 to R13 represents the recombinant individuals between M15 and M8. Their derived segregating populations were used to validate TRD of TRD4.1. **(A)** TRD analysis of five (R1 to R5) recombinant-derived segregating populations delimited TRD4.1 into a 590-kb interval. The polymorphic marker of M8 was used to investigate the genotypes of R1 to R5-derived populations. **(B)** Finely mapping of TRD4.1 to a 100-kb region between M8 and M19 using eight (R6 to R13) recombinant-derived populations. The marker M8 was used to investigate the genotypes of R6- and R7-derived populations. The marker M7 was used to investigate the genotypes of R8 to R13-derived populations. ZS97, H, and NIP denote ZS97, heterozygous, and NIP genotypes at TRD4.1, respectively. "**" Denotes significant distortion of the allele and genotype frequencies by chi-square test at $P < 0.01$. NS, no significance.

narrowed down to an approximately 34.1-kb region. This region contains nine predicted genes based on the reference genome[2] (**Figure 4C**). Of them, LOC_Os04g33150 was the only one expressed gene, after removing those annotated as unknown, transposons/retrotransposons, or hypothetical proteins. It was specifically and highly expressed in the pre-emergence inflorescence, 5-d seed, and 25-d endosperm.[3] This gene encodes a desiccation-related protein. Sequence differences that existed in the coding region between NIP and ZS97 may cause two amino acid changes.[4] Therefore, LOC_Os04g33150 is the most likely candidate gene for *TRD4.2*.

DISCUSSION

The present study identified two linked loci *TRD4.1* and *TRD4.2*, and their epistatic interaction caused transmission ratio distortion in several CSSL-derived segregating populations, in which only a single or two NIP segments were targeted within the ZS97 background in rice. Both loci showed a preferential transmission of the ZS97 alleles to the progeny in heterozygotes. This severe TRD with a bias toward the *indica* allele is consistent with previous results in the NIP/9311 BIL population and other subspecific crosses (Wang et al., 2009; Reflinur et al., 2014; Zhang et al., 2020a).

One of the findings is that the TRD effect on *TRD4.2* was dependent on *TRD4.1* (**Figure 1**). *TRD4.2* showed severe TRD in the CSSL-derived populations in which the progeny harbored homozygous $TRD4.1^{NIP}$, but normal segregation in the progeny when it carried with homozygous $TRD4.1^{ZS97}$ (**Table 2**). Furthermore, *TRD4.2* was detected only in SubN that had a fixed NIP genotype at the *TRD4.1* region in the NIP/9311 BIL population (**Supplementary Table 3**). However, *TRD4.1* displayed several allelic transmission biases no matter whether *TRD4.2* was homozygous or heterozygous NIP alleles (**Figure 2**). In addition, *TRD4.1* from the NIP decreased the incidence of TRD (63%) in a subset (SubN) of the BIL population. These results suggest that epistatic interaction plays an important determinant in the segregation patterns of TRD loci. This is the possible reason that some loci like *TRD4.2* have largely gone unnoticed in any previous study on rice.

Another finding in this study is that *TRD4.1* is possible only through male gametic selection among the progeny. TRD is a selection mechanism that could be caused by gametic and zygotic factors. In the present studies, using a series of CSSL- and NIL-derived populations, we found that the allele frequencies at *TRD4.1* and *TRD4.2* in the populations were significantly skewed toward ZS97 (**Tables 1**, **2**); however, the heterozygous genotypes showed fertile pollen and normal segregation in the progeny (**Supplementary Table 4**). These results indicate that gametic

[2]http://rice.plantbiology.msu.edu/

[3]http://rice.plantbiology.msu.edu/expression.shtml

[4]http://rice.hzau.edu.cn/rice_rs3/

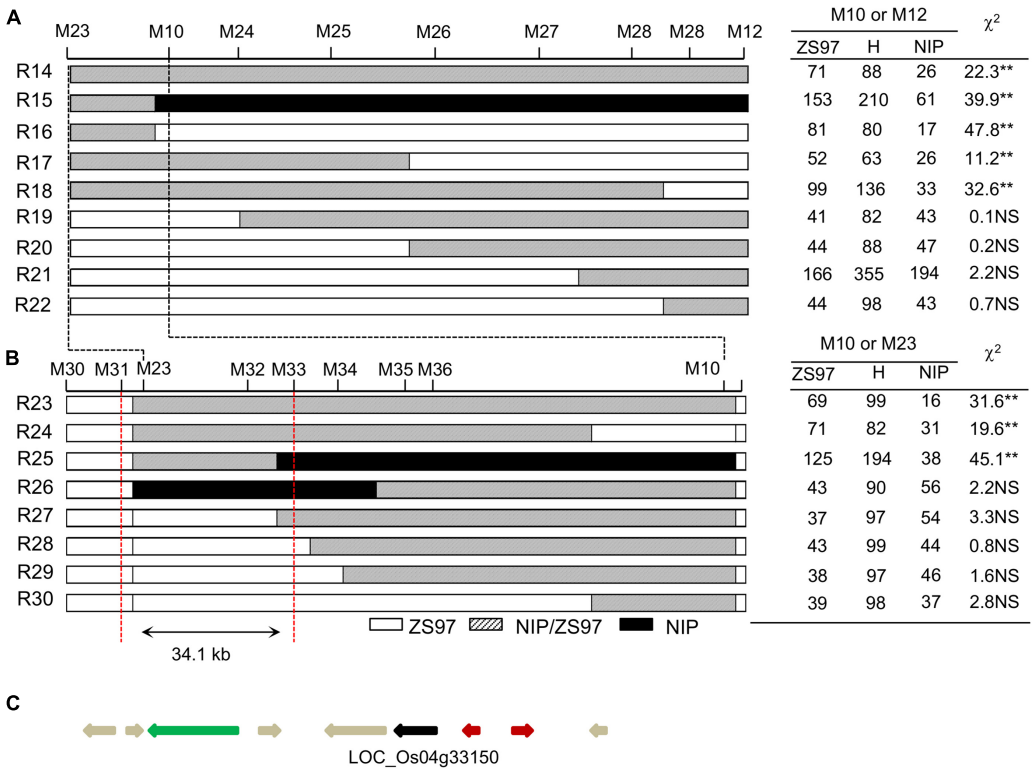

FIGURE 4 | Fine-mapping and candidate gene analysis of *TRD4.2*. R14 to R30 represent the recombinant individuals between M23 and M12. **(A)** TRD analysis of recombinant-derived segregating populations delimited *TRD4.2* into a 600-kb interval. **(B)** Fine-mapping of *TRD4.2* to a 34.1-kb region between M31 and M34. ZS97, H, and NIP denote ZS97, heterozygous, and NIP genotypes at *TRD4.2*, respectively. "**" Denote significant distortion of the allele and genotype frequencies according to chi-square test at $P < 0.01$. NS, no significant distortion. **(C)** Annotated genes for *TRD4.2* based on the Nipponbare reference genome (http://rice.plantbiology.msu.edu/index.shtml), showing a candidate gene that is indicated by the dark arrow. The red, gray, and green arrows represent hypothetical proteins, expression protein, and transposon or retrotransposon, respectively.

selection was involved in *TRD4.1* and *TRD4.2*. Notably, we used a three-allele reciprocal crossing test to investigate whether there was a female or male gametic selection in *TRD4.1* and found that the severe transmission bias of *TRD4.1* was due to the male gametic selection of the NIP alleles. It is notable that *TRD4.1* was mapped in the common region where *SD4.1* was reported to affect segregation distortion and spikelet fertility in both inter- and intra-specific populations (Wan et al., 1996; Li et al., 2017; Zhang et al., 2020a). We further delimited *TRD4.1* into a 100-kb interval and *TRD4.2* into a 34.1-kb region with one candidate gene (**Figures 3**, **4**), which would facilitate cloning the genes underlying the gametic factors for TRD. Further characterization and functional analysis of *TRD4.1* will be required to better understand the male gametic selection through gamete killer, gamete competition, and/or differential fertilization success.

CONCLUSION

Two TRD regions (*TRD4.1* and *TRD4.2*) on chromosome 4 were identified and validated using CSSL-derived secondary populations. A significant digenic interaction between *TRD4.1* and *TRD4.2* affected TRD. Of them, *TRD4.2*-mediated TRD was dependent on the presence of *TRD4.1* alleles, but *TRD4.1* was not affected by *TRD4.2*. Moreover, *TRD4.1* and *TRD4.2* were delimited to approximately 100-kb and 34.1-kb intervals, respectively. Furthermore, we found that *TRD4.1* is male gametic in action with the preferential transmission of the *indica* ZS97 allele to the progeny. These findings would be helpful for cloning candidate genes and characterizing the molecular mechanisms underlying TRD.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

SY designed and conceived the research. CZ, JW, XX, DW, and WS developed the populations. CZ, JW, XZ, XX, and ZY conducted the experiments. CZ and JW analyzed the data. CZ, JW, and SY wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This research was funded by grants from the National Natural Science Foundation of China (31971864), the Hubei Special Major Projects for Technological Innovation (2020ABA016), and the Earmarked Fund for China Agricultural Research System (CARS-01-06).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.866276/full#supplementary-material

Supplementary Figure 1 | Primarily mapping TRD regions on chromosome 4 using CSSL-derived populations. **(A)** Graphical genotype of CSSL91 showing four introduced Nipponbare segments encompassing TRD4.1 and TRD4.2 in the ZS97 background. **(B)** The allele and genotype frequencies at target regions (markers) showing non-Mendelian segregation by Chi-square test. ZZ, ZN, and NN represent ZS97, heterozygote, and NIP genotype, respectively.

Supplementary Table 1 | Primers used in this study.

Supplementary Table 2 | Nine genotypic frequencies at the loci TRD4.1 and TRD4.2 in a segregated population.

Supplementary Table 3 | The genomic regions detected for transmission ratio distortion in two subpopulations fixed TRD4.1Nip or TRD4.19311.

Supplementary Table 4 | Pollen fertility percentage of nine genotypes at TRD4.1 and TRD4.2.

REFERENCES

Baack, E., Melo, M. C., Rieseberg, L. H., and Ortiz-Barrientos, D. (2015). The origins of reproductive isolation in plants. *New Phytol.* 207, 968–984. doi: 10.1111/nph.13424

Bauer, H., Véron, N., Willert, J., and Herrmann, B. G. (2007). The *t*-complex-encoded guanine nucleotide exchange factor *Fgd2* reveals that two opposing signaling pathways promote transmission ratio distortion in the mouse. *Genes Dev.* 21, 143–147. doi: 10.1101/gad.414807

Casellas, J., Gularte, R. J., Farber, C. R., Varona, L., Mehrabian, M., Schadt, E. E., et al. (2012). Genome scans for transmission ratio distortion regions in mice. *Genetics* 191, 247–259. doi: 10.1534/genetics.111.135988

Chandnani, R., Wang, B., Draye, X., Rainville, L. K., Auckland, S., Zhuang, Z. M., et al. (2017). Segregation distortion and genome-wide digenic interactions affect transmission of introgressed chromatin from wild cotton species. *Theor. Appl. Genet.* 130, 2219–2230. doi: 10.1007/s00122-017-2952-y

Chen, J., Wang, J. J., Chen, W., Sun, W. Q., Peng, M., Yuan, Z. Y., et al. (2018). Metabolome analysis of multi-connected biparental chromosome segment substitution line populations. *Plant Physiol.* 178, 612–625. doi: 10.1104/pp.18.00490

Corbett-Detig, R. B., Zhou, J., Clark, A. G., Hartl, D. L., and Hartl, J. F. (2013). Genetic incompatibilities are widespread within species. *Nature* 504, 135–137. doi: 10.1038/nature12678

Dai, B. S., Guo, H. L., Huang, C., Ahmed, M. M., and Lin, Z. X. (2017). Identification and characterization of segregation distortion loci on cotton chromosome 18. *Front. Plant Sci.* 7:2037. doi: 10.3389/fpls.2016.02037

Fishman, L., Aagaard, J., and Tuthill, J. C. (2008). Toward the evolutionary genomics of gametophytic divergence: patterns of transmission ratio distortion in monkeyflower (*Mimulus*) hybrids reveal a complex genetic basis for conspecific pollen precedence. *Evolution* 62, 2958–2970. doi: 10.1111/j.1558-5646.2008.00475.x

Fishman, L., and McIntosh, M. (2019). Standard deviations: the biological bases of transmission ratio distortion. *Annu. Rev. Genet.* 53, 347–372. doi: 10.1146/annurev-genet-112618-043905

Giesbers, A. K. J., Boer, E. D., Ulen, J. J. W. E. H., Kaauwen, M. P. W., Visser, R. G. F., Niks, R. E., et al. (2019). Patterns of transmission ratio distortion in interspecific lettuce hybrids reveal a sex-independent gametophytic barrier. *Genetics* 211, 263–276. doi: 10.1534/genetics.118.301566

Huang, L. O., Labbe, A., and Infante-Rivard, C. (2013). Transmission ratio distortion: review of concept and implications for genetic association studies. *Hum. Genet.* 132, 245–263. doi: 10.1007/s00439-012-1257-0

Jenczewski, E., Gherardi, M., Bonnin, I., Prosperi, J. M., Olivieri, I., and Huguet, T. (1997). Insight on segregation distortions in two intraspecific crosses between annual species of *Medicago* (Leguminosae). *Theor. Appl. Genet.* 94, 682–691. doi: 10.1007/s001220050466

Koide, K., Shinya, Y., Ikenaga, M., Sawamura, N., Matsubara, K., Onishi, K., et al. (2012). Complex genetic nature of sex-independent transmission ratio distortion in Asian rice species: the involvement of unlinked modifiers and sex-specific mechanisms. *Heredity* 108, 242–247. doi: 10.1038/hdy.2011.64

Koide, Y., Onishi, K., Nishimoto, D., Baruah, A. R., Kanazawa, A., and Sano, Y. (2008a). Sex-independent transmission ratio distortion system responsible for reproductive barriers between Asian and African rice species. *New Phytol.* 179, 888–900. doi: 10.1111/j.1469-8137.2008.02490.x

Koide, Y., Ikenaga, M., Sawamura, N., Nishimoto, D., Matsubara, K., Onishi, K., et al. (2008b). The evolution of sex-independent transmission ratio distortion involving multiple allelic interactions at a single locus in rice. *Genetics* 180, 409–420. doi: 10.1534/genetics.108.090126

Kubo, T., Yoshimura, A., and Kurata, N. (2011). Hybrid male sterility in rice is due to epistatic interactions with a pollen killer locus. *Genetics* 189, 1083–1092. doi: 10.1534/genetics.111.132035

Kumar, S., Gill, B. S., and Faris, J. D. (2007). Identification and characterization of segregation distortion loci along chromosome 5B in tetraploid wheat. *Mol. Genet. Genomics* 278, 187–196. doi: 10.1007/s00438-007-0248-7

Leppälä, J., Bokma, F., and Savolainen, O. (2013). Investigating incipient speciation in *Arabidopsis lyrata* from patterns of transmission ratio distortion. *Genetics* 194, 697–708. doi: 10.1534/genetics.113.152561

Li, G. W., Jin, J. Y., Zhou, Y., Bai, X. F., Mao, D. H., Tan, C., et al. (2019). Genome wide dissection of segregation distortion using multiple inter-subspecific crosses in rice. *Sci. China Life. Sci.* 62, 507–516. doi: 10.1007/s11427-018-9452-8

Li, G. W., Li, X. T., Wang, Y., Mi, J. M., Xing, F., Zhang, D. H., et al. (2017). Three representative inter and intra-subspecific crosses reveal the genetic architecture of reproductive isolation in rice. *Plant J.* 92, 349–362. doi: 10.1111/tpj.13661

Li, X. M., Chao, D. Y., Wu, Y., Huang, X. H., Chen, K., Cui, L. G., et al. (2015). Natural alleles of a proteasome α2 subunit gene contribute to thermotolerance and adaptation of African rice. *Nat. Gent.* 47, 827–833. doi: 10.1038/ng.3305

Mizuta, Y., Harushima, Y., and Kurata, N. (2010). Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20417–20422. doi: 10.1073/pnas.1003124107

Nguyen, G. N., Yamagata, Y., Shigematsu, Y., Watanabe, M., Miyazaki, Y., Doi, K., et al. (2017). Duplication and loss of function of genes encoding RNA polymerase III subunit C4 causes hybrid incompatibility in rice. *G3* 7, 2565–2575. doi: 10.1534/g3.117.043943

Ouyang, Y. D., and Zhang, Q. F. (2013). Understanding reproductive isolation based on the rice model. *Annu. Rev. Plant Biol.* 64, 111–135. doi: 10.1146/annurev-arplant-050312-120205

Panaud, O., Chen, X., and McCouch, S. R. (1996). Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* 252, 597–607. doi: 10.1007/BF02172406

Reflinur, Kim, B., Jang, S. M., Chu, S., Bordiya, Y., Akter, M. B., et al. (2014). Analysis of segregation distortion and its relationship to hybrid barriers in rice. *Rice* 7:3. doi: 10.1186/s12284-014-0003-8

Seymour, D. K., Chae, E., Arioz, B. I., Koenig, D., and Weigel, D. (2019). Transmission ratio distortion is frequent in *Arabidopsis thaliana* controlled crosses. *Heredity* 122, 294–304. doi: 10.1038/s41437-018-0107-9

Sun, W. Q., Zhou, Q. L., Yao, Y., Qiu, X. J., Xie, K., and Yu, S. (2015). Identification of genomic regions and the isoamylase gene for reduced grain chalkiness in rice. *PLoS One* 10:e0122013. doi: 10.1371/journal.pone.0122013

Vogl, C., and Xu, S. (2000). Multipoint mapping of viability and segregation distorting loci using molecular markers. *Genetics* 155, 1439–1447. doi: 10.1093/genetics/155.3.1439

Wan, J., Yamaguchi, Y., Kato, H., and Ikehashi, H. (1996). Two new loci for hybrid sterility in cultivated rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 92, 183–190. doi: 10.1007/BF00223375

Wang, S. H., Tan, X. L., Zhang, Z. L., Wen, J. C., and Kou, S. Y. (2009). Segregation distortion detected in six rice F_2 populations generated from reciprocal hybrids at three altitudes. *Genome Res.* 91, 345–353. doi: 10.1017/S0016672309990176

Xie, Y., Xu, P., Huang, J., Ma, S., Xie, X., Tao, D., et al. (2017). Interspecific hybrid sterility in rice is mediated by *OgTPR1* at the *S1* locus encoding a peptidase-like protein. *Mol. Plant* 10, 1137–1140. doi: 10.1016/j.molp.2017.05.005

Xie, Y. Y., Tang, J. T., Xie, X. R., Li, X. J., Huang, J. L., Fei, Y., et al. (2019). An asymmetric allelic interaction drives allele transmission bias in interspecific rice hybrids. *Nat. Commun.* 10:2501. doi: 10.1038/s41467-019-10488-3

Xu, X., Li, L., Dong, X., Jin, W., Melchinger, A. E., and Chen, S. (2013). Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize. *J. Exp. Bot.* 64, 1083–1096. doi: 10.1093/jxb/ers393

Yamagata, Y., Yamamoto, E., Aya, K., Win, K. T., Doi, K., Sobrizal, et al. (2010). Mitochondrial gene in the nuclear genome inducesreproductive barrier in rice. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1494–1499. doi: 10.2307/40536356

Yang, J. Y., Zhao, X. B., Cheng, K., Du, H. Y., Ouyang, Y. D., Chen, J. J., et al. (2012). A killer-protector system regulates both hybrid sterility and segregation distortion in rice. *Science* 337, 1336–1340. doi: 10.1126/science.1223702

Yu, X. W., Zhang, Z. G., Zheng, X. M., Zhou, J. W., Kong, W. Y., Wang, P. R., et al. (2018). A selfish genetic element confers non-Mendelian inheritance in rice. *Science* 360, 1130–1132. doi: 10.1126/science.aar4279

Yuan, Z. Y., Fan, K., Xia, L. F., Ding, X. L., Tian, L., Sun, W. Q., et al. (2019). Genetic dissection of seed storability and validation of candidate gene associated with antioxidant capability in rice (*Oryza sativa* L.). *Int. J. Mol. Sci.* 20:4442. doi: 10.3390/ijms20184442

Zhang, C. P., Wang, D. W., Wang, J. L., Sun, Q., Tian, L., Tang, X. X., et al. (2020a). Genetic dissection and validation of chromosomal regions for transmission ratio distortion in intersubspecific crosses of rice. *Front. Plant Sci.* 11:563548. doi: 10.3389/fpls.2020.563548

Zhang, C. P., Yuan, Z. Y., Wang, Y. T., Sun, W. Q., Tang, X. X., Sun, Y. J., et al. (2020b). Genetic dissection of seed dormancy in rice (*Oryza sativa* L.) by using two mapping populations derived from common parents. *Rice* 13:52. doi: 10.1186/s12284-020-00413-4

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

ORIGINAL RESEARCH
published: 05 April 2022
doi: 10.3389/fpls.2022.866165

A Genetic Network Underlying Rhizome Development in *Oryza longistaminata*

Wenfei Li†, Shilai Zhang†, Guangfu Huang, Liyu Huang, Jing Zhang, Zheng Li and Fengyi Hu**

State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Research Center for Perennial Rice Engineering and Technology of Yunnan, School of Ecology and Environmental Science, School of Agriculture, Yunnan University, Kunming, China

The rhizome is an important organ through which many perennial plants are able to propagate vegetatively. Its ecological role has been thoroughly studied on many grass species while the underlying genetic basis is mainly investigated using a rhizomatous wild rice species—*Oryza longistaminata*. Previous studies have revealed that the rhizome trait in *O. longistaminata* is jointly controlled by multiple loci, yet how these loci interact with each other remains elusive. Here, an F_2 population derived from *Oryza sativa* (RD23) and *O. longistaminata* was used to map loci that affect rhizome-related traits. We identified 13 major-effect loci that may jointly control rhizomatousness in *O. longistaminata* and a total of 51 quantitative trait loci (QTLs) were identified to affect rhizome abundance. Notably, some of these loci were found to have effects on more than one rhizome-related trait. For each trait, a genetic network was constructed according to the genetic expectations of the identified loci. Furthermore, to gain an overview of the genetic regulation on rhizome development, a comprehensive network integrating all these individual networks was assembled. This network consists of three subnetworks that control different aspects of rhizome expression. Judging from the nodes' role in the network and their corresponding traits, we speculated that *qRHZ-3-1*, *qRHZ-4*, *qRHI-2*, and *qRHI-5* are the key loci for rhizome development. Functional verification using rhizome-free recombinant inbred lines (RILs) suggested that *qRHI-2* and *qRHI-5*, two multi-trait controlling loci that appeared to be critical in our network analyses, are likely both needed for rhizome formation. Our results provide more insights into the genetic basis of rhizome development and may facilitate identification of key rhizome-related genes.

Keywords: ***Oryza longistaminata*, rhizomes, development, genetic networks, quantitative trait loci**

OPEN ACCESS

Edited by:
Kenneth Lee McNally, International Rice Research Institute (IRRI), Philippines

Reviewed by:
Binying Fu, Institute of Crop Sciences (CAAS), China
Jiangyi Yang, Guangxi University, China

****Correspondence:***
Zheng Li
zhengli@ynu.edu.cn
Fengyi Hu
hfengyi@ynu.edu.cn

† These authors have contributed equally to this work

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

Received: *30 January 2022*
Accepted: *09 March 2022*
Published: *05 April 2022*

Citation:
Li W, Zhang S, Huang G, Huang L, Zhang J, Li Z and Hu F (2022) A Genetic Network Underlying Rhizome Development in Oryza longistaminata. Front. Plant Sci. 13:866165. doi: 10.3389/fpls.2022.866165

INTRODUCTION

Rhizomes are modified subterranean stems that grow horizontally and can produce roots and shoots on their nodes, enabling vigorous asexual proliferation in various perennial species (Gizmawy et al., 1985; Li et al., 2022). From a physiological point of view, the rhizome is also the main energy storage organ of many perennial plants and plays a determinate role in their survival

in harsh environment (Paterson et al., 1995). In agriculture, the rhizome trait is of significance due to both its positive and negative effects. On the one hand, strong rhizomes largely contribute to the competitiveness and invasiveness of weeds (Jang et al., 2006). On the other hand, rhizomes are also dispersal in many forage crops and may be utilized for developing perennial grain crops (Hu et al., 2011). Notably, it has been speculated that genes controlling rhizome development could be used to prolong the nutritional life cycle of certain plant species, potentially converting major annual grain crops into perennial ones (Paterson et al., 1995; Hu et al., 2003; Cox et al., 2006; Glover et al., 2010). As such, understanding the genetic mechanisms underlying rhizome development is not only instrumental in devising control or productivity-enhancing strategies for rhizomatous plants, but also beneficial to sustainable food production and ecosystem maintenance given the multiple ecological threats posed by our current annual crop dominated agricultural system (Hu et al., 2003; Cox et al., 2006; Glover et al., 2010).

Previous efforts to dissect the genetic basis of the rhizome trait were largely hindered by the sparsity of genomic resources. Recently, *Oryza longistaminata* is emerging as a useful model system for exploring rhizome development (Hu et al., 2003; Li et al., 2022). Originated from Africa, *O. longistaminata* is the only rhizomatous wild *Oryza* species that has the same AA genome type as the cultivated rice species *O. sativa* (Morishima, 1967; Tao and Sripichitt, 2000; Hu et al., 2003; Sacks et al., 2003; Ghesquiere, 2008). Hence, it can be crossed with well-studied rice cultivars for which high-quality genomic information is available, a strategy that is widely used for mapping genes or quantitative trait loci (QTLs) for agronomic traits. It is noteworthy that through such strategy, known as wide hybridization, a number of perennial rice lines have been successfully developed and are currently on trial in many Chinese provinces, representing an environmentally and economically sound rice production system (Zhang et al., 2017, 2019; Huang et al., 2018; Zhang et al., 2021).

An early study (Meakawa et al., 1998) that took advantage of the hybridization between *O. sativa* and *O. longistaminata* has suggested that the rhizomatous growth habit of *O. longistaminata* was segregated as a single dominant trait. The single dominant allele responsible for the rhizome trait was termed *Rhz*, and it was loosely linked to the *liguleless* (*lg*) locus on chromosome 4 with a recombination value of 37 ± 3.6% (Meakawa et al., 1998). However, the rhizomatous phenotypes displayed pronounced variations in the F_2 population, indicating the presence of other modifying genes. Based on a complete simple sequence-repeat map, Hu et al. (2003) identified two dominant-complementary loci, termed *Rhz2* and *Rhz3*, that predominantly control rhizomatousness in *O. longistaminata*, and the authors also revealed many QTLs affecting rhizome abundance (Hu et al., 2003). Recently, through entire population genotyping mapping and selective genotyping mapping using three F_2 populations, over 10 major- or minor-effect rhizome-regulating QTLs were identified; however, none of these QTLs could be able to function alone, indicating that interactions among multiple QTLs are required for proper rhizome development (Fan et al., 2020). Nevertheless, to our knowledge, no reports have detailed how the rhizome-related QTLs interact with each other.

Understanding the relationships between complex genotypes and their underlying phenotypes is still one of the main challenges in modern genetics (Mackay et al., 2009). To better characterize the genetic networks underlying complex traits, a theoretical framework founded upon knowledge of signal transduction pathways has been proposed (Zhang et al., 2011). In this framework, the principle of hierarchy is defined as one-way functional dependency of downstream genes on upstream regulators, and functional genetic units (FGUs) refer to a group of functionally dependent genes acting at each level of a signaling pathway (Zhang et al., 2011). This framework was later employed to explore the pleiotropic effects of *SD1*, whose mutant alleles greatly contribute to the Green Revolution, and three genetic systems (*SD1*-mediated, -repressed, and -independent) were revealed, comprising 43, 38, and 64 FGUs, respectively, and jointly controlling growth, development and productivity of rice (Zhang et al., 2013). The functionality of this framework was further demonstrated in an investigation of the genetic basis of submergence tolerance (Wang et al., 2015). The putative networks consisted of 296 loci that were grouped into 167 FGUs, and the directional links between and among the nodes (the detected loci) suggested that submergence tolerance in rice is genetically controlled by a number of positively regulated signaling pathways (Wang et al., 2015).

Here, to gain insights into the genetic basis of rhizomatousness in *O. longistaminata*, a large F_2 population containing 818 individuals was used to map loci that control various rhizome-related traits. These identified loci were further interwoven into putative networks to elucidate their interconnections. Our results highlight the intricate genetic regulation on rhizome development and may provide information for pinpointing key rhizome-related genes that may be utilized in future perennial rice breeding programs.

MATERIALS AND METHODS

Plant Materials

The mapping population was prepared from a cross between an unnamed *O. longistaminata* accession featuring long and strong rhizomes and the *O. sativa* cultivar RD23 (an indica cultivar from Thailand) (Tao and Sripichitt, 2000). The F_1 plant was grown at the Perennial Rice Research Station of Yunnan University located in Jinghong, Yunnan Province, China, a typically double rice cropping region with a tropical monsoon climate (20°57′ E, 100°45′ N, at an altitude of 555 m). A large number of F_2 seeds were obtained by bagged self-pollination and vegetative propagation using tillers. All F_2 plants were grown at the same region in Jinghong during the first cropping season of 2016, and the distance between F_2 plants was 50 cm. At the time of flowering, all plants were dug up, and the underground parts were washed free of soil for phenotypic evaluation.

Recombinant inbred lines were developed in our laboratory using the single seed descent technique. Since 2007, one single individual plant from the previous generation was randomly selected for selfing. The seeds were harvested and planted as a new generation. These steps were repeated, and finally, a RIL population consisting of 133 lines was obtained. This RIL population was employed to preliminarily validate the key loci identified in this study.

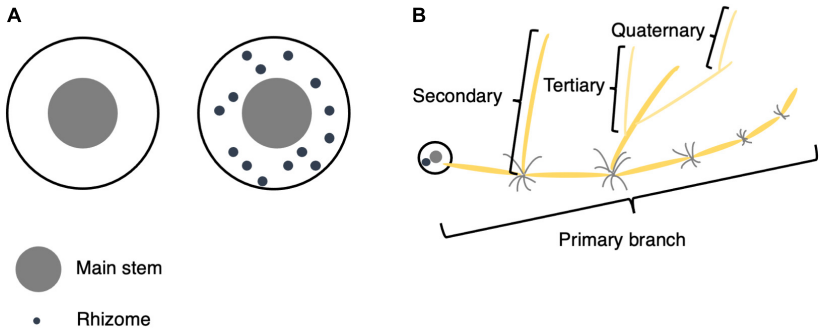

FIGURE 1 | Phenotypic evaluation of rhizomes. **(A)** Top view of plants with and without rhizomes. If a plant's underground part contained only the main stem and tillers, its Rhz was recorded as 0 (without rhizomes), and if there were obvious internodes in the underground part, its Rhz was recorded as 1 (with rhizomes). **(B)** Schematic illustration for primary, secondary, tertiary, and quaternary branches.

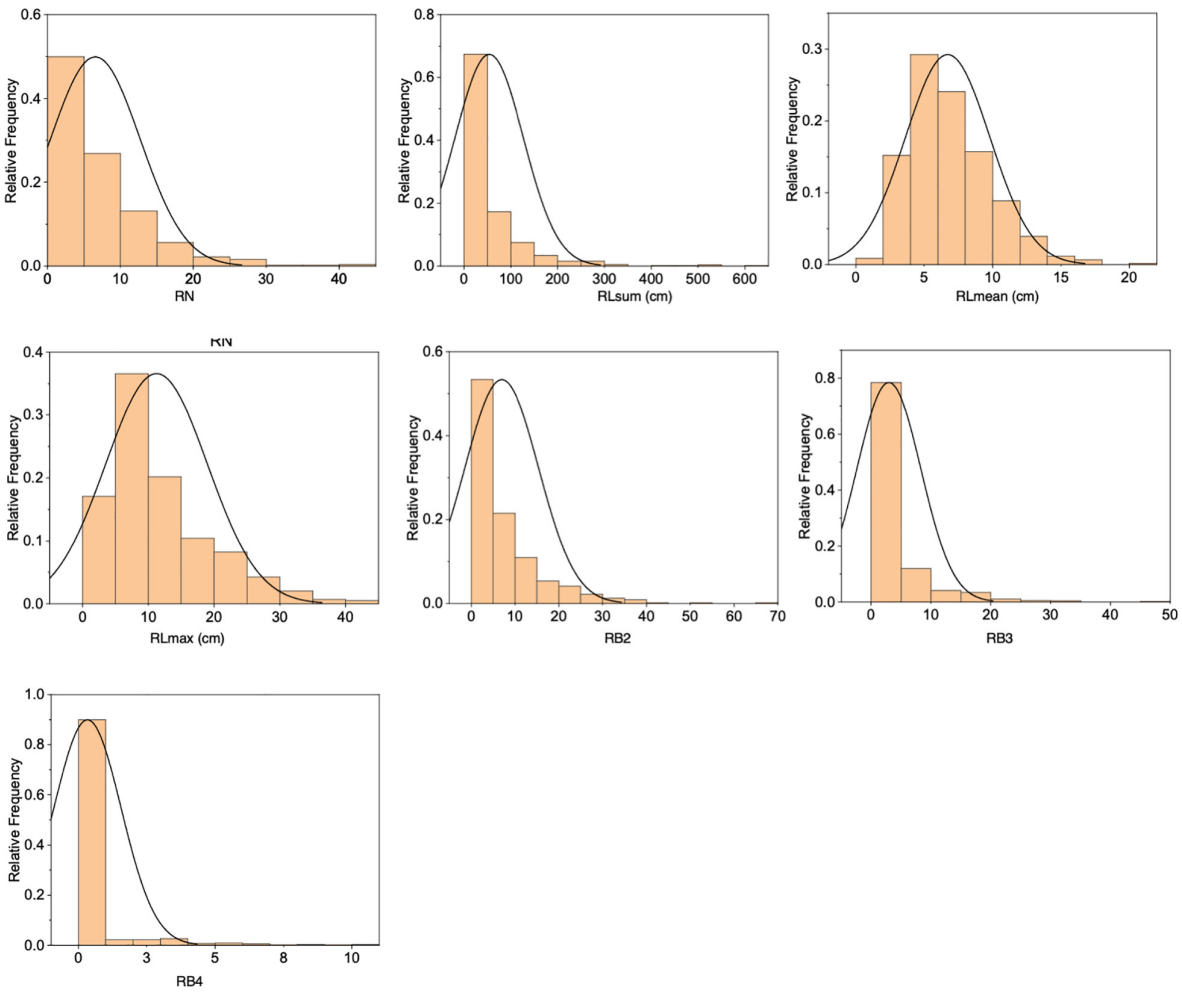

FIGURE 2 | Distributions of seven rhizome-related traits in the F_2 population.

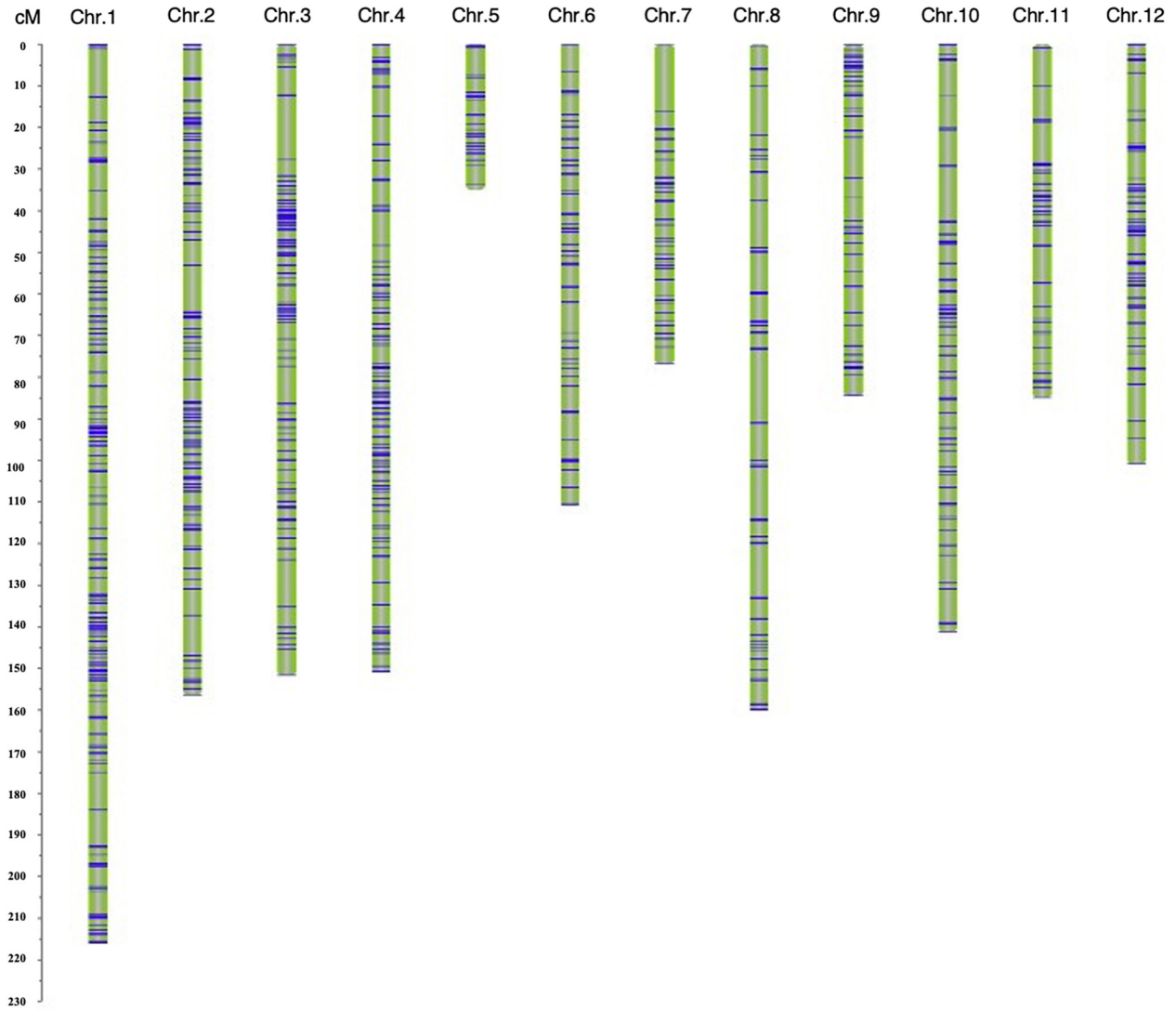

FIGURE 3 | The genetic linkage map of the F_2 population.

Phenotypic Evaluation

Each plant was first evaluated for the presence or absence of rhizomes (hereafter referred to as Rhz). For individuals with rhizomes, rhizome-related quantitative traits were measured, including rhizome number (RN), cumulative length of primary branches (RLsum), average length of primary branches (RLmean), maximum length of primary branches (RLmax), number of secondary branches (RB2), number of tertiary branches (RB3), and number of quaternary branches (RB4) per plant (**Figure 1**).

Genotyping-by-Sequencing Library Construction and Single Nucleotide Polymorphism Detection

DNA was extracted using the cetyltrimethylammonium bromide method (Doyle, 1987). A genotyping-by-sequencing library was prepared following the protocol developed by Clark et al. (2014). Briefly, DNA samples were first digested by the restriction enzymes *Pst*I and *Msp*I, and then sheared to 200–400 bp for restriction site-associated DNA library construction. Sequencing was performed on an Illumina HiSeq 2000 machine and 150-base paired-end reads were obtained.

Raw read files were decoded for each sample by using "process_radtags" in Stacks v2.3 according to their barcode files. Subsequently, the raw reads for each sample were aligned to the *O. sativa* genome (IRGSP V1.0 *O. sativa* Nipponbare, Kawahara et al., 2013) by using the Burrows-Wheeler Alignment software (parameter: mem −t 4 −M −R). The alignment results were formatted and converted into input files using SAMtools. The programs "pstacks," "cstacks," and "genotypes" in Stacks were used to perform SNP calling and obtain the genotypic data based on SNPs. The raw SNPs were filtered by an R script. First, SNPs that were not present in over 20% of individuals in

the population were filtered, and the second round of filtering was performed according to segregation ratios and parental genotypes: for each SNP, it should be homozygous in both RD23 and *O. longistaminata* genomes, and be able to differentiate *O. longistaminata* and RD23. Finally, 1665 high-quality SNPs were obtained for genetic map construction.

Qualitative Trait Locus Identification and Quantitative Trait Loci Analysis

Chi-square tests were performed to detect qualitative trait loci associated with Rhz, with the significance level setting to 5%. On one chromosome, the locus with the maximum chi-square value was regarded as a hypothetical major-effect locus. To find all independent loci on the same chromosome, correlation analyses between the hypothetical major-effect locus and other loci were carried out. All loci that were significantly associated with the hypothetical main locus ($P < 0.05$) were considered as dependent loci and removed, and the remaining loci were identified as independent major-effect loci.

The linkage map was constructed by using Joinmap 4.0 (Van Ooijen, 2006). Inclusive composite interval mapping was performed for QTL analysis using QTL IciMapping 4.2 (Lei et al., 2015). The LOD score threshold for QTL identification was set to 2.50. QTLs were named as described by McCouch et al. (1997). Note that the corresponding QTLs for RLmean and RLmax were named as *qRLM* and *qRLMAX*, respectively.

Network Construction

Using the molecular-quantitative genetic model developed by Zhang et al. (2011), a genetic network underlying each measured trait was constructed according to genetic expectations of the identified loci. Briefly, FGUs were identified by performing multiple comparisons of the corresponding phenotypic data of one genotype represented by two loci. When there was a significant phenotypic difference ($P < 0.05$) between one genotype and the other genotypes, and the interaction effect between the two loci was equivalent to the main effect, these two loci were grouped into an FGU. According to whether one locus was significant responsible for the phenotypical difference or not, the locus and the other locus in a locus pair were regarded as hierarchy QTLs (directional links) or epistatic QTLs/loci (E-QTLs/loci, non-directional links), respectively. The integrated network of all identified loci was assembled based on loci that were found to affect more than one rhizome-related trait.

Verification of Key Quantitative Trait Loci

Locus sequences were obtained from the National Centre for Biotechnology Information and our unpublished *O. longistaminata* genome data (data not shown). Sequence-tagged site (STS) markers were designed by using Primer 5.0. The primer sequences were *qRHZ-3-1*-F (5′-CTACCAGGTTCGTT GATGTC-3′) and *qRHZ-3-1*-R (5′-CGAGGTACATCGTCTT GGA-3′) for *qRHZ-3-1*, *qRHZ-4*-F (5′-CAGACGGATTGAATC GATACCA-3′), and *qRHZ-4*-R (5′-CCATTTTCCCTGTTC ATCCATC-3′) for *qRHZ-4*, *qRHI-2*-F (5′-ATAAAATGGTATG GTGTAATGG-3′) and *qRHI-2*-R (5′-TGTTTCGCATTGCA TCTG-3′) for *qRHI-2*, *qRHI-5*-F (5′-TTAGCTCTCACAAA TGAATATC-3′), and *qRHI-5*-R (5′-TATCCAACCCTTCAA ACG-3′) for *qRHI-5*. PCR was performed in a 20 μL reaction containing 1 μL genomic DNA extracted from one RIL individual, 0.5 μL Primer-F/R, 2 μL 10 × Taq Buffer, 1.6 μL dNTP, 0.5 μL Taq polymerase, and 13.9 μL ddH_2O. Amplification was programmed for 3 min at 94°C for initial denaturation and 30–35 cycles consisting of 30 s at 94°C, 30 s at 56–60°C, 20 s at 72°C, followed by a final 5 min at 72°C. Amplified DNA fragments were detected using 8% polyacrylamide gel electrophoresis.

TABLE 1 | Major-effect loci for Rhz.

Locus	Chr	Dis (cM)	Marker	Physical location	χ^2	*P*
qRHZ-1	1	183.96	mk939	C0315063831	18.04	0.003
qRHZ-2	2	115.73	mk522	C026188261	33.69	0.000
qRHZ-3-1	3	38.19	mk876	C034764149	46.06	0.000
qRHZ-3-2	3	121.15	mk1010	C0335005202	14.75	0.012
qRHZ-4	4	116.23	mk1097	C0419314643	119.16	0.000
qRHZ-5	5	8.21	mk981	C0326657391	17.82	0.003
qRHZ-6	6	110.44	mk1347	C0617172779	14.33	0.014
qRHZ-7	7	27.76	mk1468	C0726307967	29.46	0.000
qRHZ-8	8	143.42	mk1536	C0824312935	15.22	0.009
qRHZ-9	9	1.44	mk468	C1120950915	13.12	0.022
qRHZ-10	10	138.8	mk115	C1023085906	33.85	0.000
qRHZ-11	11	42.79	mk420	C1110822965	24.04	0.000
qRHZ-12	12	60.94	mk793	C125003428	30.69	0.000

Chr, chromosome; Dis, distance.

RESULTS

Phenotypes of the F_2 Population

The F_2 population derived from a cross between RD23 and *O. longistaminata* contained 818 individuals, of which 585 individuals had rhizomes. Significant variations were observed in the seven measured rhizome-related traits (RN, RLsum, RLmean, RLmax, RB2, RB3, and RB4), all exhibiting a skewed distribution (**Figure 2** and **Supplementary Table 1**). The correlation analysis of seven rhizome-related quantitative traits revealed that there were significant correlations between these traits, with the correlation coefficients of RN and RLsum, and RLmax and RLmean reached 0.934 and 0.873, respectively (**Supplementary Table 2**).

Construction of Linkage Map of the F_2 Population

A genetic map was constructed based on 1487 SNP markers spanning 1469.27 cM and covering all 12 chromosomes. The length of a single chromosome was between 35.04 and 216.00 cM, and the average distance between makers was 0.99 cM (**Supplementary Table 3**). Among the 12 chromosomes, chromosome 1 harbored the largest number (267) of markers, and the average distance between markers was 0.81 cM, whereas the number of markers on chromosome 5 was the lowest

TABLE 2 | Loci affecting rhizome-related traits in the F_2 population.

QTL	Trait	Chr	Marker interval	LOD/χ^2	PVE (%)	Add	Dom
qRLSUM-1-1	RLsum	1	mk817-mk816	8.29	3.94	22.87	0.45
qRN-1-1	RN	1	mk816-mk145	8.27	4.61	2.16	−0.05
qRB2-1-1	RB2	1	mk166-mk163	6.72	3.72	2.46	0.81
qRHI-1	RLmax	1	mk171-mk174	3.81	2.1	1.83	0.17
	RB3			5.68	3.57	1.66	0.06
qRN-1-2	RN	1	mk286-mk287	4.12	2.21	0.25	1.78
qRLMAX-1-2	RLmax	1	mk300-mk305	2.71	1.53	1.43	0.29
qRLSUM-1-2	RLsum	1	mk330-mk332	4.15	2.04	5.1	19.54
qRLM-1	RLmean	1	mk340-mk343	4.13	2.46	0.24	0.88
qRB2-1-2	RB2	1	mk358-mk359	2.8	1.54	−0.47	2.09
qRHZ-1	Rhz	1	mk939	18.04			
qRB4-2	RB4	2	mk695-mk692	34.27	42.62	3.01	−2.98
qRB2-2	RB2	2	mk566-mk572	4.79	7.95	14.31	−13.8
qRB3-2	RB3	2	mk532-mk525	23.05	27.32	9.26	−9.92
qRHI-2	Rhz	2	mk522-mk521	33.69			
	RN			15.55	8.45	3.6	−0.27
qRLSUM-2	RLsum	2	mk519-mk517	117.21	89.81	−143.39	17.44
qRHZ-3-1	Rhz	3	mk876	46.06			
qRN-3	RN	3	mk895-mk894	2.92	1.38	−2.1	−3.26
qRB3-3	RB3	3	mk926-mk930	4.39	3.1	−6.89	−7.42
qRHI-3	RLsum	3	mk997-mk971	2.56	7.88	−82.69	−89.46
	RB2			2.79	11.23	−8.1	−8.63
	RB4			53.94	47.99	−2.77	−2.76
qRHZ-3-2	Rhz	3	mk1010	14.75			
qRHI-4-1	RB3	4	mk1029-mk1053	15.48	22.31	9.11	−9.83
	RB4			58.04	47.95	3.26	−3.28
qRB2-4-1	RB2	4	mk1063-mk1258	6.04	11.26	11.44	−12.39
qRHI-4-2	RN	4	mk1257-mk1045	8.98	8.1	8.46	−8.73
	RLsum			19.76	13.76	166.4	−165.43
qRB2-4-2	RB2	4	mk1118-mk1116	11.67	6.66	3.13	0.49
qRLSUM-4-2	RLsum	4	mk1100-mk1098	8.87	4.38	22.56	−5.36
qRLMAX-4	RLmax	4	mk1098-mk1096	20.28	12.42	4	−0.72
qRHZ-4	Rhz	4	mk1097	119.16			
qRN-4-2	RN	4	mk1082-mk1086	6.3	3.12	1.56	−0.64
qRLM-4	RLmean	4	mk1085-mk1088	15.76	10.39	1.57	−0.23
qRHI-4-3	RB4	4	mk1091-mk1078	61.91	52.25	2.75	−2.9
	RB3		mk1091-mk1078	19.31	23.3	9.54	−10.08
qRB2-5	RB2	5	mk460-mk980	10.7	5.88	1.31	2.87
qRHI-5	Rhz	5	mk981-mk1270	17.82			
	RLmean			8.13	5.83	−0.12	1.63
	RLmax			6.84	4.19	0.99	2.22
qRN-5	RN	5	mk1268-mk1266	13.59	6.87	1.46	1.83
qRB3-5	RB3	5	mk1266-mk1265	7.08	13.66	7.23	−7.27
qRLSUM-5	RLsum	5	mk1220-mk1219	17.53	15.96	124.29	−129.87
qRB3-6	RB3	6	mk1636-mk1367	8.74	27.59	10.02	−10.1
qRHZ-6	Rhz	6	mk1347	14.33			
qRLSUM-7	RLsum	7	mk1435-mk1440	7.79	11.13	−104.8	−96.25
qRB4-7	RB4	7	mk1440-mk1458	26.37	27.89	−3.29	−3.07
qRHZ-7	Rhz	7	mk1468	29.46			
qRHI-8	RN	8	mk1272-mk512	7.34	8.27	8.92	−8.66
	RLsum			22.93	18.09	137.2	−137.76
	RB3			21.14	26.44	9.39	−9.37
qRB2-8	RB2	8	mk1274-mk1535	5.24	14.15	9.14	−9.65

(Continued)

TABLE 2 | (Continued)

QTL	Trait	Chr	Marker interval	LOD/χ^2	PVE (%)	Add	Dom
qRB4-8-1	RB4	8	mk1474-mk1350	11.98	28.96	2.48	−2.65
qRLM-8	RLmean	8	mk1481-mk941	3.13	1.92	−1	0.22
qRB4-8-2	RB4	8	mk1510-mk503	88.25	61.16	2.76	−2.79
qRHZ-8	Rhz	8	mk1536	15.22			
qRHZ-9	Rhz	9	mk468	13.12			
qRHI-9-1	RB2	9	mk1648-mk1650	5.07	7.77	8.84	−7.5
	RB3			9.47	20.66	11.62	−11.18
	RLmean			4.38	2.78	1.66	−0.87
	RLmax			5.12	2.93	3.34	−1
qRHI-9-2	RN	9	mk1649-mk1230	6.31	10.45	6.62	−6.89
	RLsum			22.03	16.98	127.18	−128.99
qRHI-10	RN	10	mk100-mk1645	5.86	8.68	4.45	−3.64
	RLsum			13.94	16.33	77.3	−68.09
	RB3			4.2	21.25	6.43	−5.68
qRHZ-10	Rhz	10	mk115	33.85			
qRB3-11	RB3	11	mk1249-mk450	10.3	21.34	10.33	−10.61
qRHI-11	RN	11	mk422-mk432	9.94	5.12	−1.92	−0.5
	RLsum			6.94	3.33	−18.35	−4.95
	RB2			6.91	3.76	−2.27	−0.5
qRHZ-11	Rhz	11	mk420	24.04			
qRB4-11	RB4	11	mk389-mk388	3.42	2.8	0.24	−0.3
qRLM-11	RLmean	11	mk388-mk387	2.69	1.75	0.57	−0.22
qRB2-12	RB2	12	mk829-mk775	5.53	11.96	−8.73	−6.84
qRB4-12	RB4	12	mk775-mk774	8.39	12.21	−2.73	−2.78
qRB3-12	RB3	12	mk763-mk738	2.56	6.49	−6.53	−6.15
qRHI-12	RN	12	mk743-mk744	5.37	3.82	−14.36	−13.82
	RLsum			5.53	3.73	−169.15	−163.57
qRHZ-12	Rhz	12	mk793	30.69			

Chr, chromosome; PVE, phenotypic variation explained; Add and Dom, QTL additive and dominance effects, respectively.

(38 markers), with the average distance being 0.92 cM (**Figure 3** and **Supplementary Table 3**).

Qualitative Trait Loci for Rhz and Quantitative Trait Loci Affecting Rhizome Development

In all, 13 regions located in 12 chromosomes were identified as qualitative trait loci that were significantly associated with Rhz (**Table 1** and **Supplementary Table 4**). Among these loci, chromosome 4-located *qRHZ-4* showed the most significant segregation distortion ($\chi^2 = 119.16$), and it was therefore hypothesized as one major Rhz locus. On this premise, the effects of the remaining 12 Rhz loci were further analyzed after the interference of *qRHZ-4* was removed. The results showed that, except for *qRHZ-2*, no other loci were significantly associated with Rhz. Hence, *qRHZ-2* could be regarded as another major hypothetical Rhz locus (**Supplementary Table 5**).

For the rhizome abundance traits measured in this study, a total of 51 QTLs were identified, while their effects on the phenotype varied greatly (**Table 2** and **Figure 4**). For example, among the 12 QTLs that were found to affect RN, the highest phenotypic variation explained (PVE) was 10.45, whereas the lowest PVE was 1.38. Interestingly, we found that 11 QTLs may control at least two traits, and they were re-termed *qRHI*. Most of these QTLs appeared to affect multiple highly correlated traits, such as *qRHI-1* for RLmax and RB3; *qRHI-3* for RLsum, RB2, and RB4; *qRHI-4-1* and *qRHI-4-3* for RB3 and RB4; and *qRHI-4-2* for RLsum and RN. It is worth noting that three regions identified to determine Rhz were also found to be located in QTL intervals affecting other rhizome-related traits, such as *qRHI-2* for Rhz and RN, and *qRHI-5* for Rhz, RLmean and RLmax.

qRHZ-3-1, *qRHZ-4*, *qRHI-2*, and *qRHI-5* Are Key Quantitative Trait Loci in the Genetic Network Underlying Rhizome Development

To further understand the relationships between and among the identified loci, we constructed a genetic network for each trait based on FGUs (**Figure 5** and **Supplementary Table 6**). A total of 12 FGUs affecting Rhz were detected, consisting of 2 pairs of hierarchical loci and 29 pairs of E-loci (**Figure 5A** and **Supplementary Table 6**). This network displayed a clear hierarchical structure with *qRHZ-4* as the highest-level node. Another similar highly hierarchical network was the RLsum network, with *qRLsum* as the top node and each layer containing

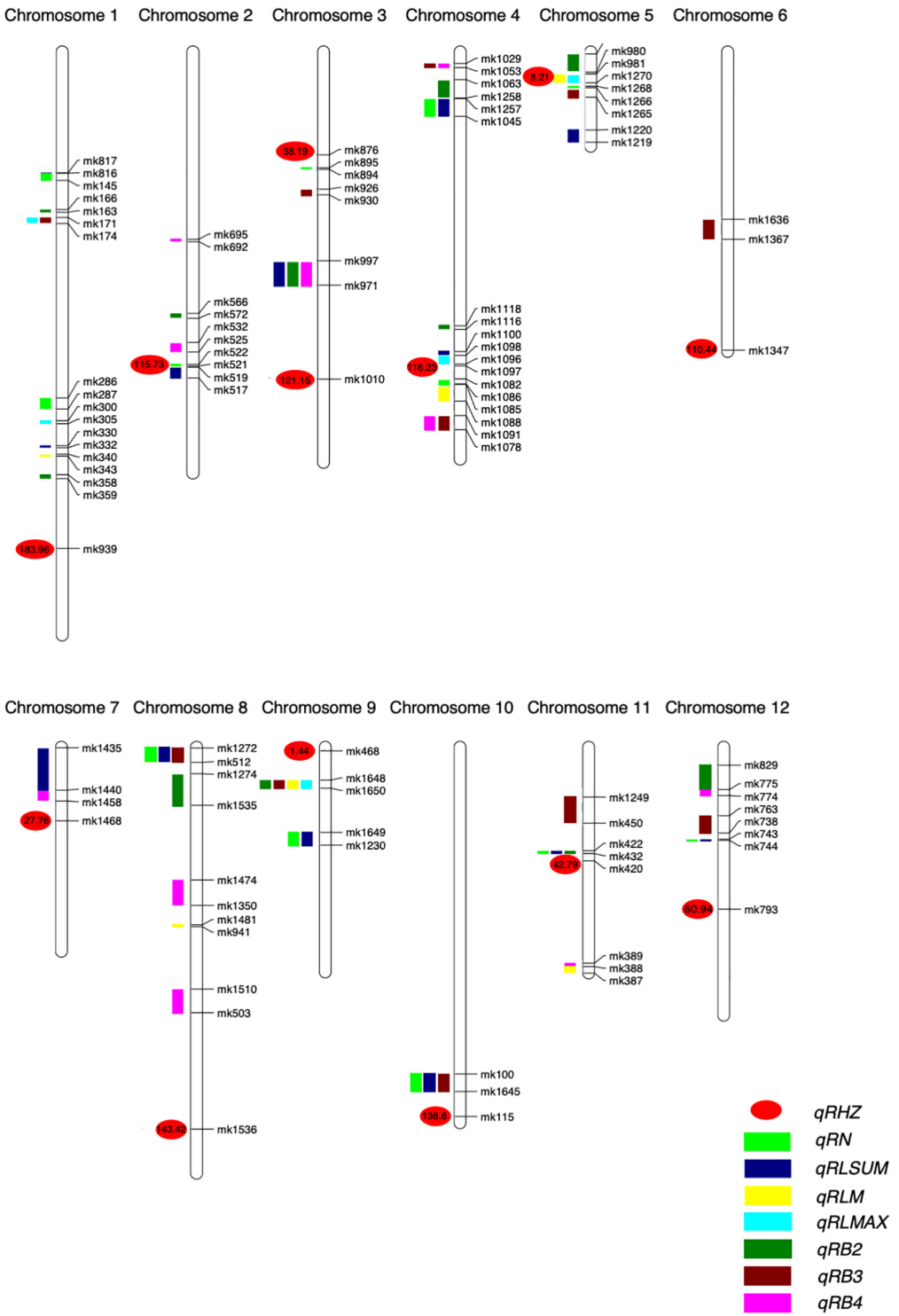

FIGURE 4 | Positions of QTLs on the genetic linkage map of the F_2 population. Numbers are genetic distances.

more nodes along with the hierarchy (**Figure 5C**). By contrast, the networks for RB and RB4 appeared to be non-hierarchical, with no obvious loci located on the upstream of other loci (**Figures 5F,H**). The network for RLmax was the network of the simplest form (containing only three nodes), in which *qRLmax-4* was hierarchically connected with other two nodes (**Figure 5E**).

Considering the possible interactions among loci for different traits, we further integrated the individual networks for single

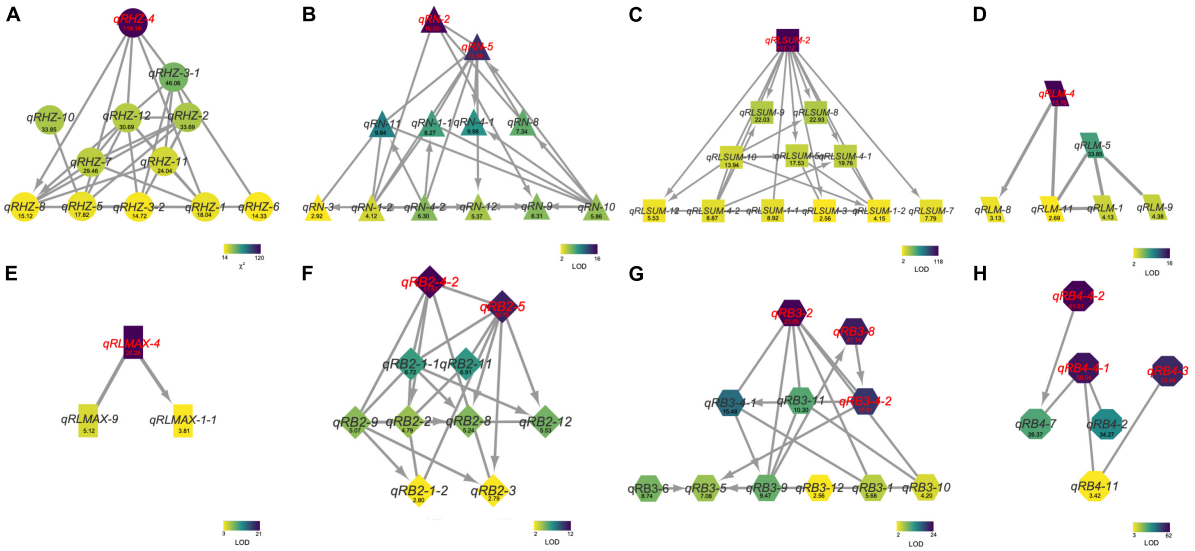

FIGURE 5 | The genetic networks for Rhz **(A)**, RN **(B)**, RLsum **(C)**, RLmean **(D)**, RLmax **(E)**, RB2 **(F)**, RB3 **(G)**, and RB4 **(H)**. A line between two nodes indicates that these two nodes are a pair of E-loci **(A)** or E-QTLs **(B–H)**. An arrow between two nodes represents a hierarchical relationship between this node pair, with the node at the arrowhead being regulated by the other node of the pair. Numbers are Chi square values **(A)** or LOD values **(B–H)**.

traits into a comprehensive network (**Figure 6**). In this network, a total of 58 FGUs were included, containing 62 pairs of hierarchy loci/QTLs and 104 pairs of E-loci/QTLs. Based on network properties, we extracted three sub-networks and termed them as network-RHZ, network-RB, and network-RN/L, which corresponded to Rhz, rhizome branching traits, and rhizome number/length, respectively. Given that network-RHZ should be the prerequisite for other two networks, its nodes with the most marked effects, *qRHZ-3-1* and *qRHZ-4*, and its connector nodes with the other two networks, *qRHI-2* and *qRHI-5* were classified as key loci in the network (**Figure 6**).

Preliminary Verification of Key Loci in the Networks

To preliminarily verify the biological function of the four key loci we identified, we detected the genotypes of a rhizome-free RIL population containing 133 individuals by using the STS markers of these loci, and observed the corresponding phenotypes. For each individual locus, most individuals of the RIL population were RD23 homozygous, but a few *O. longistaminata* homozygous individuals were also detected (**Supplementary Table 7**). There were five individuals harboring two pairs of *O. longistaminata* homozygous alleles, namely, RIL 129, 117, 303, 304, and 322 (**Figure 7**; as RIL 303 and 304 are of the same genotype, only the phenotype of RIL 303 is shown). The observation that all RILs with homozygous *O. longistaminata* alleles at single locus were rhizome-free confirmed that no single locus can ensure the presence of rhizomes. Moreover, none of these rhizome-free individuals were detected to harbor more than two of these key loci with *O. longistaminata* alleles, suggesting that three or more loci with *O. longistaminata* alleles are required for the presence of rhizomes. Given that we failed to detect any genotype with homozygous *O. longistaminata* alleles at both *qRHI-2* and *qRHI-5*, two multi-trait controlling loci that appeared to be critical in our network analyses, we hypothesize that these two loci are likely both needed for proper rhizome growth.

DISCUSSION

In the present study, we substantially expanded the genetic repertoire for rhizome development. A recent manuscript (Fan et al., 2020) identified over 10 loci related to rhizome growth, including five major-effect loci—*qRED1.2*, *qRED3.1*, *qRED3.3*, *qRED4.1*, and *qRED4.2*, some of which are partially overlapped with two previously mapped loci determining the presence of rhizome, *Rhz2* and *Rhz3* (Hu et al., 2003). Here, we identified 13 Rhz-regulating loci and as many as 51 QTLs controlling rhizome abundance. Not surprisingly, some of them are overlapped with those previously reported QTLs. Comparing with Fan et al. (2020), eight of our identified loci (*qRN-1-2*, *qRLSUM-2*, *qRHZ-3-1*, *qRHZ-3-1*, *qRHI-3*, *qRB2-4-2*, *qRHI-11*, and *qRHZ-11*) are overlapped with their QTL intervals. Among these, two Rhz-regulating loci, *qRHZ-3-1* and *qRHZ-11*, are located in the intervals of *qRED3.1* and *qRED-11*, respectively, highlighting the potential prerequisite role of these loci, and *qRHI-3* and *qRHI-11*, two intervals controlling multiple rhizome traits, were found to coincide with *qRED3.3* and *qRED11* (Fan et al., 2020). Meanwhile, many rhizome abundance QTLs we identified are closely linked with their major-effect loci, such as *qRB2-2* and *qRED2.2*, *qRN-3* and *qRED3.1*, and *qRLSUM-4-2* and *qRED4.1* (Fan et al., 2020). In general, these two studies are mutually confirmative; however, it is important to note that, instead of using the average rhizome extension distance (measured in soil surface, Fan et al., 2020), we treated the presence of rhizomes

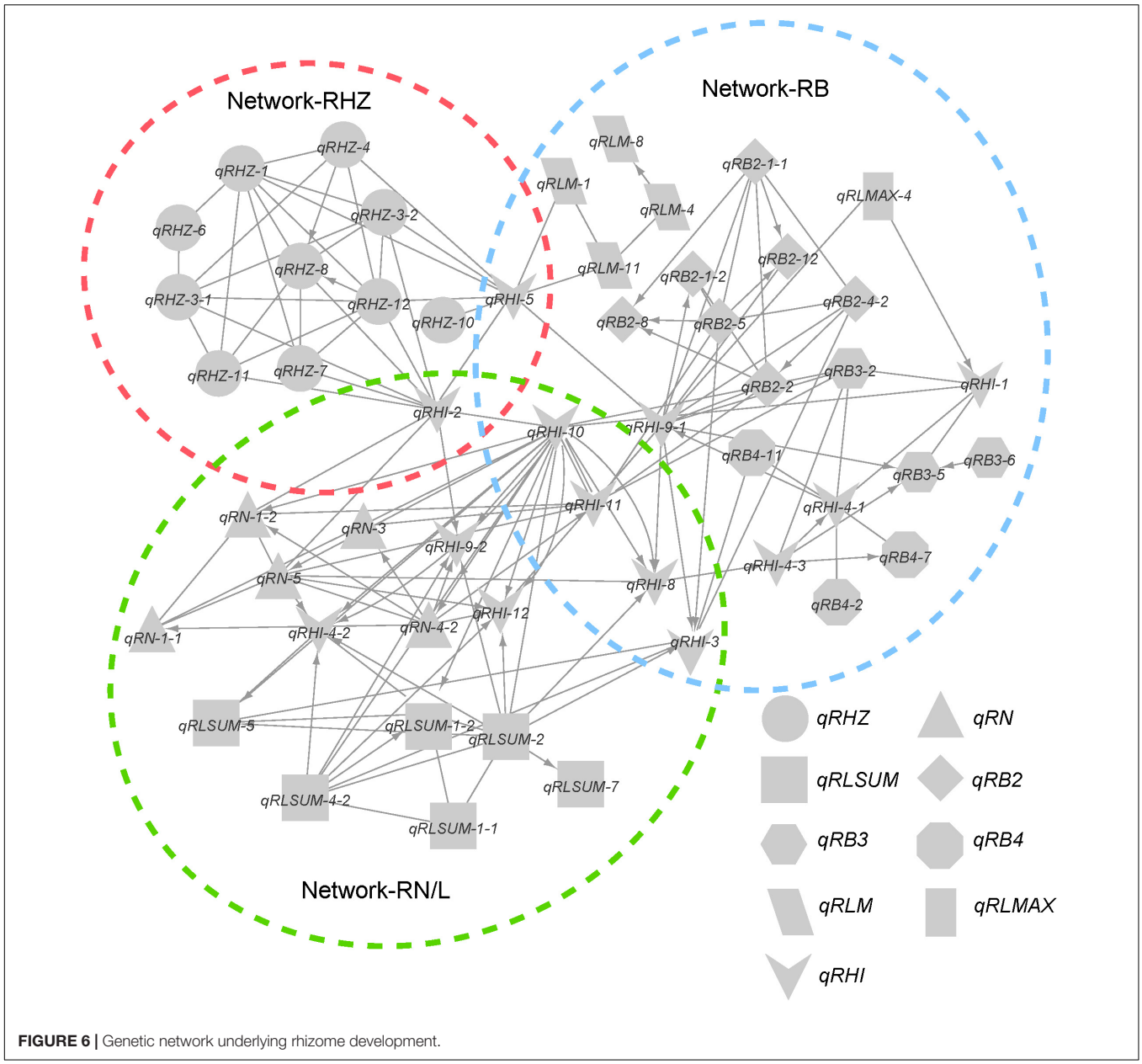

FIGURE 6 | Genetic network underlying rhizome development.

as a qualitative trait and carefully evaluated specific rhizome-related traits by digging the plants up. This may explain why our results do not completely include the previously identified QTLs, and such comprehensive phenotypic evaluation also evidently gave rise a larger number of identified QTLs in our study. Furthermore, the results presented here are also largely consistent with our previous study (Hu et al., 2003). A total of six pairs of loci are overlapped, and many of these loci pairs were identified to control the same or similar trait, such as *qRN-1-2* and *QRl1*, and *qRBD3-2* and *QRbd2*. The disparity between our current and previous results may stem from the distinct populations. Moreover, compared with previously used simple-sequence repeat (Hu et al., 2003) and insertion/deletion markers (Fan et al., 2020), the SNP markers used in this study have relatively wider coverage and higher density in the whole genome; thus, more QTLs have emerged.

Multiple genetic networks were established on the basis of classical QTL mapping (Zhang et al., 2011; **Figures 5**, **6**). We observed a clear hierarchical structure in some networks for individual traits (**Figure 5**). This indicates that some loci, such as *qRHZ-4* in the Rhz network (**Figure 5A**) and *qRLsum-2* in the RLsum network (**Figure 5C**), may play a more significant regulatory role than the other loci in the same network since upstream loci are likely prerequisite for downstream loci to execute their functions. Given that the other traits measured in this study all depend on the presence of rhizomes, the two top-located nodes in the RHZ network are presumably of primary importance for proper rhizome development. Furthermore, in

FIGURE 7 | Phenotypes of representative rhizome-free RILs. "+" and "–" represent *O. longistaminata* and RD23 homozygous alleles at the locus, respectively.

the comprehensive network containing all the identified loci, *qRHI-2* and *qRHI-5* were the hub loci connecting network-RHZ with network-RB and network-RN/L. In network analyses, hub nodes are generally considered to be critical (Zhang et al., 2011). In our case, considering that no other networks would exist if network-RHZ collapses, *qRHI-2* and *qRHI-5* are therefore considered as secondarily important loci for rhizome development. It is attempting to speculate that these four loci might be the minimal set of loci required for rhizome development, and this is worth of further investigation. Of note, the concept of FGU in this study is defined as a pair of mutual functional dependent genes, whose encoding proteins may not necessarily physically interact with each other, and FGUs in our networks were generated by comparing the genes' contributions to the phenotype. Compared with RIL and doubled haploid populations, F_2 populations usually have a larger number of distinct genotypes, and the efficacy of detecting genetic networks using F_2 populations is weakened due to the possible errors occurred in multiple comparisons (Zhang et al., 2011). Such weaking effect would be amplified if there are a large number of loci controlling the trait (Zhang et al., 2011). Therefore, even though our preliminary functional verification suggested that the two hub loci, *qRHI-2* and *qRHI-5*, are likely both needed for the rhizome phenotype, we could not exclude the possibility that other loci are also required here. Further verification using rhizomatous accessions and functional characterization of these loci would be needed to unravel the complex genetic basis of rhizome development.

CONCLUSION

In this study, we detected 62 loci that are putative regulators for rhizome development. By grouping these loci into FGUs, we provided evidence that rhizome growth is controlled by a multi-locus, multi-layered genetic network. Based upon the network structure and the interactions among the nodes, we predicted four loci, two major-effect loci for rhizome initiation and two hub loci connecting the individual genetic networks, as key loci. Functional verification using rhizome-free RILs confirmed that none of these key loci could solely initiate rhizome morphogenesis and suggested that *qRHI-2* and *qRHI-5* are likely both needed for rhizome initiation.

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

FH and SZ designed the experiments. WL, SZ, LH, GH, and JZ performed the experiments. FH, ZL, and WL wrote the manuscript. WL, SZ, and ZL analyzed the data. All authors have read and approved the final manuscript.

FUNDING

This work was funded by the National Natural Science Foundation of China (U1302264 and 31660387).

ACKNOWLEDGMENTS

We thank members in our laboratory for their help with phenotypical evaluation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.866165/full#supplementary-material

REFERENCES

Clark, L. V., Brummer, J. E., Głowacka, K., Hall, M. C., Heo, K., Peng, J., et al. (2014). A footprint of past climate change on the diversity and population structure of *Miscanthus sinensis*. *Ann. Bot.* 114, 97–107. doi: 10.1093/aob/mcu084

Cox, T. S., Glover, J. D., Van Tassel, D. L., Cox, C. M., and DeHaan, L. R. (2006). Prospects for developing perennial grain crops. *BioScience* 56, 649–659. doi: 10.1641/0006-3568(2006)56[649:pfdpgc]2.0.co;2

Doyle, J. J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19, 11–15.

Fan, Z., Wang, K., Rao, J., Cai, Z., Tao, L., Fan, Y., et al. (2020). Interactions among multiple quantitative trait loci underlie rhizome development of perennial rice. *Front. Plant Sci.* 11:591157. doi: 10.3389/fpls.2020.591157

Ghesquiere, A. (2008). Evolution of oryza longistaminata. *Rice Genet.* 1, 15–25. doi: 10.1142/9789812814265_0002

Gizmawy, I., Kigel, J., Koller, D., and Ofir, M. (1985). Initiation, orientation and early development of primary rhizomes in *Sorghum halepense* (L.) pers. *Ann. Bot.* 55, 343–350. doi: 10.1093/oxfordjournals.aob.a086911

Glover, J. D., Reganold, J. P., Bell, L. W., Borevitz, J., Brummer, E. C., Buckler, E. S., et al. (2010). Increased food and ecosystem security via perennial grains. *Science* 328, 1638–1639. doi: 10.1126/science.1188761

Hu, F. Y., Tao, D. Y., Sacks, E., Fu, B. Y., Xu, P., Li, J., et al. (2003). Convergent evolution of perenniality in rice and sorghum. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4050–4054. doi: 10.1073/pnas.0630531100

Hu, F., Wang, D., Zhao, X., Zhang, T., Sun, H., Zhu, L., et al. (2011). Identification of rhizome-specific genes by genome-wide differential expression analysis in oryza longistaminata. *BMC Plant Biol.* 11:18. doi: 10.1186/1471-2229-11-18

Huang, G., Qin, S., Zhang, S., Cai, X., Wu, S., Dao, J., et al. (2018). Performance, economics and potential impact of perennial rice PR23 relative to annual rice cultivars at multiple locations in yunnan province of China. *Sustainability* 10:1086. doi: 10.3390/su10041086

Jang, C. S., Kamps, T. L., Skinner, D. N., Schulze, S. R., Vencill, W. K., and Paterson, A. H. (2006). Functional classification, genomic organization, putatively cis-acting regulatory elements, and relationship to quantitative trait loci, of sorghum genes with rhizome-enriched expression. *Plant Physiol.* 142, 1148–1159. doi: 10.1104/pp.106.082891

Kawahara, Y., de la Bastide, M., Hamilton, J. P., Kanamori, H., McCombie, W. R., Ouyang, S., et al. (2013). Improvement of the oryza sativa nipponbare reference genome using next generation sequence and optical map data. *Rice* 6:4. doi: 10.1186/1939-8433-6-4

Li, Z., Lathe, R. S., Li, J., He, H., and Bhalerao, R. P. (2022). Towards understanding the biological foundations of perenniality. *Trends Plant Sci.* 27, 56–68. doi: 10.1016/j.tplants.2021.08.007

Mackay, T. F. C., Stone, E. A., and Ayroles, J. F. (2009). The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* 10, 565–577. doi: 10.1038/nrg2612

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

Meakawa, M., Inukai, T., Rikiishi, K., and Govindaraj, K. G. (1998). Inheritance of the rhizomatous traits in hybrid of *Oryza longistaminata* Chev. et Roehr. and *O. sativa* L. *SABRAO J. Breed. Genet.* 30, 67–72.

Lei, M, Li, H., Zhang, L., and Wang, J. (2015). QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in bi-parental populations. *Crop J.* 3, 269–283. doi: 10.1016/j.cj.2015.01.001

Morishima, O. H. (1967). Variation in the breeding system of a wild rice, oryza perennis. *Evolution* 21, 249–258. doi: 10.2307/2406673

Paterson, A. H., Schertz, K. F., Lin, Y. R., Liu, S. C., and Chang, Y. L. (1995). The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.). *Proc. Natl. Acad. Sci. U.S.A.* 92, 6127–6131. doi: 10.1073/pnas.92.13.6127

Sacks, E., Roxas, J. P., and Cruz, M. S. T. (2003). Developing perennial upland rice II: field performance of S families from an intermated *Oryza sativa*/*Oryza longistaminata* population. *Crop Sci.* 43, 129–134. doi: 10.2135/cropsci2003.1290

Tao, D., and Sripichitt, P. (2000). Preliminary report on transfer traits of vegetative propagation from wild rice species to oryza sativa via distant hybridization and embryo rescue. *Kasetsart J.* 34, 1–11.

Van Ooijen, J. (2006). *JoinMap 4, Software for the Calculation of Genetic Linkage Maps in Experimental Populations*. Wageningen: Kyazma BV.

Wang, W., Fu, B., Ali, J., Xu, J., Gao, Y., Zheng, T., et al. (2015). Genome-wide responses to selection and genetic networks underlying submergence tolerance in rice. *Plant Genome* 8:66. doi: 10.3835/plantgenome2014.10.0066

Zhang, F., Jiang, Y. Z., Yu, S. B., Ali, J., Paterson, A. H., Khush, G. S., et al. (2013). Three genetic systems controlling growth, development and productivity of rice (*Oryza sativa* L.): a reevaluation of the 'green revolution'. *Theor. Appl. Genet.* 126, 1011–1024. doi: 10.1007/s00122-012-2033-1

Zhang, F., Zhai, H.-Q., Paterson, A. H., Xu, J.-L., Gao, Y.-M., Zheng, T.-Q., et al. (2011). Dissecting genetic networks underlying complex phenotypes: the theoretical framework. *PLoS One* 6:e14541. doi: 10.1371/journal.pone.0014541

Zhang, S., Hu, J., Yang, C., Liu, H., Yang, F., Zhou, J., et al. (2017). Genotype by environment interactions for grain yield of perennial rice derivatives (*Oryza sativa* L./ *Oryza longistaminata*) in southern China and laos. *Field Crops Res.* 207, 62–70. doi: 10.1016/j.fcr.2017.03.007

Zhang, S., Huang, G., Zhang, J., Huang, L., Cheng, M., Wang, Z., et al. (2019). Genotype by environment interactions for performance of perennial rice genotypes (*Oryza sativa* L./*Oryza longistaminata*) relative to annual rice genotypes over regrowth cycles and locations in southern China. *Field Crops Res.* 241:107556. doi: 10.1016/j.fcr.2019.107556

Zhang, Y., Huang, G., Zhang, S., Zhang, J., Gan, S., Cheng, M., et al. (2021). An innovated crop management scheme for perennial rice cropping system and its impacts on sustainable rice production. *Eur. J. Agron.* 122:126186. doi: 10.1016/j.eja.2020.126186

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

REVIEW
published: 26 April 2022
doi: 10.3389/fpls.2022.863789

Introgression Lines: Valuable Resources for Functional Genomics Research and Breeding in Rice (*Oryza sativa* L.)

*Bo Zhang[1], Ling Ma[1], Bi Wu[1], Yongzhong Xing[1] and Xianjin Qiu[2]**

[1] National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research, Huazhong Agricultural University, Wuhan, China, [2] College of Agriculture, Yangtze University, Jingzhou, China

OPEN ACCESS

Edited by:
Dayun Tao,
Yunnan Academy of Agricultural Sciences, China

Reviewed by:
Zhang Yu,
Yunnan Academy of Agricultural Sciences, China
Sung-Ryul Kim,
International Rice Research Institute (IRRI), Philippines

****Correspondence:***
Xianjin Qiu
xjqiu216@yangtzeu.edu.cn

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

Received: *27 January 2022*
Accepted: *01 April 2022*
Published: *26 April 2022*

Citation:
Zhang B, Ma L, Wu B, Xing Y and Qiu X (2022) Introgression Lines: Valuable Resources for Functional Genomics Research and Breeding in Rice (Oryza sativa L.). Front. Plant Sci. 13:863789. doi: 10.3389/fpls.2022.863789

The narrow base of genetic diversity of modern rice varieties is mainly attributed to the overuse of the common backbone parents that leads to the lack of varied favorable alleles in the process of breeding new varieties. Introgression lines (ILs) developed by a backcross strategy combined with marker-assisted selection (MAS) are powerful prebreeding tools for broadening the genetic base of existing cultivars. They have high power for mapping quantitative trait loci (QTLs) either with major or minor effects, and are used for precisely evaluating the genetic effects of QTLs and detecting the gene-by-gene or gene-by-environment interactions due to their low genetic background noise. ILs developed from multiple donors in a fixed background can be used as an IL platform to identify the best alleles or allele combinations for breeding by design. In the present paper, we reviewed the recent achievements from ILs in rice functional genomics research and breeding, including the genetic dissection of complex traits, identification of elite alleles and background-independent and epistatic QTLs, analysis of genetic interaction, and genetic improvement of single and multiple target traits. We also discussed how to develop ILs for further identification of new elite alleles, and how to utilize IL platforms for rice genetic improvement.

Keywords: introgression line, QTL detection, genetic interaction, elite allele identification, breeding by design

INTRODUCTION

Introgression lines (ILs) are a set of lines constructed by using a combination of continuous backcrossing and selfing to replace chromosome fragments of the recipient parent (RP) with chromosome fragments of the donor parent (DP), which can directly transfer the desired traits of interest from exotic varieties to adapted varieties. IL sets containing overlapping introgression segments covering the entire genome of DP in the RP background are referred to as chromosome segment substitution lines (CSSLs).

The number of backcrosses, number of backcrossed lines in each generation and number of selfings vary in different rice ILs construction programs, depending on the purpose of ILs construction, crossing methodology, genetic distance between parents, compatibility of genotypes, and number or size of target chromosome segment introgressions (Cavanagh et al., 2008; Balakrishnan et al., 2019). The backcrossed progenies can be selected either by phenotyping or by genotyping during the ILs construction, and then are continuously backcrossed with the RP to produce ILs. When screening progenies of backcrosses based on phenotype, if the trait of interest is controlled by a recessive gene, selfing is required before each backcross to identify plants with the recessive allele of interest (Vogel, 2009). With the development of rice functional genomics and molecular markers, marker-assisted selection (MAS) provides a more direct and effective way to obtain the desired rice ILs (Agarwal et al., 2008), and a series of IL/CSSL libraries have been constructed for QTL identification, gene cloning and variety improvement (Zhang et al., 2021c). The rapid development of high-throughput and high-density SNP genotyping technologies has facilitated the precise identification of introgression segments in ILs (Michael, 2014).

ILs are not only powerful prebreeding tools for broadening the genetic base of existing cultivars, but also valuable resources for quantitative trait loci (QTLs) mapping, gene effects evaluation, favorable alleles identification and genetic interactions analysis. The genetic background noise of ILs with a few chromosome-segment substitutions is low, so ILs with significant differences from the recurrent parents can be rapidly identified by evaluating the phenotype of target trait. Although the construction of ILs is time-consuming and labor-intensive, ILs have been widely reported in cereals, oil crops, vegetables, and commercial fiber crops due to their advantages (Balakrishnan et al., 2019). Harlan and Pope (1922) first used the backcrossing to transfer the smooth awn trait to an elite barley cultivar a century ago. Since then, backcross breeding has gradually developed into a widely used crop breeding method, which is often used to enhance crop resistance to disease and insects to ensure high and stable yield. Rice breeding has undergone two important revolutions including semidwarf breeding and heterosis utilization, both of which are based on gene introgression (Rao et al., 2014; Wu et al., 2018).

Cultivated rice, especially Asian rice, mainly consists of two subspecies, namely *indica* and *japonica* (Kovach et al., 2007; Wang et al., 2018). The two subspecies are further classified into 9 subpopulations with substantial genetic divergence and geographical distribution differences, including XI-1A, XI-1B, XI-2, XI-3, GJ-trp, GJ-sbtrp, GJ-tmp, cA, and cB (Wang et al., 2018). Inter- and intra-specific genetic diversity in mapping and breeding populations is fundamental. To date, a large number of rice ILs have been developed with distant hybridizations or intraspecific crosses (Balakrishnan et al., 2019; Zhang, 2021). The characteristics and construction processes of ILs/CSSLs were well described in previous reviews (Cavanagh et al., 2008; Ali et al., 2010; Balakrishnan et al., 2019). In this paper, we discussed the importance of parental selection during ILs construction, reviewed the recent achievements from rice ILs in functional genomics and molecular breeding, and proposed how to develop ILs for further exploring novel alleles and facilitating genetic improvement of rice.

SELECTION OF RECURRENT AND DONOR PARENTS FOR INTROGRESSION LINES CONSTRUCTION

The selection of parents is important for ILs construction because genetic differences between parents are the basis of breeding and functional genomics studies using ILs. A traditional backcross scheme for ILs construction with MAS is showed in **Figure 1A**. Rice varieties with excellent comprehensive traits but defects in one or few traits are often used as RPs for ILs construction with the goal of breeding improvement, and varieties that excel in these traits are selected as DPs. To date, the excellent *indica* rice varieties, including Huajingxian 74, Zhenshan 97B, 93-11, IR24, and IR64, and the elite *japonica* rice varieties, including Nipponbare, Koshihikari and Taichung 65, have been used as RPs (**Table 1**). Of course, these varieties are also used as donors in ILs construction. With the reference genomes of parents, ILs/CSSLs are ideal materials with which to comprehensively dissect the genetic basis of agronomic traits, and any genomic variations that cause phenotypic changes can be verified by these materials. Several sets of CSSLs have been constructed using parents with reference genomes (**Table 1**), such as CSSLs derived from the crosses between Zhenshan 97B and Nipponbare (Li et al., 2011), between Zhenshan 97B and Minghui 63 (Shen and Xing, 2014), and between 93 and 11 and Nipponbare (Zhu et al., 2009; Yuan et al., 2019).

Wild species of rice contain many useful allelic variations or haplotypes for improving cultivated rice yield and resistance to biotic and abiotic stresses (Kovach et al., 2007). Using wild rice as donors can significantly increase the genetic diversity of cultivars. Although a number of genes from wild species, especially biotic stress resistance genes, have been used in various breeding programs, most of the elite allelic variations in wild rice remains to be exploited (Ali et al., 2010; Li and Zheng, 2013). The genus *Oryza* is composed of 2 domesticated (*O. sativa* and *O. glaberrima*) and 22 wild species (Sanchez et al., 2013). All of six AA-genome wild species and 5 distantly related species with CC, BBCC, CCDD, EE, and FF genomes have been successfully used to construct ILs (Ali et al., 2010; Sanchez et al., 2013; Zhang et al., 2022). To date, more than twenty sets of ILs derived from crosses between wild relatives and cultivated rice have been developed (Ali et al., 2010; Balakrishnan et al., 2019), and some public resource platforms, such as CSSLs created with introgression from four wild rice into the *japonica* rice variety Taichung 65 (Yamagata et al., 2019) and ILs developed from crosses between two varieties (IR64 and Cybonnet) and three wild accessions (Singh et al., 2020), have been initially established.

Although hybrids between *indica* and *japonica* subspecies commonly exhibit varying degrees of hybrid sterility and hybrid breakdown, many researchers have focused on constructing

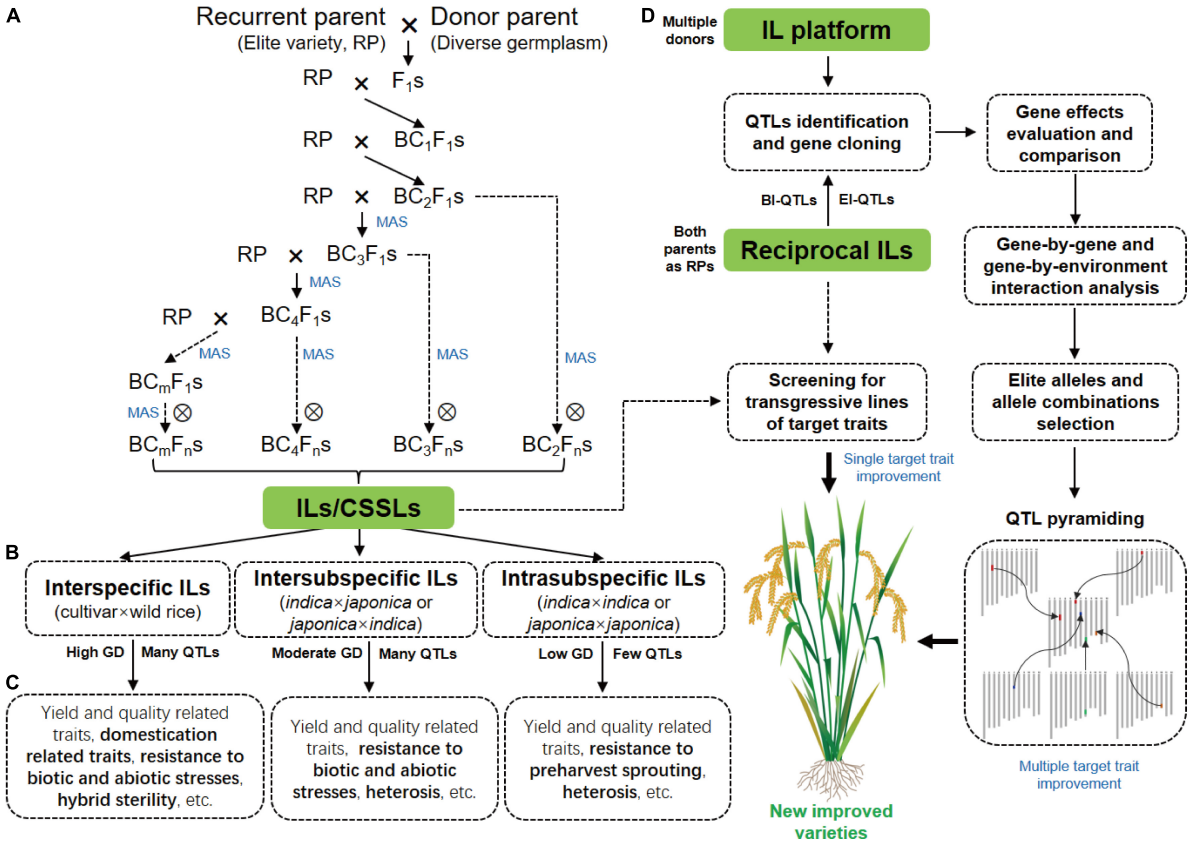

FIGURE 1 | Construction of rice ILs and their application in functional genomics research and breeding. **(A)** A traditional backcross scheme for ILs/CSSLs development with MAS. **(B)** ILs with different crossing types. **(C)** The main traits for genetic dissection by ILs derived from different crosses. The specific traits for the genetic dissection by different ILs are highlighted in bold. **(D)** The strategy of rice breeding by using ILs or IL platform. GD, genetic diversity; BI-QTLs, background-independent QTLs; E-QTLs, epistatic QTLs.

ILs by using *indica-japonica* crosses to perform genetic dissection and improvement of important agronomic traits due to the large degree of genetic divergence between the two subspecies (Balakrishnan et al., 2019). *Japonica* accessions experience harsher high-altitude and/or high-latitude environments, compared with the less harsh but more diverse environments experienced by *indica* rice, resulting in a great difference in agronomic traits between the two subspecies (Wang et al., 2018). Some favorable alleles introgressed between subpopulations, such as the long-grain haplotype of *GS3* (Fan et al., 2006), the wide-grain haplotype of *qSW5/GW5* (Shomura et al., 2008; Weng et al., 2008), the low-amylose haplotype of *Wx* (Tian et al., 2009) and the haplotype of *OsNRT1.1B* with high nutrient use efficiency (Hu et al., 2015), are of great significance for the improvement of rice agronomic traits.

Rice subpopulations originating in specific regions, such as cA and cB, have a larger number of "private" alleles than other subpopulations (Wang et al., 2018). Subpopulation cA encompasses the Aus, Boro and Rayada ecotypes from Bangladesh and India and subpopulation cB comprises the famous Basmati and Sadri aromatic varieties (Wang et al., 2018). The two subpopulations show many elite characteristics, such as the early maturity and drought and heat resistance of Aus varieties and superb fragrance and good cooking quality of aromatic varieties (Civáò et al., 2015). Several Aus varieties, such as Nagina 22 (Lin et al., 2015) and Chuan 7 (Zhang et al., 2021a), have been used as donors to construct ILs, while the aromatic varieties preferred for consumption, such as Khao Dawk Mali 105 (Nounjan et al., 2016; Chutimanukul et al., 2018), Basmati (Grover et al., 2020; Yadav et al., 2020), and Mushk Budji (Khan et al., 2018), are usually used as RPs.

In addition, African cultivated varieties (*Oryza glaberrima*) with an independent origin of domestication are also a valuable exotic gene pool for Asian cultivated rice and have been used as DPs for IL development (Li et al., 2015). The interspecific upland New Rice for Africa (NERICA) varieties have been developed by introgression of chromosome segments from *O. glaberrima* (CG 14) into three *japonica* rice varieties. Blending traits from the African traditional and Asian varieties, NERICA varieties have superior agronomic characteristics such as better weed resistance, resilience against major African biotic and abiotic stresses, high fertilizer returns, and high yields (Jones et al., 1997; Kijima et al., 2006; Fukuta et al., 2012).

ILs constructed by using interspecific and intersubspecific crosses can significantly expand the gene pools of RPs and the large genetic differences between parents also increase

TABLE 1 | List of introgression lines developed with several elite cultivars as recipient parents.

Recipient parent	Donor parent	Line	Generation	Trait*	References
Indica					
Huajingxian 74#	*Indica* (cv. IR64, CLSJM) and *japonica* (cv. Suyunuo, IRAT261, Lemont, IAPAR9)	86 SSSLs	BC_3F_2, BC_3F_3	Heading date, Plant height	He et al., 2005a,b,c
		217 SSSLs	BC_4F_2	Tiller and panicle number	Xi et al., 2006; Liu et al., 2008, 2009
	Indica (cv. Basmati385)	153 SSSLs	–	Grain width	Wang S. et al., 2012
	O. meridionalis (IRGC104093, 105286, 105293, 105291)	99 SSSLs	BC_4F_2 to BC_7F_2	10 Yield-related traits	He et al., 2017
Zhenshan 97B#	*Indica* (cv. Minghui 63#)	202 CSSLs	BC_4F_2	Heading date, plant height, heterosis	Shen and Xing, 2014; Shen et al., 2014
	Indica (cv. Pokkali)	172 ILs	BC_4F_2	Plant height	Kovi et al., 2011
	Japonica (cv. Nipponbare)	143 CSSLs	BC_4F_2	Grain size, seed dormancy, heterosis	Yuan S. et al., 2020; Wang et al., 2021; Xiong et al., 2021
	O. rufipogon (IRGC105491)	111 CSSLs	BC_2F_4 to BC_6F_3	Chlorophyll content	Gao et al., 2020
93-11#	*Indica* (cv. PA64s)	81 CSSLs	BC_4F_2 to BC_6F_2	11 Agronomic traits	Zhang et al., 2019b
		156 CSSLs	BC_3F_4 to BC_5F_3	8 Yield-related traits, heterosis	Liu et al., 2016; Lin et al., 2020
	Japonica (cv. Nipponbare)	103 CSSLs	BC_4F_2	Seed shattering, grain size	Zhu et al., 2009
		119 CSSLs	BC_4F_2	Tiller angle	Zhao et al., 2012
		122 CSSLs	BC_3F_2, BC_4F_3	Seed dormancy	Zhang et al., 2020a
		128 CSSLs	BC_4F_3 to BC_6F_2	Root traits, heterosis	Xu et al., 2010; Tao et al., 2016; Zhou et al., 2016
		55 CSSLs	BC_4F_2 to BC_6F_2	11 Agronomic traits	Zhang et al., 2019b
	Japonica (cv. C418)	108 CSSLs	BC_3F_8	Grain weight	Bian et al., 2010
	O. rufipogon (DP30)	132 CSSLs	BC_4F_2 to BC_6F_2	Plant architecture, cold tolerance	Yuan R. et al., 2020
	O. rufipogon (CWR276)	198 CSSLs	BC_4F_5 to BC_7F_4	10 Agronomic traits	Qiao et al., 2016
	O. rufipogon (CWR274)	133 CSSLs	BC_3F_4 to BC_5F_3	Grain size, grain weight	Qi et al., 2017
IR24	*Japonica* (cv. Asominori)	66 CSSLs	$BC_1F_{2\ (RILs \times IR24)}$	Ferrous iron toxicity	Wan et al., 2003
		70 CSSLs	$BC_2F_{2(RILs \times IR24)}$	Heading date, grain size	Kubo et al., 2002
	Oryza minuta (No. 101133)	131 ILs	BC_4F_6	10 Yield-related traits	Guo et al., 2013
IR64	*Japonica* (cv. Koshihikari)	39 CSSLs	–	28 Agronomic traits	Ujiie et al., 2016
	Japonica (IRGC23364)	26 CSSLs	BC_4F_4	Root angle	Uga et al., 2015
	Japonica (cv. Binam)	99 ILs	BC_2F_8	Salt tolerance	Zang et al., 2008
	10 Donor varieties	334 ILs	BC_3F_8	Spikelet number	Fujita et al., 2012
	O. rufipogon (IRGC105491)	105 ILs	BC_2F_5	7 Agronomic traits	Cheema et al., 2008
	O. rufipogon (W1944, IRGC106148, 105567)	218 CSSLs	BC_3F_3 to BC_6F_3	–	Singh et al., 2020
	O. glaberrima (RAM54, 90)	200 ILs	BC_2F_4	Root traits	Kijoji et al., 2014
Japonica					
Nipponbare#	*Indica* (cv. 93-11)	57 CSSLs	BC_4F_4, BC_5F_3	Grain weight, panicle architecture, amylose content	Zhang et al., 2011, 2019c; Su et al., 2021
	Indica (cv. PA64S)	61 CSSLs	BC_4F_2 to BC_6F_2	11 Agronomic traits	Zhang et al., 2019b
	Indica (cv. Kasalath)	54 CSSLs	–	Root system development	Suralta et al., 2008
	Japonica (cv. Koshihikari)	48 CSSLs	BC_4F_4	Pre-harvest sprouting	Hori et al., 2010
	O. rufipogon (C35)	104 CSSLs	BC_5F_6	Panicle-related traits	Ma et al., 2019
Koshihikari	*Indica* (cv. Habataki)	32 CSSLs	BC_3F_3 to BC_5F_4	Grain quality traits	Murata et al., 2014
	Indica (cv. Kasalath)	39 CSSLs	BC_1F_4, SBC_3F_3	15 Agronomic traits	Ebitani et al., 2005
	indica (cv. Takanari)	41 CSSLs	BC_4F_2 to BC_4F_4	6 Yield-related traits	Takai et al., 2014
	Indica (cv. IR64)	43 CSSLs	–	28 Agronomic traits	Ujiie et al., 2016
	Indica (cv. Nona Bokra)	44 CSSLs	BC_3F_3, BC_3F_4	Bacterial seedling rot	Takai et al., 2007; Mizobuchi et al., 2013
		154 CSSLs	BC_3F_2	Grain quality traits	Hao et al., 2009
	Japonica (cv. Nipponbare)	41 CSSLs	BC_4F_4	Pre-harvest sprouting	Hori et al., 2010
	Japonica (cv. Owarihatamochi)	44 CSSLs	BC_4F_5	Pre-harvest sprouting	Mizuno et al., 2018
	O. rufipogon (W0106)	33 CSSLs	BC_4F_2 to BC_7F_2	10 Agronomic traits	Furuta et al., 2014

(Continued)

TABLE 1 | (Continued)

Recipient parent	Donor parent	Line	Generation	Trait*	References
	O. glaberrima (IRGC104038)	34 CSSLs	BC_4F_7 to BC_7F_7	10 Yield-related traits	Shim et al., 2010
	O. barthii (W0009)	40 CSSLs	BC_4F_2 to BC_7F_2	10 Agronomic traits	Bessho-Uehara et al., 2017
Taichung 65	*Indica* (cv. ZYQ, DGWG)	10 SSSLs	–	21 Agronomic traits	Liu et al., 2004
	Indica (cv. DV85, ARC10313)	89 CSSLs	BC_3F_5	–	Yasui et al., 2010
	O. rufipogon (W1962)	44 CSSLs	BC_4F_4	–	Yamagata et al., 2019
	O. glaberrima (IRGC103777)	26 CSSLs	BC_4F_4	–	Yamagata et al., 2019
	O. nivara (IRGC 105715)	33 CSSLs	BC_4F_4	–	Yamagata et al., 2019
	O. meridionalis (W1625)	119 CSSLs	BC_4F_6	Hybrid breakdown	Munguambe et al., 2021
	O. longistaminata (IRGC110404)	40 CSSLs	BC_1F_7 to BC_4F_7	8 Yield-related traits	Ramos et al., 2016
	O. longistamina (W1508)	24 CSSLs	BC_3F_2	Anther length	Ogami et al., 2019

*Traits have been evaluated and/or genetically dissected in the corresponding IL populations.
#The reference genomes of these varieties have been released.
CLSJM, Chenglongshuijingmi; ZYQ, Zhaiyeqing; DGWG, Dee-Geo-Woo-Gen.

the available molecular markers for MAS, which is conducive to mining QTLs. However, the cross incompatibility and F_1 hybrid sterility increase the difficulties of interspecific and intersubspecific ILs/CSSLs construction (Ali et al., 2010). In contrast, ILs from intrasubspecific crosses are easier to construct but have lower genetic diversity. A number of intrasubspecific ILs have been developed by using elite *indica* or *japonica* varieties as the RPs (**Table 1**).

PROGRESS IN USING INTROGRESSION LINES IN RICE FUNCTIONAL GENOMICS RESEARCH

ILs are powerful tools with which to dissect complex traits into a set of monogenic loci, and are appropriate for detecting QTLs with both large and small effects (Xi et al., 2006). Over the past three decades, a large number of IL sets derived-from different crosses have been constructed for mining QTLs/genes for agronomic traits (**Figures 1B,C** and **Table 1**).

Genetic Dissection of Complex Traits by Introgression Lines With the Wild Relatives as Donor Parents

Wild species have acquired many elite alleles for resistance to biotic and abiotic stresses during the long period of natural selection (Atwell et al., 2014; Mammadov et al., 2018). A large number of IL libraries constructed from wild and cultivated rice were used to detect QTLs for cold (Yuan R. et al., 2020), heat (Lei et al., 2013; Prasanth et al., 2016; Cao et al., 2020), drought (Zhang et al., 2006; Zhou et al., 2006), and salt (Tian et al., 2011; Yang et al., 2012) tolerance. For example, a set of 132 CSSLs were developed from a cultivated rice, 93–11, and the Guangxi common wild rice (GXCWR), and the cold-tolerance QTL *qCT2.1* on chromosome 2 was identified using the secondary mapping population (Yuan R. et al., 2020). Using a set of 90 ILs derived from the cross between Teqing and Yuanjiang common wild rice (YJCWR), five QTLs related to heat response (*qHST1-1*, *qHST1-2*, *qHST2*, *qHST3,* and *qHST8*) in the seedling stage were detected, and *O. rufipogon*-derived alleles at one locus reduced sensitivity to heat (Lei et al., 2013). QTLs for resistance to biotic stress, such as that imposed by the green rice leafhopper (Thein et al., 2019), rice stripe necrosis virus (Gutiérrez et al., 2010) and rice blast fungus (Xu et al., 2015), were also mapped using ILs carrying wild rice donor fragments in cultivated rice backgrounds.

ILs developed from wild and cultivated rice are valuable materials for the genetic dissection of domestication-related traits such as an erect growth habit, inflorescence architecture, and awn length. Several domestication genes have been cloned using such IL-derived populations (**Table 2**). *PROG1*, a prostrate growth gene, was identified in a set of ILs developed from Teqing and YJCWR. Its non-functional allele determines the critical transition of rice from prostrate to erect growth (Tan et al., 2007, 2008). Using the same IL sets, a domestication gene, *OsLG1*, controlling inflorescence architecture was cloned (Zhu et al., 2013). Using a set of 354 ILs derived from the cross between 93 and 11 and YJCWR, an IL with long and barbed awns was identified. Further fine-mapping indicated that a frameshift deletion in *LABA1* of cultivated rice reduces the cytokinin concentration in awn primordia, disrupting barb formation and awn elongation (Fu et al., 2010; Hua et al., 2015). Another domestication gene controlling plant architecture, *TIG1*, was also cloned in an IL population with the wild rice (W2014) segments in the 93–11 background (Zhang et al., 2019d).

There has also been a burst of studies on mapping novel genomic regions and QTLs for yield-related traits (Furuta et al., 2014, 2016; Ma et al., 2016; Qiao et al., 2016; Rao et al., 2018) and quality traits (Garcia-Oliveira et al., 2009; Fasahat et al., 2012; Yun et al., 2016; Qi et al., 2017) by using IL populations derived from the crosses between wild and cultivated rice, generally with the elite varieties as RPs. For example, using 33 CSSLs of *O. rufipogon* (W0106) in the background of the elite *japonica* rice variety Koshihikari, a total of 15 major QTLs for eight yield-related traits were detected, and a novel QTL controlling the number of grains per panicle were identified on chromosome 10 (Furuta et al., 2014). In the same background of Koshihikari, a set of 26 CSSLs with segments from *O. nivara* were constructed, and new QTLs associated with yield-related traits were identified

TABLE 2 | List of genes with natural variations identified by ILs and cloned by IL-derived populations in rice.

Gene	Locus	Recipient parent	Donor parent	Trait	References
NOG1	LOC_Os01g54860	*Indica* (cv. Guichao 2)	*O. rufipogon* (DXCWR)	Grain number	Huo et al., 2017
OsGUX1	LOC_Os01g65780	*Indica* (cv. Zhenshan 97)	*O. rufipogon* (ACC10)	Chlorophyll content	Gao et al., 2020
OsHMA4	LOC_Os02g10290	*Japonica* (cv. Lemont)	*Indica* (cv. Teqing)	Copper accumulation	Huang et al., 2016
CAL1	LOC_Os02g41904	*Japonica* (cv. Chunjiang06)	*Indica* (cv. Tainan1)	Cadmium accumulation	Luo et al., 2018
DTH3	LOC_Os03g03070	*Indica* (cv. Huajingxian74)	Multiple varieties	Heading date	Zhu et al., 2018
BOC1	LOC_Os03g12820	*Indica* (cv. Teqing)	*O. rufipogon* (YJWCR)	Callus browning	Zhang et al., 2020b
TT1	LOC_Os03g26970	*Japonica* (cv. Wuyunjing)	*O. glaberrima* (CG14)	Thermotolerance	Li et al., 2015
GNP1	LOC_Os03g63970	*Japonica* (cv. Lemont)	*indica* (cv. Teqing)	Grain number	Wu et al., 2016
CTB4a	LOC_Os04g04330	*Japonica* (cv. Towada)	*japonica* (cv. KMXBG)	Cold tolerance	Zhang et al., 2017
An-1	LOC_Os04g28280	*Indica* (cv. Guangluai4)	*O. rufipogon* (W1943)	Awn length; Grain size	Luo et al., 2013
LABA1/An-2	LOC_Os04g43840	*Indica* (cv. 93-11)	*O. rufipogon* (YJWCR)	Awn length	Hua et al., 2015
GW5	LOC_Os05g09520	*Japonica* (cv. Asominori)	*indica* (cv. IR24)	Grain width	Weng et al., 2008
OsEBS	LOC_Os05g51360	*Indica* (cv. Guichao2)	*O. rufipogon* (DXCWR)	Plant biomass; Spikelet number	Dong et al., 2013
Hd17	LOC_Os06g05060	*Japonica* (cv. Nipponbare)	*Japonica* (cv. Koshihikari)	Heading date	Matsubara et al., 2012
GW6	LOC_Os06g15620	*Indica* (cv. Huajingxian74)	*Japonica* (cv. Nanyangzhan)	Grain width	Shi et al., 2020
PROG1	LOC_Os07g05900	*Indica* (cv. Teqing)	*O. rufipogon* (YJWCR)	Prostrate growth	Tan et al., 2008
OsHMA3	LOC_Os07g12900	*Indica* (cv. Huajingxian74)	*Indica* (cv. BG367)	Cadmium accumulation	Sui et al., 2019
COS1/FZP	LOC_Os07g47330	*Japonica* (cv. C418)	*O. rufipogon* (DXCWR)	Inflorescence branching	Huang et al., 2018
Hd18	LOC_Os08g04780	*Japonica* (cv. Koshihikari)	*japonica* (cv. Hayamasari)	Heading date	Shibaya et al., 2016
TIG1	LOC_Os08g33530	*Indica* (cv. 93-11)	*O. rufipogon* (W2014)	Tiller angle	Zhang et al., 2019d
GAD1	LOC_Os08g37890	*Indica* (cv. NA93-11)	*O. rufipogon* (W2014)	Awn length; Grain size	Jin et al., 2016
OsSPL16/GW8	LOC_Os08g41940	*Indica* (cv. Huajingxian74)	*indica* (Basmati385)	Grain width	Wang S. et al., 2012
GS9	LOC_Os09g27590	*Japonica* (cv. Nipponbare)	Qingluzan11	Grain size	Zhao et al., 2018
DTE9/OsMADS8	LOC_Os09g32948	*Japonica* (cv. Hwayeong)	*O. rufipogon* (W1944)	Hybrid weakness	Kim et al., 2021
TAC1	LOC_Os09g35980	*Indica* (cv. IR24)	*japonica* (cv. Asominori)	Tiller angle	Yu et al., 2007
OsGluA2	LOC_Os10g26060	*Japonica* (cv. Sasanishiki)	*Indica* (cv. Habataki)	Grain protein content	Yang et al., 2019
OsLTPL159	LOC_Os10g36160	*Indica* (cv. Guichao2)	*O. rufipogon* (DXCWR)	Cold tolerance	Zhao et al., 2020
GL10/OsMADS56	LOC_Os10g39130	*Indica* (cv. Huajingxian74)	*Japonica* (cv. Lemont)	Grain length	Zhan et al., 2022
NRT1.1B	LOC_Os10g40600	*Japonica* (cv. Nipponbare)	*Indica* (cv. IR24)	Nitrogen-use efficiency	Hu et al., 2015
OsGH3.13	LOC_Os11g32520	*Indica* (cv. Zhenshan 97)	*Japonica* (cv. Nipponbare)	Grain length	Wang et al., 2021
TOND1	LOC_Os12g43440	*Indica* (cv. Teqing)	*O. rufipogon* (YJWCR)	Tolerance to nitrogen deficiency	Zhang et al., 2015

DXCWR, Dongxiang common wild rice; YJWCR, Yuanjiang common wild rice; KMXBG, Kunmingxiaobaigu.

(Furuta et al., 2016). A set of 198 CSSLs were developed by introgressing *O. rufipogon* segments into the background of the elite *indica* rice variety 93–11, and a new QTL associated with the heading date was detected in a 78-kb region on chromosome 10 (Qiao et al., 2016). Similarly, a set of 131 ILs were developed by introducing *O. nivara* segments into the 93–11 background, and 65 QTLs for 13 yield-related traits were detected by bin-map, with ∼36.9% of the alleles at the detected QTLs from *O. nivara* leading to improved yield-associated traits (Ma et al., 2016). QTLs associated with 12 grain quality traits were detected using 96 ILs developed from an interspecific cross between the Korean elite *japonica* rice variety Hwaseong and *O. rufipogon* (IRGC 105491). A total of 48 QTLs for these traits were identified, and most wild alleles of these detected QTLs had negative effects on the traits (Yun et al., 2016).

In addition to the traits mentioned above, IL sets developed from wild and cultivated rice have also been used for genetic dissection of other traits, such as interspecific hybrid sterility (Yang et al., 2016; Li et al., 2018a; Zhang et al., 2018), anther length (Ogami et al., 2019), the stigma exsertion rate (Tan et al., 2020), out-crossing rate (Prahalada et al., 2021), and photosynthetic efficiency (Rao et al., 2018).

Genetic Dissection of Complex Traits by Using Introgression Lines Derived From Intersubspecific Crosses

Multiple IL sets derived from *indica-japonica* crosses have been developed to mine QTLs for rice yield-related traits (Ando et al., 2008; Sun et al., 2017; Mitsuya et al., 2019), quality-related traits (Hao et al., 2009; Sun et al., 2015; Wang et al., 2021), and resistance to biotic (Park et al., 2007) and abiotic (Zang et al., 2008; Zhao et al., 2016; Feng et al., 2018) stresses. Several QTLs for these traits have been detected and causal natural variations for the detected QTLs have been identified by using IL-derived F_2 populations (**Table 2**). For example, two sets of reciprocal ILs derived from a *japonica* rice variety Lemont and an *indica* variety Teqing were constructed, and multiple QTLs for grain number per panicle (GNP) were identified form these two ILs. Among these detected QTLs, a stable QTL *GNP1* was further finely mapped to a 33.7-kb region and a GA biosynthesis gene *GA20-oxidase 1* was identified as the candidate gene for *GNP1* (Wu et al., 2016). QTL analysis of seven panicle- and grain-related traits was performed in a set of genotype-defined CSSLs derived from a cross of two genome-sequenced varieties, Nipponbare

and Zhenshan 97, and a total of 43 QTLs for these traits were identified. Furthermore, the novel locus *qGL11* for grain length and thousand-grain weight was finely mapped to a 25-kb region, and the IAA-amido synthetase gene *OsGH3.13* was verified as the causal gene for *qGL11* (Wang et al., 2021).

Heterosis is one of the most important characteristics of F_1 plants derived from *indica-japonica* crosses (Dan et al., 2014). Many ILs and IL-derived backcross and/or testcross populations have been used to dissect the genetic basis of heterosis (Xin et al., 2011; Wang Z. et al., 2012; Tao et al., 2016; Xiong et al., 2021; Yang et al., 2021c). For example, a set of 70 ILs and corresponding testcross F_1 populations were developed from a cross between the *indica* rice variety IR24 (RP) and *japonica* rice variety Asominori to investigate heterotic loci (HLs) associated with six yield-related traits (Xin et al., 2011). A total of 41 HLs were detected on the basis of mid-parent heterosis values with single-point analysis, and most of QTLs were overdominant, suggesting that heterotic effects at the single-locus level are mainly overdominant in rice (Xin et al., 2011). Tao et al. (2016) constructed a set of 128 CSSLs derived from a cross between the *indica* rice variety 93–11 and *japonica* rice variety Nipponbare to investigate the genetic basis of heterosis. Using the testcross populations derived from CSSLs and three P-TGMS lines, a total of 97 HLs associated with yield components were detected. Genetic analysis revealed that the contribution of different genetic effects to heterosis might vary among traits (Tao et al., 2016; Yang et al., 2021c). In a recent study, *Ghd8* was identified and verified as a major HL for yield components by using a set of intersubspecific CSSLs and two test populations (Xiong et al., 2021). This HL has also been identified in the two-line rice hybrid system (Li et al., 2016), indicating its important role in hybrid breeding.

QTLs have been mapped in ILs derived from *indica-japonica* crosses for other traits, including hybrid sterility (Zhao et al., 2011), nutrient utilization (Kabange et al., 2021), seed germination and dormancy (Li et al., 2011; Zhang et al., 2020a), mature seed culturability (Zhao et al., 2009), carbon isotope discrimination (Takai et al., 2009), lodging resistance (Mulsanti et al., 2018), and cell wall characteristics (Xu et al., 2017).

Genetic Dissection of Complex Traits by Using Introgression Lines Derived From Intrasubspecific Crosses

ILs developed from intrasubspecific crosses are favored by breeders because there are no incompatibility barriers in these crosses. Although the genetic diversity within rice subspecies is lower than that between subspecies, there are still large phenotypic differences among varieties from different regions with different genealogical relationships in the same subspecies. Several IL libraries from intrasubspecific crosses have been developed in rice to genetically dissect complex traits such as yield-related traits (Shen and Xing, 2014; Hori et al., 2021a; Kato and Hirayama, 2021), quality-related traits (Yang et al., 2021b), root traits (Anis et al., 2019), resistance to preharvest sprouting (Hori et al., 2010; Mizuno et al., 2018), and heterosis (Lin et al., 2020). For example, a set of 202 CSSLs of an elite hybrid, Shanyou 63, were developed in the Zhenshan 97 background, and QTLs for heading date and plant height were detected (Shen and Xing, 2014; Shen et al., 2014). A total of 15 partial dominance QTLs for plant height were identified in 15 CSSL-derived F_2 populations, and these QTLs were further identified as HLs for plant height that acted dominantly and epistatically (Shen et al., 2014). Furthermore, multiple QTLs and HLs for yield and spikelets per panicle were identified in these CSSLs, and the hybrids pyramiding these detected HLs in the combined parental genome background showed yield performance similar to that of Shanyou 63, indicating that heterosis might be successfully achieved by manipulating several major dominant HLs (Shen et al., 2022).

Identification of Elite Alleles by Ranking the Effects of Different Alleles With Introgression Lines

ILs derived from advanced backcrosses generally carry only one or a few chromosome fragments from the donor, which greatly minimizes the genetic background noise and facilitates the evaluations of the genetic effects of QTLs/genes on corresponding traits. The target introgressed segment of an IL is homozygous, conferring its genomic stability, so ILs can be repeatedly planted in multiple sites to evaluate the genetic effects of QTLs across different environments (Zhou et al., 2017). A set of ILs covering the entire DP genome can be applied to identify multiple QTLs associated with the trait of interest, and the genetic effects of these detected QTLs can be compared based on their additive effect in the same background, especially using single segment substitution lines (SSSLs) (Ye et al., 2010; Zhu et al., 2018). ILs are not only beneficial for comparing the genetic effects of different QTLs in the same background but also beneficial for comparing the effects of different alleles at one QTL in the fixed background, so as to identify the favorable alleles or allele combinations. For example, the heading date QTL *qHD-3* was identified by using a library consisting of 1123 CSSLs from multiple donors in the same genetic background, Huajingxian 74 (HJX74). Compared with the HJX74 allele, the *qHD-3* allele from donor variety Lemont delayed rice heading, while *qHD-3* from another donor parent, IR64, promoted rice heading. Further sequencing analysis revealed that different variants of *DTH3* were identified between Lemont and IR64 (Zhu et al., 2018). To investigate the interaction effect of *Ghd8* on yield mid-parent heterosis (MPH), Xiong et al. (2021) constructed five near-isogenic lines (NILs) with each carrying an introgression segment covering *Ghd8* from a particular donor in the same background of ZS97, and identified the combination ($Ghd8^{ACC10}$/$Ghd8^{MH63}$) with the highest MPH for both spikelet number and grain yield per plant using a half-diallel mating design.

Identification of Background-Independent and Epistatic Quantitative Trait Loci Using Reciprocal Introgression Lines

Reciprocal ILs developed in both parental backgrounds have the advantage of enabling evaluation of differences in allelic effects of QTLs in both genetic backgrounds (Kubo et al., 2002).

Reciprocal ILs are not only conducive to the detection and fine mapping of QTLs that have both large and small effects but also appropriate for evaluating gene activity as a single factor or in epistatic interactions (Takai et al., 2014). Reciprocal ILs have been used to identify background-independent QTLs (BI-QTLs) and epistatic QTLs (E-QTLs) for grain yield- and quality-related traits (Takai et al., 2014; Qiu et al., 2017; Hori et al., 2021b), salt tolerance (Cheng et al., 2012), lodging resistance (Ookawa et al., 2016), the leaf net photosynthetic rate (Adachi et al., 2019), and sink- and source-related traits (Wang et al., 2020). For example, using the reciprocal CSSLs derived from a cross between MH63 and 02428, a total of nine BI-QTLs for appearance quality were identified. Thirteen and 10 stably expressed QTLs (SE-QTLs) were also detected in the MH63 and 02428 backgrounds, respectively (Qiu et al., 2017). Compared with E-QTLs, BI-QTLs, and SE-QTLs are more easily applied for rice genetic improvement because their function does not depend on background or the environment.

Genetic Interaction Analysis Conducted by Using Introgression Lines and Introgression Line-Derived Populations

Most agronomic traits of crops are complex traits controlled by multiple genes and affected by environmental factors. Understanding the genetic and environmental bases of QTL × QTL and gene-by-environment (G × E) interactions is of fundamental importance in plant breeding (Monteverde et al., 2019). As in recombinant inbred lines (RILs), genetic interactions between QTLs/genes, such as additive-by-dominance and dominance-by-dominance interactions, also cannot be analyzed using ILs because all sites in ILs are homozygous. However, the offspring-separated populations derived from the crosses between ILs or backcrosses between ILs and RPs are ideal populations for genetic interaction analysis due to their low genetic background noise. Genetic interaction analysis among four major rice heading date genes, *Ghd7*, *Ghd8*, *OsPRR37,* and *Hd1*, was performed in a 4-gene segregating population in the near-isogenic background under both natural long-day (NLD) and natural short-day (NSD) conditions (Zhang et al., 2019a). Tetragenic, trigenic and digenic interactions among these four genes for heading date were observed under both conditions but were more significant under NLD conditions. Further analysis showed that the differences in digenic interactions between *Ghd7* and *OsPRR37* under different conditions were essential to the alternative function of *OsPRR37* (Zhang et al., 2019a). Genetic interaction analysis between QTLs for other agronomic traits, such as plant height (Shen et al., 2014), grain size (Xia et al., 2018), tillering (Zhou et al., 2020), cold tolerance (Liang et al., 2018), and drought tolerance (Yadav et al., 2019), was also performed in IL-derived populations.

Most the traits with low heritability are easily affected by the environment, and there is an obvious interaction between gene and the environment. Understanding G × E interactions can help breeders in deciding which QTL to use in their breeding programs while tailoring crop cultivars for specific or more diverse environments (Liu et al., 2008). G × E interactions were analyzed in ILs or IL-derived populations planted in a variety of environments. To dissect the genetic basis of the genetic main effect (G) and G × E interaction effect (GE) for panicle number (PN) in rice, a population consisting of 35 SSSLs derived from originating from crosses between the RP, HJX74, and 17 DPs was grown in six cropping season environments (Liu et al., 2008). The total genetic effect was partitioned into G and GE by using the mixed linear-model approach, and then QTL analyses of these effects were conducted separately. A single QTL effect was divided into two components: The additive effect (**a**) and additive × environment interaction effect (**ae**). A total of 18 QTLs for PN were identified, and three types of QTLs were suggested according to their effects expressed. Two QTLs expressed stably across environments due to the association with only **a**, nine QTLs with only **ae** were unstable, and the remaining seven QTLs were identified with both **a** and **ae**, which were also unstable across environments (Liu et al., 2008). Notably, when QTLs have large **ae** values in opposing directions in different environments, selection for an allele with favorable effects in certain environments could lead to undesired results in other environments.

INTROGRESSION LINES ARE VALUABLE RESOURCES FOR RICE BREEDING

Rice breeding is based on the use of valuable alleles and elite germplasm resources. Fortunately, rice has a large number of germplasms containing all kinds of excellent genes for high yield, good quality, resistance to biotic and abiotic stresses, and etc. (Tanksley and McCouch, 1997). Transferring these excellent genes into modern varieties absolutely could broaden the gene pool and greatly improve the productivity. However, many of these valuable alleles have linkage drags (Ali et al., 2010). IL populations have the advantage of breaking the linkage between valuable genes and linkage drags. To date, many traits in rice have been greatly improved by developing ILs in the backgrounds of modern varieties (Hu et al., 2016; Balakrishnan et al., 2019; Zhang, 2021; Zhang et al., 2021c).

Single Target Trait Improvement by Using Introgression Lines

Each line of ILs carries only one or a few donor chromosome segments, which usually affects only one or a few traits of the RP. Therefore, using ILs to improve single-target trait in rice is efficient. In the elite genetic backgrounds, IL with improved agronomic performance can be released as a new variety.

Increasing rice yield to meet the need for a rapidly growing global population is one of the most important goals of rice breeding. Many yield-related QTLs/genes have been introgressed into elite backgrounds to increase rice yield. For example, *GN4-1*, a QTL for grain number per panicle, was introgressed into ZH8006 from WYJ6, and grain yield increased by more than 10% (Zhou et al., 2018). To achieve the high yield potential in Kongyu 131, a minute chromosome fragment carrying the favorable *Gn1a* allele from the donor parent GKBR was introgressed into

the genome of Kongyu 131, which resulted in larger panicles and a subsequent yield increase in the new Kongyu 131 (Feng et al., 2017). In addition, introgression of a functional epigenetic *OsSPL14*^{WFP} allele into elite *indica* cultivars can greatly improve panicle traits and grain yield (Kim et al., 2018).

With the increasing development of people's living standards, rice quality has received attention from both rice producers and consumers. There have been many successful cases of improving rice grain quality through ILs. For example, to improve the eating and cooking quality of Zhenshan 97 and its hybrid Shanyou 63, the *Wx/Waxy* gene region from Minghui 63, a restorer line with medium amylose content (AC), soft gel consistency (GC), and high gelatinization temperature (GT), was introgressed into Zhenshan 97B and then transferred into Zhenshan 97A (Zhou et al., 2003). In recent years, the discovery of different alleles of *Wx* and *ALK* has provided more choices for rice grain quality improvement via gene introgression (Chen et al., 2020; Huang et al., 2021; Zhang et al., 2021b; Zhou et al., 2021). *GS9* is an important gene for both grain shape and chalkiness, and appearance quality could be largely improved by developing an IL for this gene (Zhao et al., 2018, 2021).

Biotic and abiotic stresses are large threats to rice production, leading to serious yield losses every year. ILs constructed by introgressing novel genes into modern varieties are an effective tool for solving this problem. For example, by combining a backcross breeding strategy combined with MAS, a total of 12 blast resistance genes were introgressed into one maintainer line of cytoplasmic male sterility (CMS) and three photo-thermo genetic male sterility (P-TGMS) lines. The blast resistance of these lines was significantly improved, while other traits were not changed (Jiang et al., 2012, 2015, 2019). Using the same strategy, brown planthopper resistance has been improved by introgressing *Bph3*, *Bph14*, *Bph15*, *Bph18*, *Bph20*, *Bph21,* and *Bph33* into 93–11 and Jin23B (Hu et al., 2013, 2018; Jiang et al., 2018), and bacterial blight resistance has been improved by pyramiding *Xa4*, *xa5*, *xa13,* and *Xa21* (Huang et al., 1997). In the two IL populations developed with C258 and ZGX1 as RPs and IR75861 as the DP, 12 and 8 lines had higher salt tolerance than the RPs, respectively, and they could be directly planted into environments characterized by salt stress (Qiu et al., 2015).

The excess application of fertilizers in rice production not only increases cost, but also causes severe environmental problems (Zhang et al., 2020c). Breeding rice varieties with high nitrogen-use efficiency (NUE) is of great significance to the practice of Green Super Rice (Yu et al., 2020). GRF4 promotes and integrates nitrogen assimilation, carbon fixation, and growth, whereas DELLA inhibits these processes. Introgression of *GRF4* into the elite variety 93–11 carrying *sd1* can significantly improve its NUE, thereby increasing yield (Li et al., 2018b). *NAL1* and the gain-of-function mutation *dep1-1* have been reported to play important roles in N-responsive tillering regulation, and pyramiding of the *dep1-1* and *NAL1*^{NJ6} alleles can achieve sustainable improvements in NUE and grain yield in *japonica* rice breeding (Sun et al., 2014; Xu et al., 2019). The introgression of favorable alleles of other NUE-related genes such as *OsTCP19* and *OsSBM1* also showed great potential for genetic improvement of high NUE in rice (Liu et al., 2021; Xu et al., 2021).

Multiple Target Trait Improvement by Using Introgression Lines

To meet people's diverse demands for rice, modern rice varieties should be improved mainly by multiple locus introgression to improve the overall features of rice in terms of grain yield and quality, resistance to biotic and abiotic stresses and nitrogen use efficiency, etc. For example, Luan et al. (2019) used four SSSLs with HJX74 as the RP to develop a new maintainer named H131B with a suitable heading date, good appearance and eating and cooking quality, and fragrant smell by pyramiding *MADS50*, *gs3*, *fgr*, *Wx*^{gl}, and *ALK*. Dixit et al. (2020) developed seven ILs possessing a combination of seven to ten QTLs for resistance to different biotic and abiotic stresses by using the MAS breeding method in the background of Swarna. These ILs were superior to the respective RPs in terms of agronomic performance and possessed superior grain quality. Luo et al. (2016) successfully developed a new restorer line, WH6725, with disease resistance to rice blast and bacterial blight, tolerance to submergence and an aromatic fragrance by pyramiding *Xa27*, *Pi9*, *Sub1A,* and *badh2.1*. Through genotyping-based identification and backcrossing with MAS, Zeng et al. (2017) introgressed 21 favorable alleles from NIP and 9311 into Teqing and then developed three lines that exhibited higher yield potential and better grain quality than their parental varieties and the super-hybrid rice Liang-you-pei-jiu. Guangzhan 63-4S is an elite two-line thermos-sensitive genic male sterile line. However, it is highly susceptible to blast and bacterial blight. Two BC_2 ILs were developed to introgress *Xa7* for bacterial blight and *Pi-2* for blast resistance. Then, the ILs were crossed and pyramided with these two genes. A new two-line thermosensitive genic male-sterile line named Hua1228S that was resistant against both rice blast and bacterial blight was developed (Mi et al., 2018).

Rice Breeding Improvement by Using Introgression Line Platform

IL platform developed by multiple donors in an elite genetic background facilitates the discovery of new QTLs and identification of multiple alleles of a gene. IL platform can be used to identify the optimal allele or allele combination by evaluating their genetic effects in the fixed background and then improve varieties by QTL pyramiding (Zhang, 2021; **Figure 1D**). In addition, new varieties can be directly developed by screening transgressive lines of target traits in the IL platform (**Figure 1D**). HJX74 is an elite variety from South China. A single segment IL platform including more than 2000 ILs was developed with HJX74 as the RP and more than 40 varieties as DPs (Luan et al., 2019; Zhang, 2021). A large number of important genes were identified using this platform (Wang S. et al., 2012; Wang et al., 2015; Yang et al., 2021a). In addition, a series of varieties have been bred by pyramiding different favorable genes in this platform, such as three-line maintainer lines (H121B, H131B, HZB, HBB, E5-HXB, and E5-HBB) (Luan et al., 2019), black rice (Huaxiaohei1 and Huaxiaohei 2), red rice (Huaxiaohong1 and Huaxiaohong 2), and high-quality rice (Huabiao1 and Huabiao 3) (Zhang, 2021). Huanghuazhan (HHZ) is an elite rice variety widely planted in the middle and downstream regions of the

Yangtze River and in South China. An IL platform including 496 ILs were developed using HHZ as the RP and 8 varieties as DPs (Ali et al., 2017). Most ILs showed significantly higher yields than the parents under both abiotic stress conditions (salt and drought) and non-abiotic stress conditions (submergence). Ultimately, at least six ILs were directly released as new green super rice varieties in the Philippines and Pakistan and have been grown on more than 1 million ha of land in different ecosystems (Ali et al., 2017).

CONCLUSION AND PERSPECTIVES

ILs simplify the genetic dissection of complex traits and accelerate whole-genome large-scale gene discovery in rice. To date, a large number of the natural variations associated with agronomic traits in rice have been identified from ILs and then cloned using IL-derived mapping populations (Li et al., 2018c; Zhang, 2021; **Table 2**). The primary mapping populations may be more cost-effective for large-effect QTLs mapping, but ILs have a greater advantage in detecting minor QTLs because of their low genetic background noise. The resolution of QTL mapping using ILs is low due to the large size of the introgression segments and the selective abandonment of some recombination events (lack of lines) during the construction of ILs (Cavanagh et al., 2008; Balakrishnan et al., 2019). Substitution mapping of QTLs using multiple ILs is an effective way to improve the mapping resolution (Zhang, 2021). In addition, any two ILs carrying different QTLs/genes controlling the same trait can be used for the epistatic interaction detection with their offspring-segregated population, which can provide useful information on how to select favorable alleles and allele combinations in rice molecular design breeding.

A large number of genetic dissections of traditional agronomic traits have been performed in rice (Li et al., 2018c; Wei et al., 2021), so it is difficult to further mine new QTLs/genes for these traits. While some rice subpopulations or distant relatives, such as cA, cB, and Africa cultivated rice, with narrow geographic or relatively independent origins, may carry many private (rare) genes and novel natural variations due to the differences in domestication and environmental pressures (Li et al., 2015; Bai et al., 2017; Wang et al., 2018). Accessions from these subpopulations were rarely used in the past, but are encouraged used as donors for constructing ILs in the future because more new QTLs might be detected using this kind of ILs. Many ILs have been constructed using AA-genome wild species as donors, but there are few successful cases of constructing ILs using other wild *Oryza* species as donors due to the strong incompatibility barriers (Sanchez et al., 2013). These wild species not only have strong biotic and abiotic stress resistance, but also have unique and excellent traits, such as the early morning flowering trait in CC-genome wild rice *Oryza officinalis* and the high biomass production in CCDD-genome wild species (Sanchez et al., 2013; Hirabayashi et al., 2015). Thus, using more different wild rice species as genetic resources to broaden the gene pool of cultivars by constructing ILs is of great significance, but it is also full of challenges. Moreover, multiple donors from different wild/cultivated rice species and subpopulations are also encouraged to develop IL platforms in one genetic background for new alleles detection. To date, the third-generation sequencing technology makes the cost of genome assembly dramatically decreased, a large number of rice varieties have been deeply sequenced, and reference genomes of dozens of varieties (including wild and cultivated rice) have been assembled (Wang et al., 2018; Wing et al., 2018; Qin et al., 2021). Thus, the accessions used for ILs development are suggested to generate the genome reference to facilitate the discovery of genetic variations between the recurrent and donor parents. Based on the difference of gene annotations between parents, the ILs contained the target genes can be used to precisely examine the potential target traits and rapidly verify the valuable natural variations. In addition, ILs combined with high-throughput phenotyping platform are encouraged to detect QTLs for traits measured dynamically and non-destructively, especially for the root associated traits and abiotic stress tolerance, which are hard to precisely measure by traditional ways (Yang et al., 2014).

As important prebreeding materials, ILs facilitate the genetic improvement of rice via gene transfer and pyramiding. However, QTL pyramiding for rice variety improvement requires multiple crosses and the avoidance of genetic drag, which usually requires several generations and large populations (Zhang et al., 2021c). Applying the IL platform in an elite background can accurately reveal multiple favorable alleles that can be used to improve the RP. This platform can also be used to evaluate the genetic effects of different alleles at the same locus and the effects of different allele combinations to identify the optimal allele or combination for target genetic improvement. In addition, the excellent allele combinations existing in the background will be kept in the process of QTL pyramiding, so target genetic improvement can be realized quickly. Moreover, the IL platform provides a chance to rapidly generate multi-line varieties by mixing the ILs with different biotic stress resistant genes, such as those for rice blast and brown planthopper resistance (Koizumi et al., 2004; Li et al., 2013). In some elite backgrounds, multiple ILs have been constructed by different institutions (**Table 1**), and global sharing of these resources would accelerate the construction of IL platforms and further facilitate the consortium of rice genetic improvement.

AUTHOR CONTRIBUTIONS

YX provided the idea for the review. BZ and XQ collected the data and wrote the manuscript. LM assisted in literature collection. YX, XQ, BZ, LM, and BW revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the National Natural Science Foundation of China (Grant Nos. 32171993 and 32101740) and the China Postdoctoral Science Foundation (Grant No. 2021M701360).

REFERENCES

Adachi, S., Yamamoto, T., Nakae, T., Yamashita, M., Uchida, M., Karimata, R., et al. (2019). Genetic architecture of leaf photosynthesis in rice revealed by different types of reciprocal mapping populations. *J. Exp. Bot.* 70, 5131–5144. doi: 10.1093/jxb/erz303

Agarwal, M., Shrivastava, N., and Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Rep.* 27, 617–631. doi: 10.1007/s00299-008-0507-z

Ali, J., Xu, J. L., Gao, Y. M., Ma, X. F., Meng, L. J., Wang, Y., et al. (2017). Harnessing the hidden genetic diversity for improving multiple abiotic stress tolerance in rice (*Oryza sativa* L.). *PLoS One* 12:e0172515. doi: 10.1371/journal.pone.0172515

Ali, M. L., Sanchez, P. L., Yu, S.-B., Lorieux, M., and Eizenga, G. C. (2010). Chromosome segment substitution lines: a powerful tool for the introgression of valuable genes from *Oryza* wild species into cultivated rice (*O. sativa*). *Rice* 3, 218–234. doi: 10.1007/s12284-010-9058-3

Ando, T., Yamamoto, T., Shimizu, T., Ma, X. F., Shomura, A., Takeuchi, Y., et al. (2008). Genetic dissection and pyramiding of quantitative traits for panicle architecture by using chromosomal segment substitution lines in rice. *Theor. Appl. Genet.* 116, 881–890. doi: 10.1007/s00122-008-0722-6

Anis, G. B., Zhang, Y., Islam, A., Zhang, Y., Cao, Y., Wu, W., et al. (2019). *RDWN6*^{XB}, a major quantitative trait locus positively enhances root system architecture under nitrogen deficiency in rice. *BMC Plant Biol.* 19:12. doi: 10.1186/s12870-018-1620-y

Atwell, B. J., Wang, H., and Scafaro, A. P. (2014). Could abiotic stress tolerance in wild relatives of rice be used to improve *Oryza sativa*? *Plant Sci.* 21, 48–58. doi: 10.1016/j.plantsci.2013.10.007

Bai, X., Huang, Y., Hu, Y., Liu, H., 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

Balakrishnan, D., Surapaneni, M., Mesapogu, S., and Neelamraju, S. (2019). Development and use of chromosome segment substitution lines as a genetic resource for crop improvement. *Theor. Appl. Genet.* 132, 1–25. doi: 10.1007/s00122-018-3219-y

Bessho-Uehara, K., Furuta, T., Masuda, K., Yamada, S., Angeles-Shim, R. B., Ashikari, M., et al. (2017). Construction of rice chromosome segment substitution lines harboring *Oryza barthii* genome and evaluation of yield-related traits. *Breed. Sci.* 67, 408–415. doi: 10.1270/jsbbs.17022

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

Cao, Z., Li, Y., Tang, H., Zeng, B., Tang, X., Long, Q., et al. (2020). Fine mapping of the *qHTB1-1* QTL, which confers heat tolerance at the booting stage, using an *Oryza rufipogon* Griff. introgression line. *Theor. Appl. Genet.* 133, 1161–1175. doi: 10.1007/s00122-020-03539-7

Cavanagh, C., Morell, M., Mackay, I., and Powell, W. (2008). From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Curr. Opin. Plant Biol.* 11, 215–221. doi: 10.1016/j.pbi.2008.01.002

Cheema, K. K., Bains, N. S., Mangat, G. S., Das, A., Vikal, Y., Brar, D. S., et al. (2008). Development of high yielding IR64 × *Oryza rufipogon* (Griff.) introgression lines and identification of introgressed alien chromosome segments using SSR markers. *Euphytica* 160, 401–409. doi: 10.1007/s10681-007-9560-x

Chen, Z., Lu, Y., Feng, L., Hao, W., Li, C., Yang, Y., et al. (2020). Genetic dissection and functional differentiation of *ALK*^{a} and *ALK*^{b}, two natural alleles of the *ALK/SSIIa* gene, responding to low gelatinization temperature in rice. *Rice* 13:39. doi: 10.1186/s12284-020-00393-5

Cheng, L., Wang, Y., Meng, L., Hu, X., Cui, Y., 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, 45–55. doi: 10.1139/g11-075

Chutimanukul, P., Kositsup, B., Plaimas, K., Buaboocha, T., Siangliw, M., Toojinda, T., et al. (2018). Data in support of photosynthetic responses in a chromosome segment substitution line of 'Khao Dawk Mali 105' rice at seedling stage. *Data Brief* 21, 307–312. doi: 10.1016/j.dib.2018.09.128

Civáò, P., Craig, H., Cox, C. J., and Brown, T. A. (2015). Three geographically separate domestications of Asian rice. *Nat. Plants* 1:15164. doi: 10.1038/nplants.2015.164

Dan, Z., Liu, P., Huang, W., Zhou, W., Yao, G., Hu, J., et al. (2014). Balance between a higher degree of heterosis and increased reproductive isolation: a strategic design for breeding inter-subspecific hybrid rice. *PLoS One* 9:e93122. doi: 10.1371/journal.pone.0093122

Dixit, S., Singh, U. M., Singh, A. K., Alam, S., Venkateshwarlu, C., Nachimuthu, V. V., et al. (2020). Marker assisted forward breeding to combine multiple biotic-abiotic stress resistance/tolerance in rice. *Rice* 13:29. doi: 10.1186/s12284-020-00391-7

Dong, X., Wang, X., Zhang, L., Yang, Z., Xin, X., Wu, S., et al. (2013). Identification and characterization of *OsEBS*, a gene involved in enhanced plant biomass and spikelet number in rice. *Plant Biotechnol. J.* 11, 1044–1057. doi: 10.1111/pbi.12097

Ebitani, T., Takeuchi, Y., Nonoue, Y., Yamamoto, T., Takeuchi, K., and Yano, M. (2005). Construction and evaluation of chromosome segment substitution lines carrying overlapping chromosome segments of indica rice cultivar 'Kasalath' in a genetic background of japonica elite cultivar 'Koshihikari'. *Breed. Sci.* 55, 65–73. doi: 10.1270/jsbbs.55.65

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., et al. (2006). *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* 112, 1164–1171. doi: 10.1007/s00122-006-0218-1

Fasahat, P., Muhammad, K., Abdullah, A., and Wickneswari, R. (2012). Identification of introgressed alien chromosome segments associated with grain quality in *Oryza rufipogon* x MR219 advanced breeding lines using SSR markers. *Genet. Mol. Res.* 11, 3534–3546. doi: 10.4238/2012.September.26.10

Feng, B., Chen, K., Cui, Y., Wu, Z., Zheng, T., Zhu, Y., et al. (2018). Genetic dissection and simultaneous improvement of drought and low nitrogen tolerances by designed QTL pyramiding in rice. *Front. Plant Sci.* 9:306. doi: 10.3389/fpls.2018.00306

Feng, X., Wang, C., Nan, J., Zhang, X., Wang, R., Jiang, G., et al. (2017). Updating the elite rice variety Kongyu 131 by improving the *Gn1a* locus. *Rice* 10:35. doi: 10.1186/s12284-017-0174-1

Fu, Q., Zhang, P., Tan, L., Zhu, Z., Ma, D., Fu, Y., et al. (2010). Analysis of QTLs for yield-related traits in Yuanjiang common wild rice (*Oryza rufipogon* Griff.). *J. Genet. Genomics* 37, 147–157. doi: 10.1016/s1673-8527(09)60033-3

Fujita, D., Tagle, A. G., Ebron, L. A., Fukuta, Y., and Kobayashi, N. (2012). Characterization of near-isogenic lines carrying QTL for high spikelet number with the genetic background of an indica rice variety IR64 *(Oryza sativa* L.). *Breed. Sci.* 62, 18–26. doi: 10.1270/jsbbs.62.18

Fukuta, Y., Konisho, K., Senoo-Namai, S., Yanagihara, S., Tsunematsu, H., Fukuo, A., et al. (2012). Genetic characterization of rainfed upland New Rice for Africa (NERICA) varieties. *Breed. Sci.* 62, 27–37. doi: 10.1270/jsbbs.62.27

Furuta, T., Uehara, K., Angeles-Shim, R. B., Shim, J., Ashikari, M., and Takashi, T. (2014). Development and evaluation of chromosome segment substitution lines (CSSLs) carrying chromosome segments derived from *Oryza rufipogon* in the genetic background of *Oryza sativa* L. *Breed. Sci.* 63, 468–475. doi: 10.1270/jsbbs.63.468

Furuta, T., Uehara, K., Angeles-Shim, R. B., Shim, J., Nagai, K., Ashikari, M., et al. (2016). Development of chromosome segment substitution lines harboring *Oryza nivara* genomic segments in Koshihikari and evaluation of yield-related traits. *Breed. Sci.* 66, 845–850. doi: 10.1270/jsbbs.16131

Gao, D., Sun, W., Wang, D., Dong, H., Zhang, R., and Yu, S. (2020). A xylan glucuronosyltransferase gene exhibits pleiotropic effects on cellular composition and leaf development in rice. *Sci. Rep.* 10:3726. doi: 10.1038/s41598-020-60593-3

Garcia-Oliveira, A. L., Tan, L., Fu, Y., and Sun, C. (2009). Genetic identification of quantitative trait loci for contents of mineral nutrients in rice grain. *J. Integr. Plant Biol.* 51, 84–92. doi: 10.1111/j.1744-7909.2008.00730.x

Grover, N., Kumar, A., Yadav, A. K., Gopala Krishnan, S., Ellur, R. K., Bhowmick, P. K., et al. (2020). Marker assisted development and characterization of herbicide tolerant near isogenic lines of a mega Basmati rice variety, "Pusa Basmati 1121". *Rice* 13:68. doi: 10.1186/s12284-020-00423-2

Guo, S.-B., Wei, Y., Li, X.-Q., Liu, K.-Q., Huang, F.-K., Chen, C.-H., et al. (2013). Development and identification of introgression lines from cross of *Oryza sativa* and *Oryza minuta*. *Rice Sci.* 20, 95–102. doi: 10.1016/S1672-6308(13)60111-0

Gutiérrez, A. G., Carabalí, S. J., Giraldo, O. X., Martínez, C. P., Correa, F., Prado, G., et al. (2010). Identification of a *Rice stripe necrosis virus* resistance locus and yield component QTLs using *Oryza sativa* x *O. glaberrima* introgression lines. *BMC Plant Biol.* 10:6. doi: 10.1186/1471-2229-10-6

Hao, W., Zhu, M. Z., Gao, J. P., Sun, S. Y., and Lin, H. X. (2009). Identification of quantitative trait loci for rice quality in a population of chromosome segment

substitution lines. *J. Integr. Plant Biol.* 51, 500–512. doi: 10.1111/j.1744-7909.2009.00822.x

Harlan, H. V., and Pope, M. N. (1922). The use and value of back-crosses in small-grain breeding. *J. Hered.* 13, 319–322. doi: 10.1093/oxfordjournals.jhered.a102237

He, F., Xi, Z., Zeng, R., Akshay, T., and Zhang, G. (2005a). Mapping of heading date QTLs in rice (*Oryza sativa*) using single segment substitution lines. *Zhongguo Nong Ye Ke Xue* 38, 1505–1513.

He, F., Xi, Z., Zeng, R., Talukdar, A., and Zhang, G. (2005b). Developing single segment substitution lines (SSSLs) in rice (*Oryza sativa* L.) using advanced backcrosses and MAS. *Yi Chuan Xue Bao* 32, 825–831.

He, F., Zhang, X., Zeng, R., Talukdar, A., and Zhang, G. (2005c). Identification of QTLs for plant height and its components by using single segment substitution lines in rice (*Oryza sativa*). *Rice Sci.* 12, 151–156.

He, N., Wu, R., Pan, X., Peng, L., Sun, K., Zou, T., et al. (2017). Development and trait evaluation of chromosome single-segment substitution lines of *O. meridionalis* in the background of *O. sativa*. *Euphytica* 213:281. doi: 10.1007/s10681-017-2072-4

Hirabayashi, H., Sasaki, K., Kambe, T., Gannaban, R. B., Miras, M. A., Mendioro, M. S., et al. (2015). *qEMF3*, a novel QTL for the early-morning flowering trait from wild rice, *Oryza officinalis*, to mitigate heat stress damage at flowering in rice, *O. sativa*. *J. Exp. Bot.* 66, 1227–1236. doi: 10.1093/jxb/eru474

Hori, K., Saisho, D., Nagata, K., Nonoue, Y., Uehara-Yamaguchi, Y., Kanatani, A., et al. (2021a). Genetic elucidation for response of flowering time to ambient temperatures in Asian rice cultivars. *Int. J. Mol. Sci.* 22:1024. doi: 10.3390/ijms22031024

Hori, K., Suzuki, K., Ishikawa, H., Nonoue, Y., Nagata, K., Fukuoka, S., et al. (2021b). Genomic regions involved in differences in eating and cooking quality other than *Wx* and *Alk* genes between indica and japonica rice cultivars. *Rice* 14:8. doi: 10.1186/s12284-020-00447-8

Hori, K., Sugimoto, K., Nonoue, Y., Ono, N., Matsubara, K., Yamanouchi, U., et al. (2010). Detection of quantitative trait loci controlling pre-harvest sprouting resistance by using backcrossed populations of japonica rice cultivars. *Theor. Appl. Genet.* 120, 1547–1557. doi: 10.1007/s00122-010-1275-z

Hu, B., Wang, W., Ou, S., Tang, J., Li, H., Che, R., et al. (2015). Variation in *NRT1.1B* contributes to nitrate-use divergence between rice subspecies. *Nat. Genet.* 47, 834–838. doi: 10.1038/ng.3337

Hu, J., Chang, X., Zou, L., Tang, W., and Wu, W. (2018). Identification and fine mapping of *Bph33*, a new brown planthopper resistance gene in rice (*Oryza sativa* L.). *Rice* 11:55. doi: 10.1186/s12284-018-0249-7

Hu, J., Cheng, M., Gao, G., Zhang, Q., Xiao, J., and He, Y. (2013). Pyramiding and evaluation of three dominant brown planthopper resistance genes in the elite indica rice 9311 and its hybrids. *Pest Manage. Sci.* 69, 802–808. doi: 10.1002/ps.3437

Hu, J., Xiao, C., and He, Y. (2016). Recent progress on the genetics and molecular breeding of brown planthopper resistance in rice. *Rice* 9:30. doi: 10.1186/s12284-016-0099-0

Hua, L., Wang, D. R., Tan, L., Fu, Y., Liu, F., Xiao, L., et al. (2015). *LABA1*, a domestication gene associated with long, barbed awns in wild rice. *Plant Cell* 27, 1875–1888. doi: 10.1105/tpc.15.00260

Huang, L., Gu, Z., Chen, Z., Yu, J., Chu, R., Tan, H., et al. (2021). Improving rice eating and cooking quality by coordinated expression of the major starch synthesis-related genes, *SSII* and *Wx*, in endosperm. *Plant Mol. Biol.* 106, 419–432. doi: 10.1007/s11103-021-01162-8

Huang, N., Angeles, E. R., Domingo, J., Magpantay, G., Singh, S., Zhang, G., et al. (1997). Pyramiding of bacterial blight resistance genes in rice: marker-assisted selection using RFLP and PCR. *Theor. Appl. Genet.* 95, 313–320. doi: 10.1007/s001220050565

Huang, X. Y., Deng, F., Yamaji, N., Pinson, S. R., Fujii-Kashino, M., Danku, J., et al. (2016). A heavy metal P-type ATPase OsHMA4 prevents copper accumulation in rice grain. *Nat. Commun.* 7:12138. doi: 10.1038/ncomms12138

Huang, Y., Zhao, S., Fu, Y., Sun, H., Ma, X., Tan, L., et al. (2018). Variation in the regulatory region of *FZP* causes increases in secondary inflorescence branching and grain yield in rice domestication. *Plant J.* 96, 716–733. doi: 10.1111/tpj.14062

Huo, X., Wu, S., Zhu, Z., Liu, F., Fu, Y., Cai, H., et al. (2017). *NOG1* increases grain production in rice. *Nat. Commun.* 8:1497. doi: 10.1038/s41467-017-01501-8

Jiang, H., Feng, Y., Bao, L., Li, X., Gao, G., Zhang, Q., et al. (2012). Improving blast resistance of Jin 23B and its hybrid rice by marker-assisted gene pyramiding. *Mol. Breed.* 30, 1679–1688. doi: 10.1007/s11032-012-9751-6

Jiang, H., Hu, J., Li, Z., Liu, J., Gao, G., Zhang, Q., et al. (2018). Evaluation and breeding application of six brown planthopper resistance genes in rice maintainer line Jin 23B. *Rice* 11:22. doi: 10.1186/s12284-018-0215-4

Jiang, H., Li, Z., Liu, J., Shen, Z., Gao, G., Zhang, Q., et al. (2019). Development and evaluation of improved lines with broad-spectrum resistance to rice blast using nine resistance genes. *Rice* 12:29. doi: 10.1186/s12284-019-0292-z

Jiang, J., Mou, T., Yu, H., and Zhou, F. (2015). Molecular breeding of thermo-sensitive genic male sterile (TGMS) lines of rice for blast resistance using *Pi2* gene. *Rice* 8:11. doi: 10.1186/s12284-015-0048-3

Jin, J., Hua, L., Zhu, Z., Tan, L., Zhao, X., Zhang, W., et al. (2016). *GAD1* encodes a secreted peptide that regulates grain number, grain length, and awn development in rice domestication. *Plant Cell* 28, 2453–2463. doi: 10.1105/tpc.16.00379

Jones, M. P., Dingkuhn, M., Aluko/snm, G. K., and Semon, M. (1997). Interspecific *Oryza Sativa* L. X *O. Glaberrima* Steud. progenies in upland rice improvement. *Euphytica* 94, 237–246. doi: 10.1023/A:1002969932224

Kabange, N. R., Park, S. Y., Lee, J. Y., Shin, D., Lee, S. M., Kwon, Y., et al. (2021). New insights into the transcriptional regulation of genes involved in the nitrogen use efficiency under potassium chlorate in rice (*Oryza sativa* L.). *Int. J. Mol. Sci.* 22:2192. doi: 10.3390/ijms22042192

Kato, K., and Hirayama, Y. (2021). Development and characterization of chromosome segment substitution lines derived from backcross between japonica donor rice cultivar Yukihikari and japonica recipient cultivar Kirara397. *Breed. Sci.* 71, 283–290. doi: 10.1270/jsbbs.20128

Khan, G. H., Shikari, A. B., Vaishnavi, R., Najeeb, S., Padder, B. A., Bhat, Z. A., et al. (2018). Marker-assisted introgression of three dominant blast resistance genes into an aromatic rice cultivar *Mushk Budji*. *Sci. Rep.* 8:4091. doi: 10.1038/s41598-018-22246-4

Kijima, Y., Sserunkuuma, D., and Otsuka, K. (2006). How revolutionary is the "Nerica Revolution"? Evidence from Uganda. *Dev. Econ.* 44, 252–267. doi: 10.1111/j.1746-1049.2006.00016.x

Kijoji, A. A., Nchimbi-Msolla, S., Kanyeka, Z. L., Serraj, R., and Henry, A. (2014). Linking root traits and grain yield for rainfed rice in sub-Saharan Africa: response of *Oryza sativa*×*Oryza glaberrima* introgression lines under drought. *Field Crops Res.* 165, 25–35. doi: 10.1016/j.fcr.2014.03.019

Kim, S. H., Ji, S. D., Lee, H. S., Jeon, Y. A., Shim, K. C., Adeva, C., et al. (2021). A novel embryo phenotype associated with interspecific hybrid weakness in rice is controlled by the MADS-domain transcription factor OsMADS8. *Front. Plant Sci.* 12:778008. doi: 10.3389/fpls.2021.778008

Kim, S. R., Ramos, J. M., Hizon, R. J. M., Ashikari, M., Virk, P. S., Torres, E. A., et al. (2018). Introgression of a functional epigenetic $OsSPL14^{WFP}$ allele into elite indica rice genomes greatly improved panicle traits and grain yield. *Sci. Rep.* 8:3833. doi: 10.1038/s41598-018-21355-4

Koizumi, S., Ashizawa, T., and Zenbayashi, K. S. (2004). "Durable control of rice blast disease with multilines," in *Rice Blast: Interaction with Rice and Control*, ed. S. Kawasaki (Dordrecht: Springer), 191–199. doi: 10.1007/978-0-306-48582-4_23

Kovach, M. J., Sweeney, M. T., and McCouch, S. R. (2007). New insights into the history of rice domestication. *Trends Genet.* 23, 578–587. doi: 10.1016/j.tig.2007.08.012

Kovi, M. R., Zhang, Y., Yu, S., Yang, G., Yan, W., and Xing, Y. (2011). Candidacy of a chitin-inducible gibberellin-responsive gene for a major locus affecting plant height in rice that is closely linked to Green Revolution gene *sd1*. *Theor. Appl. Genet.* 123, 705–714. doi: 10.1007/s00122-011-1620-x

Kubo, T., Aida, Y., Nakamura, K., Tsunematsu, H., Doi, K., and Yoshimura, A. (2002). Reciprocal chromosome segment substitution series derived from japonica and indica cross of rice (*Oryza sativa* L.). *Breed. Sci.* 52, 319–325. doi: 10.1270/jsbbs.52.319

Lei, D., Tan, L., Liu, F., Chen, L., and Sun, C. (2013). Identification of heat-sensitive QTL derived from common wild rice (*Oryza rufipogon* Griff.). *Plant Sci.* 20, 121–127. doi: 10.1016/j.plantsci.2012.12.001

Li, D., Huang, Z., Song, S., Xin, Y., Mao, D., Lv, Q., et al. (2016). Integrated analysis of phenome, genome, and transcriptome of hybrid rice uncovered multiple heterosis-related loci for yield increase. *Proc. Natl. Acad. Sci. U.S.A.* 113, E6026–E6035. doi: 10.1073/pnas.1610115113

Li, M., Sun, P., Zhou, H., Chen, S., and Yu, S. (2011). Identification of quantitative trait loci associated with germination using chromosome segment substitution lines of rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 123, 411–420. doi: 10.1007/s00122-011-1593-9

Li, X., Chao, D., Wu, Y., Huang, X., Chen, K., Cui, L., et al. (2015). Natural alleles of a proteasome α2 subunit gene contribute to thermotolerance and adaptation of African rice. *Nat. Genet.* 47, 827–833. doi: 10.1038/ng.3305

Li, X., Liu, Z., Lu, D., Liu, Y., Mao, X., Li, Z., et al. (2013). Development and evaluation of multi-genotype varieties of rice derived from MAGIC lines. *Euphytica* 192, 77–86. doi: 10.1007/s10681-013-0879-1

Li, J., Zhou, J., Xu, P., Deng, X., Deng, W., Zhang, Y., et al. (2018a). Mapping five novel interspecific hybrid sterility loci between *Oryza sativa* and *Oryza meridionalis*. *Breed. Sci.* 68, 516–523. doi: 10.1270/jsbbs.18001

Li, S., Tian, Y., Wu, K., Ye, Y., Yu, J., Zhang, J., et al. (2018b). Modulating plant growth-metabolism coordination for sustainable agriculture. *Nature* 560, 595–600. doi: 10.1038/s41586-018-0415-5

Li, Y., Xiao, J., Chen, L., Huang, X., Cheng, Z., Han, B., et al. (2018c). Rice functional genomics research: past decade and future. *Mol. Plant* 11, 359–380. doi: 10.1016/j.molp.2018.01.007

Li, Z., and Zheng, T. (2013). "Utilization of exotic germplasm," in *Genetics and Genomics of Rice*, eds Q. Zhang and R. A. Wing (New York, NY: Springer New York), 349–361. doi: 10.1007/978-1-4614-7903-1_23

Liang, Y., Meng, L., Lin, X., Cui, Y., Pang, Y., Xu, J., et al. (2018). QTL and QTL networks for cold tolerance at the reproductive stage detected using selective introgression in rice. *PLoS One* 13:e0200846. doi: 10.1371/journal.pone.0200846

Lin, Q., Wang, W., Ren, Y., Jiang, Y., Sun, A., Qian, Y., et al. (2015). Genetic dissection of seed storability using two different populations with a same parent rice cultivar N22. *Breed. Sci.* 65, 411–419. doi: 10.1270/jsbbs.65.411

Lin, T., Zhou, C., Chen, G., Yu, J., Wu, W., Ge, Y., et al. (2020). Heterosis-associated genes confer high yield in super hybrid rice. *Theor. Appl. Genet.* 133, 3287–3297. doi: 10.1007/s00122-020-03669-y

Liu, G., Zeng, R., Zhu, H., Zhang, Z., Ding, X., Zhao, F., et al. (2009). Dynamic expression of nine QTLs for tiller number detected with single segment substitution lines in rice. *Theor. Appl. Genet.* 118, 443–453. doi: 10.1007/s00122-008-0911-3

Liu, G., Zhang, Z., Zhu, H., Zhao, F., Ding, X., Zeng, R., et al. (2008). Detection of QTLs with additive effects and additive-by-environment interaction effects on panicle number in rice (*Oryza sativa* L.) with single-segment substitution lines. *Theor. Appl. Genet.* 116, 923–931. doi: 10.1007/s00122-008-0724-4

Liu, G. M., Li, W. T., Zeng, R. Z., Zhang, Z. M., and Zhang, G. Q. (2004). [Identification of QTLs on substituted segments in single segment substitution lines of rice]. *Yi Chuan Xue Bao* 31, 1395–1400.

Liu, X., Zhao, Z., Liu, L., Xiao, Y., Tian, Y., Liu, S.-J., et al. (2016). Construction of chromosomal segment substitution lines and genetic dissection of introgressed segments associated with yield determination in the parents of a super-hybrid rice. *Plant Breed.* 135, 63–72. doi: 10.1111/pbr.12329

Liu, Y., Wang, H., Jiang, Z., Wang, W., Xu, R., Wang, Q., et al. (2021). Genomic basis of geographical adaptation to soil nitrogen in rice. *Nature* 590, 600–605. doi: 10.1038/s41586-020-03091-w

Luan, X., Dai, Z., Yang, W., Tan, Q., Lu, Q., Guo, J., et al. (2019). Breeding by design of CMS lines on the platform of SSSL library in rice. *Mol. Breed.* 39:126. doi: 10.1007/s11032-019-1028-x

Luo, J., Liu, H., Zhou, T., Gu, B., Huang, X., Shangguan, Y., et al. (2013). *An-1* encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. *Plant Cell* 25, 3360–3376. doi: 10.1105/tpc.113.113589

Luo, J. S., Huang, J., Zeng, D. L., Peng, J. S., Zhang, G. B., Ma, H. L., et al. (2018). A defensin-like protein drives cadmium efflux and allocation in rice. *Nat. Commun.* 9:645. doi: 10.1038/s41467-018-03088-0

Luo, Y., Ma, T., Zhang, A., Ong, K. H., Li, Z., Yang, J., et al. (2016). Marker-assisted breeding of the rice restorer line Wanhui 6725 for disease resistance, submergence tolerance and aromatic fragrance. *Rice* 9:66. doi: 10.1186/s12284-016-0139-9

Ma, X., Fu, Y., Zhao, X., Jiang, L., Zhu, Z., Gu, P., et al. (2016). Genomic structure analysis of a set of *Oryza nivara* introgression lines and identification of yield-associated QTLs using whole-genome resequencing. *Sci. Rep.* 6:27425. doi: 10.1038/srep27425

Ma, X., Han, B., Tang, J., Zhang, J., Cui, D., Geng, L., et al. (2019). Construction of chromosome segment substitution lines of Dongxiang common wild rice (*Oryza rufipogon* Griff.) in the background of the japonica rice cultivar Nipponbare (*Oryza sativa* L.). *Plant Physiol. Biochem.* 144, 274–282. doi: 10.1016/j.plaphy.2019.09.041

Mammadov, J., Buyyarapu, R., Guttikonda, S. K., Parliament, K., Abdurakhmonov, I. Y., and Kumpatla, S. P. (2018). Wild relatives of maize, rice, cotton, and soybean: treasure troves for tolerance to biotic and abiotic stresses. *Front. Plant Sci.* 9:886. doi: 10.3389/fpls.2018.00886

Matsubara, K., Ogiso-Tanaka, E., Hori, K., Ebana, K., Ando, T., and Yano, M. (2012). Natural variation in *Hd17*, a homolog of *Arabidopsis ELF3* that is involved in rice photoperiodic flowering. *Plant Cell Physiol.* 53, 709–716. doi: 10.1093/pcp/pcs028

Mi, J., Yang, D., Chen, Y., Jiang, J., Mou, H., Huang, J., et al. (2018). Accelerated molecular breeding of a novel P/TGMS line with broad-spectrum resistance to rice blast and bacterial blight in two-line hybrid rice. *Rice* 11:11. doi: 10.1186/s12284-018-0203-8

Michael, J. T. (2014). High-throughput SNP genotyping to accelerate crop improvement. *Plant Breed. Biotechnol.* 2, 195–212. doi: 10.9787/PBB.2014.2.3.195

Mitsuya, S., Murakami, N., Sato, T., Kazama, T., Toriyama, K., Skoulding, N. S., et al. (2019). Evaluation of rice grain yield and yield components of Nona Bokra chromosome segment substitution lines with the genetic background of Koshihikari, in a saline paddy field. *AoB Plants* 11:plz040. doi: 10.1093/aobpla/plz040

Mizobuchi, R., Sato, H., Fukuoka, S., Tsushima, S., Imbe, T., and Yano, M. (2013). Identification of *qRBS1*, a QTL involved in resistance to bacterial seedling rot in rice. *Theor. Appl. Genet.* 126, 2417–2425. doi: 10.1007/s00122-013-2145-2

Mizuno, Y., Yamanouchi, U., Hoshino, T., Nonoue, Y., Nagata, K., Fukuoka, S., et al. (2018). Genetic dissection of pre-harvest sprouting resistance in an upland rice cultivar. *Breed. Sci.* 68, 200–209. doi: 10.1270/jsbbs.17062

Monteverde, E., Gutierrez, L., Blanco, P., Pérez de Vida, F., Rosas, J. E., Bonnecarrère, V., et al. (2019). Integrating molecular markers and environmental covariates to interpret genotype by environment interaction in rice (*Oryza sativa* L.) grown in subtropical areas. *G3* 9, 1519–1531. doi: 10.1534/g3.119.400064

Mulsanti, I. W., Yamamoto, T., Ueda, T., Samadi, A. F., Kamahora, E., Rumanti, I. A., et al. (2018). Finding the superior allele of japonica-type for increasing stem lodging resistance in indica rice varieties using chromosome segment substitution lines. *Rice* 11:25. doi: 10.1186/s12284-018-0216-3

Munguambe, N. E., Inoue, S., Demeter, Z., Yamagata, Y., Yasui, H., Zheng, S. H., et al. (2021). Substitution mapping of a locus responsible for hybrid breakdown in populations derived from interspecific introgression line. *Front. Plant Sci.* 12:633247. doi: 10.3389/fpls.2021.633247

Murata, K., Iyama, Y., Yamaguchi, T., Ozaki, H., Kidani, Y., and Ebitani, T. (2014). Identification of a novel gene (*Apq1*) from the indica rice cultivar 'Habataki' that improves the quality of grains produced under high temperature stress. *Breed. Sci.* 64, 273–281. doi: 10.1270/jsbbs.64.273

Nounjan, N., Siangliw, J. L., Toojinda, T., Chadchawan, S., and Theerakulpisut, P. (2016). Salt-responsive mechanisms in chromosome segment substitution lines of rice (*Oryza sativa* L. cv. KDML105). *Plant Physiol. Biochem.* 103, 96–105. doi: 10.1016/j.plaphy.2016.02.038

Ogami, T., Yasui, H., Yoshimura, A., and Yamagata, Y. (2019). Identification of anther length QTL and construction of chromosome segment substitution lines of *Oryza longistaminata*. *Plants* 8:388. doi: 10.3390/plants8100388

Ookawa, T., Aoba, R., Yamamoto, T., Ueda, T., Takai, T., Fukuoka, S., et al. (2016). Precise estimation of genomic regions controlling lodging resistance using a set of reciprocal chromosome segment substitution lines in rice. *Sci. Rep.* 6:30572. doi: 10.1038/srep30572

Park, D. S., Lee, S. K., Lee, J. H., Song, M. Y., Song, S. Y., Kwak, D. Y., et al. (2007). The identification of candidate rice genes that confer resistance to the brown planthopper (*Nilaparvata lugens*) through representational difference analysis. *Theor. Appl. Genet.* 115, 537–547. doi: 10.1007/s00122-007-0587-0

Prahalada, G., Marathi, B., Vinarao, R., Kim, S., Diocton, R., Ramos, J., et al. (2021). QTL mapping of a novel genomic region associated with high out-crossing rate derived from *Oryza longistaminata* and development of new CMS lines in rice, *O. sativa* L. *Rice* 14:80. doi: 10.1186/s12284-021-00521-9

Prasanth, V. V., Basava, K. R., Babu, M. S., Vt, G. N. V., Devi, S. J., Mangrauthia, S. K., et al. (2016). Field level evaluation of rice introgression lines for heat tolerance and validation of markers linked to spikelet fertility. *Physiol. Mol. Biol. Plants* 22, 179–192. doi: 10.1007/s12298-016-0350-6

Qi, L., Sun, Y., Li, J., Su, L., Zheng, X., Wang, X., et al. (2017). Identify QTLs for grain size and weight in common wild rice using chromosome segment substitution lines across six environments. *Breed. Sci.* 67, 472–482. doi: 10.1270/jsbbs.16082

Qiao, W., Qi, L., Cheng, Z., Su, L., Li, J., Sun, Y., et al. (2016). Development and characterization of chromosome segment substitution lines derived from *Oryza rufipogon* in the genetic background of *O. sativa* spp. *indica* cultivar 9311. *BMC Genomics* 17:580. doi: 10.1186/s12864-016-2987-5

Qin, P., Lu, H., Du, H., Wang, H., Chen, W., Chen, Z., et al. (2021). Pan-genome analysis of 33 genetically diverse rice accessions reveals hidden genomic variations. *Cell* 184, 3542–3558.e16. doi: 10.1016/j.cell.2021.04.046

Qiu, X., Chen, K., Lv, W., Ou, X., Zhu, Y., Xing, D., et al. (2017). Examining two sets of introgression lines reveals background-independent and stably expressed QTL that improve grain appearance quality in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 130, 951–967. doi: 10.1007/s00122-017-2862-z

Qiu, X., Yuan, Z., Liu, H., Xiang, X., Yang, L., He, W., et al. (2015). Identification of salt tolerance-improving quantitative trait loci alleles from a salt-susceptible rice breeding line by introgression breeding. *Plant Breed.* 134, 653–660. doi: 10.1111/pbr.12321

Ramos, J., Furuta, T., Uehara, K., Chihiro, N., Angeles-Shim, R., Shim, J., et al. (2016). Development of chromosome segment substitution lines (CSSLs) of *Oryza longistaminata* A. Chev. & Röhr in the background of the elite japonica rice cultivar, Taichung 65 and their evaluation for yield traits. *Euphytica* 210, 151–163. doi: 10.1007/s10681-016-1685-3

Rao, Y., Li, Y., and Qian, Q. (2014). Recent progress on molecular breeding of rice in China. *Plant Cell Rep.* 33, 551–564. doi: 10.1007/s00299-013-1551-x

Rao, Y. V., Balakrishnan, D., Addanki, K. R., Mesapogu, S., Kiran, T. V., Subrahmanyam, D., et al. (2018). Characterization of backcross introgression lines derived from *Oryza nivara* accessions for photosynthesis and yield. *Physiol. Mol. Biol. Plants* 24, 1147–1164. doi: 10.1007/s12298-018-0575-7

Sanchez, P. L., Wing, R. A., and Brar, D. S. (2013). "The wild relative of rice: Genomes and genomics," in *Genetics and Genomics of Rice*, eds Q. Zhang and R. A. Wing (New York, NY: Springer New York), 9–25. doi: 10.1007/978-1-4614-7903-1_2

Shen, G., Hu, W., Wang, X., Zhou, X., Han, Z., Sherif, A., et al. (2022). Assembly of yield heterosis of an elite rice hybrid is promising by manipulating dominant QTLs. *J. Integr. Plant Biol.* 64, 688–701. doi: 10.1111/jipb.13220

Shen, G., and Xing, Y. (2014). Two novel QTLs for heading date are identified using a set of chromosome segment substitution lines in rice (*Oryza sativa* L.). *J. Genet. Genomics* 41, 659–662. doi: 10.1016/j.jgg.2014.08.008

Shen, G., Zhan, W., Chen, H., and Xing, Y. (2014). Dominance and epistasis are the main contributors to heterosis for plant height in rice. *Plant Sci.* 215–216, 11–18. doi: 10.1016/j.plantsci.2013.10.004

Shi, C. L., Dong, N. Q., Guo, T., Ye, W. W., Shan, J. X., and Lin, H. X. (2020). A quantitative trait locus *GW6* controls rice grain size and yield through the gibberellin pathway. *Plant J.* 103, 1174–1188. doi: 10.1111/tpj.14793

Shibaya, T., Hori, K., Ogiso-Tanaka, E., Yamanouchi, U., Shu, K., Kitazawa, N., et al. (2016). *Hd18*, encoding histone acetylase related to *Arabidopsis* FLOWERING LOCUS D, is involved in the control of flowering time in rice. *Plant Cell Physiol.* 57, 1828–1838. doi: 10.1093/pcp/pcw105

Shim, R. A., Angeles, E. R., Ashikari, M., and Takashi, T. (2010). Development and evaluation of *Oryza glaberrima* Steud. chromosome segment substitution lines (CSSLs) in the background of *O. sativa* L. cv. Koshihikari. *Breed. Sci.* 60, 613–619. doi: 10.1270/jsbbs.60.613

Shomura, A., Izawa, T., Ebana, K., Ebitani, T., Kanegae, H., Konishi, S., et al. (2008). Deletion in a gene associated with grain size increased yields during rice domestication. *Nat. Genet.* 40, 1023–1028. doi: 10.1038/ng.169

Singh, N., Wang, D. R., Ali, L., Kim, H., Akther, K. M., Harrington, S. E., et al. (2020). A coordinated suite of wild-introgression lines in indica and japonica elite backgrounds. *Front. Plant Sci.* 11:564824. doi: 10.3389/fpls.2020.564824

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, 2304–2318. doi: 10.1111/pbi.13661

Sui, F., Zhao, D., Zhu, H., Gong, Y., Tang, Z., Huang, X. Y., et al. (2019). Map-based cloning of a new total loss-of-function allele of *OsHMA3* causes high cadmium accumulation in rice grain. *J. Exp. Bot.* 70, 2857–2871. doi: 10.1093/jxb/erz093

Sun, H., Qian, Q., Wu, K., Luo, J., Wang, S., Zhang, C., et al. (2014). Heterotrimeric G proteins regulate nitrogen-use efficiency in rice. *Nat. Genet.* 46, 652–656. doi: 10.1038/ng.2958

Sun, W., Zhou, Q., Yao, Y., Qiu, X., Xie, K., and Yu, S. (2015). Identification of genomic regions and the isoamylase gene for reduced grain chalkiness in rice. *PLoS One* 10:e0122013. doi: 10.1371/journal.pone.0122013

Sun, Z., Yin, X., Ding, J., Yu, D., Hu, M., Sun, X., et al. (2017). QTL analysis and dissection of panicle components in rice using advanced backcross populations derived from *Oryza sativa* cultivars HR1128 and 'Nipponbare'. *PLoS One* 12:e0175692. doi: 10.1371/journal.pone.0175692

Suralta, R. R., Inukai, Y., and Yamauchi, A. (2008). Utilizing chromosome segment substitution lines (CSSLs) for evaluation of root responses to transient moisture stresses in rice. *Plant Prod. Sci.* 11, 457–465. doi: 10.1626/pps.11.457

Takai, T., Ikka, T., Kondo, K., Nonoue, Y., Ono, N., Arai-Sanoh, Y., et al. (2014). Genetic mechanisms underlying yield potential in the rice high-yielding cultivar Takanari, based on reciprocal chromosome segment substitution lines. *BMC Plant Biol.* 14:295. doi: 10.1186/s12870-014-0295-2

Takai, T., Nonoue, Y., Yamamoto, S.-I., Yamanouchi, U., Matsubara, K., Liang, Z.-W., et al. (2007). Development of chromosome segment substitution lines derived from backcross between indica donor rice cultivar 'Nona Bokra' and japonica recipient cultivar 'Koshihikari'. *Breed. Sci.* 57, 257–261. doi: 10.1270/jsbbs.57.257

Takai, T., Ohsumi, A., San-oh, Y., Laza, M. R., Kondo, M., Yamamoto, T., et al. (2009). Detection of a quantitative trait locus controlling carbon isotope discrimination and its contribution to stomatal conductance in japonica rice. *Theor. Appl. Genet.* 118, 1401–1410. doi: 10.1007/s00122-009-0990-9

Tan, L., Li, X., Liu, F., Sun, X., Li, C., Zhu, Z., et al. (2008). Control of a key transition from prostrate to erect growth in rice domestication. *Nat. Genet.* 40, 1360–1364. doi: 10.1038/ng.197

Tan, L., Liu, F., Xue, W., Wang, G., Ye, S., Zhu, Z., et al. (2007). Development of *Oryza rufipogon* and *O. sativa* introgression lines and assessment for yield-related quantitative trait loci. *J. Integr. Plant Biol.* 49, 871–884. doi: 10.1111/j.1744-7909.2007.00497.x

Tan, Q., Zou, T., Zheng, M., Ni, Y., Luan, X., Li, X., et al. (2020). Substitution mapping of the major quantitative trait loci controlling stigma exsertion rate from *Oryza glumaepatula*. *Rice* 13:37. doi: 10.1186/s12284-020-00397-1

Tanksley, S. D., and McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277, 1063–1066. doi: 10.1126/science.277.5329.1063

Tao, Y., Zhu, J., Xu, J., Wang, L., Gu, H., Zhou, R., et al. (2016). Exploitation of heterosis loci for yield and yield components in rice using chromosome segment substitution lines. *Sci. Rep.* 6:36802. doi: 10.1038/srep36802

Thein, H. W., Yamagata, Y., Van Mai, T., and Yasui, H. (2019). Four resistance alleles derived from *Oryza longistaminata* (A. Chev. & Roehrich) against green rice leafhopper, *Nephotettix cincticeps* (Uhler) identified using novel introgression lines. *Breed. Sci.* 69, 573–584. doi: 10.1270/jsbbs.19060

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

Tian, Z., Qian, Q., Liu, Q., Yan, M., Liu, X., Yan, C., et al. (2009). Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21760. doi: 10.1073/pnas.0912396106

Uga, Y., Kitomi, Y., Yamamoto, E., Kanno, N., Kawai, S., Mizubayashi, T., et al. (2015). A QTL for root growth angle on rice chromosome 7 is involved in the genetic pathway of *DEEPER ROOTING 1*. *Rice* 8:8. doi: 10.1186/s12284-015-0044-7

Ujiie, K., Yamamoto, T., Yano, M., and Ishimaru, K. (2016). Genetic factors determining varietal differences in characters affecting yield between two rice (*Oryza sativa* L.) varieties, Koshihikari and IR64. *Genet. Resour. Crop Evol.* 63, 97–123. doi: 10.1007/s10722-015-0237-3

Vogel, K. E. (2009). Backcross breeding. *Methods Mol. Biol.* 526, 161–169. doi: 10.1007/978-1-59745-494-0_14

Wan, J. L., Zhai, H. Q., Wan, J. M., Yasui, H., and Yoshimura, A. (2003). Mapping QTL for traits associated with resistance to ferrous iron toxicity in rice (*Oryza sativa* L.), using japonica chromosome segment substitution lines. *Yi Chuan Xue Bao* 30, 893–898.

Wang, D., Sun, W., Yuan, Z., Sun, Q., Fan, K., Zhang, C., et al. (2021). Identification of a novel QTL and candidate gene associated with grain size using chromosome segment substitution lines in rice. *Sci. Rep.* 11:189. doi: 10.1038/s41598-020-80667-6

Wang, S., Li, S., Liu, Q., Wu, K., Zhang, J., Wang, 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, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., 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, 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, 43–49. doi: 10.1038/s41586-018-0063-9

Wang, Y., Wang, J., Zhai, L., Liang, C., Chen, K., and Xu, J. (2020). Identify QTLs and candidate genes underlying source-, sink-, and grain yield-related traits in rice by integrated analysis of bi-parental and natural populations. *PLoS One* 15:e0237774. doi: 10.1371/journal.pone.0237774

Wang, Z., Yu, C., Liu, X., Liu, S., Yin, C., Liu, L., et al. (2012). Identification of indica rice chromosome segments for the improvement of japonica inbreds and hybrids. *Theor. Appl. Genet.* 124, 1351–1364. doi: 10.1007/s00122-012-1792-z

Wei, X., Qiu, J., Yong, K., Fan, J., Zhang, Q., Hua, H., et al. (2021). A quantitative genomics map of rice provides genetic insights and guides breeding. *Nat. Genet.* 53, 243–253. doi: 10.1038/s41588-020-00769-9

Weng, J., Gu, S., Wan, X., Gao, H., Guo, T., Su, N., et al. (2008). Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. *Cell Res.* 18, 1199–1209. doi: 10.1038/cr.2008.307

Wing, R. A., Purugganan, M. D., and Zhang, Q. (2018). The rice genome revolution: from an ancient grain to Green Super Rice. *Nat. Rev. Genet.* 19, 505–517. doi: 10.1038/s41576-018-0024-z

Wu, B., Hu, W., and Xing, Y. Z. (2018). The history and prospect of rice genetic breeding in China. *Yi Chuan* 40, 841–857. doi: 10.16288/j.yczz.18-213

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

Xi, Z. Y., He, F. H., Zeng, R. Z., Zhang, Z. M., Ding, X. H., Li, W. T., et al. (2006). Development of a wide population of chromosome single-segment substitution lines in the genetic background of an elite cultivar of rice (*Oryza sativa* L.). *Genome* 49, 476–484. doi: 10.1139/g06-005

Xia, D., Zhou, H., Liu, R., Dan, W., Li, P., Wu, B., et al. (2018). *GL3.3*, a novel QTL encoding a GSK3/SHAGGY-like kinase, epistatically Interacts with *GS3* to produce extra-long grains in rice. *Mol. Plant* 11, 754–756. doi: 10.1016/j.molp.2018.03.006

Xin, X. Y., Wang, W. X., Yang, J. S., and Luo, X. J. (2011). Genetic analysis of heterotic loci detected in a cross between indica and japonica rice (*Oryza sativa* L.). *Breed. Sci.* 61, 380–388. doi: 10.1270/jsbbs.61.380

Xiong, Y., Zhang, C., Zhou, H., Sun, W., Wang, P., Wang, D., et al. (2021). Identification of heterotic loci with desirable allelic interaction to increase yield in rice. *Rice* 14:97. doi: 10.1186/s12284-021-00539-z

Xu, J., Shang, L., Wang, J., Chen, M., Fu, X., He, H., et al. (2021). The *SEEDLING BIOMASS 1* allele from indica rice enhances yield performance under low-nitrogen environments. *Plant Biotechnol. J.* 19, 1681–1683. doi: 10.1111/pbi.13642

Xu, J., Zhao, Q., Du, P., Xu, C., Wang, B., Feng, Q., et al. (2010). Developing high throughput genotyped chromosome segment substitution lines based on population whole-genome re-sequencing in rice (*Oryza sativa* L.). *BMC Genomics* 11:656. doi: 10.1186/1471-2164-11-656

Xu, P., Dong, L., Zhou, J., Li, J., Zhang, Y., Hu, F., et al. (2015). Identification and mapping of a novel blast resistance gene *Pi57*(t) in *Oryza longistaminata*. *Euphytica* 205, 95–102. doi: 10.1007/s10681-015-1402-7

Xu, X., Wu, K., Xu, R., Yu, J., Wang, J., Zhao, Y., et al. (2019). Pyramiding of the *dep1-1* and *NAL1*^{NJ6} alleles achieves sustainable improvements in nitrogen-use efficiency and grain yield in japonica rice breeding. *J. Genet. Genomics* 46, 325–328. doi: 10.1016/j.jgg.2019.02.009

Xu, Z., Li, S., Zhang, C., Zhang, B., Zhu, K., Zhou, Y., et al. (2017). Genetic connection between cell-wall composition and grain yield via parallel QTL analysis in indica and japonica subspecies. *Sci. Rep.* 7:12561. doi: 10.1038/s41598-017-12903-5

Yadav, A. K., Kumar, A., Grover, N., Ellur, R. K., Krishnan, S. G., Bollinedi, H., et al. (2020). Marker aided introgression of '*Saltol*', a major QTL for seedling stage salinity tolerance into an elite Basmati rice variety 'Pusa Basmati 1509'. *Sci. Rep.* 10:13877. doi: 10.1038/s41598-020-70664-0

Yadav, S., Sandhu, N., Majumder, R. R., Dixit, S., Kumar, S., Singh, S. P., et al. (2019). Epistatic interactions of major effect drought QTLs with genetic background loci determine grain yield of rice under drought stress. *Sci. Rep.* 9:2616. doi: 10.1038/s41598-019-39084-7

Yamagata, Y., Win, K. T., Miyazaki, Y., Ogata, C., Yasui, H., and Yoshimura, A. (2019). Development of introgression lines of AA genome Oryza species, *O. glaberrima*, *O. rufipogon*, and *O. nivara*, in the genetic background of *O. sativa* L. cv. Taichung 65. *Breed. Sci.* 69, 359–363. doi: 10.1270/jsbbs.19002

Yang, C., Zhang, T., Wang, H., Zhao, N., and Liu, B. (2012). Heritable alteration in salt-tolerance in rice induced by introgression from wild rice (*Zizania latifolia*). *Rice* 5:36. doi: 10.1186/1939-8433-5-36

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., Zhang, F., Zafar, S., Wang, J., Lu, H., Naveed, S., et al. (2021c). Genetic dissection of heterosis of indica-japonica by introgression line, recombinant inbred line and their testcross populations. *Sci. Rep.* 11:10265. doi: 10.1038/s41598-021-89691-6

Yang, W., Xiong, L., Liang, J., Hao, Q., Luan, X., Tan, Q., et al. (2021b). Substitution mapping of two closely linked QTLs on chromosome 8 controlling grain chalkiness in rice. *Rice* 14:85. doi: 10.1186/s12284-021-00526-4

Yang, W., Liang, J., Hao, Q., Luan, X., Tan, Q., Lin, S., et al. (2021a). Fine mapping of two grain chalkiness QTLs sensitive to high temperature in rice. *Rice* 14:33. doi: 10.1186/s12284-021-00476-x

Yang, Y., Guo, M., Sun, S., Zou, Y., Yin, S., Liu, Y., et al. (2019). Natural variation of *OsGluA2* is involved in grain protein content regulation in rice. *Nat. Commun.* 10:1949. doi: 10.1038/s41467-019-09919-y

Yang, Y., Zhou, J., Li, J., Xu, P., Zhang, Y., and Tao, D. (2016). Mapping QTLs for hybrid sterility in three AA genome wild species of *Oryza*. *Breed. Sci.* 66, 367–371. doi: 10.1270/jsbbs.15048

Yasui, H., Yamagata, Y., and Yoshimura, A. (2010). Development of chromosome segment substitution lines derived from indica rice donor cultivars DV85 and ARC10313 in the genetic background of japonica cultivar Taichung 65. *Breed. Sci.* 60, 620–628. doi: 10.1270/jsbbs.60.620

Ye, G., Liang, S., and Wan, J. (2010). QTL mapping of protein content in rice using single chromosome segment substitution lines. *Theor. Appl. Genet.* 121, 741–750. doi: 10.1007/s00122-010-1345-2

Yu, B., Lin, Z., Li, H., Li, X., Li, J., Wang, Y., 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, S., Ali, J., Zhang, C., Li, Z., and Zhang, Q. (2020). Genomic breeding of Green Super Rice varieties and their deployment in Asia and Africa. *Theor. Appl. Genet.* 133, 1427–1442. doi: 10.1007/s00122-019-03516-9

Yuan, R., Zhao, N., Usman, B., Luo, L., Liao, S., Qin, Y., et al. (2020). Development of chromosome segment substitution lines (CSSLs) derived from Guangxi wild rice (*Oryza rufipogon* Griff.) under rice (*Oryza sativa* L.) background and the identification of QTLs for plant architecture, agronomic traits and cold tolerance. *Genes* 11:980. doi: 10.3390/genes11090980

Yuan, S., Wang, Y., Zhang, C., He, H., and Yu, S. (2020). Genetic dissection of seed dormancy using chromosome segment substitution lines in rice (*Oryza sativa* L.). *Int. J. Mol. Sci.* 21:1344. doi: 10.3390/ijms21041344

Yuan, Z., Fan, K., Xia, L., Ding, X., Tian, L., Sun, W., et al. (2019). Genetic dissection of seed storability and validation of candidate gene associated with antioxidant capability in rice (*Oryza sativa* L.). *Int. J. Mol. Sci.* 20:4442. doi: 10.3390/ijms20184442

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 *Oryza rufipogon* chromosome segments in japonica rice. *Rice* 9:62. doi: 10.1186/s12284-016-0135-0

Zang, J., Sun, Y., Wang, Y., Yang, J., Li, F., Zhou, Y., et al. (2008). Dissection of genetic overlap of salt tolerance QTLs at the seedling and tillering stages using backcross introgression lines in rice. *Sci. China C Life Sci.* 51, 583–591. doi: 10.1007/s11427-008-0081-1

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

Zhan, P., Ma, S., Xiao, Z., Li, F., Wei, X., Lin, S., et al. (2022). Natural variations in *grain length 10* (*GL10*) regulate rice grain size. *J. Genet. Genomics*. [Epub ahead of print]. doi: 10.1016/j.jgg.2022.01.008

Zhang, F., Shi, Y., Ali, J., Xu, J., and Li, Z. (2021c). Breeding by selective introgression: theory, practices, and lessons learned from rice. *Crop J.* 9, 646–657. doi: 10.1016/j.cj.2021.03.006

Zhang, B., Qi, F., Hu, G., Yang, Y., Zhang, L., Meng, J., et al. (2021a). BSA-seq-based identification of a major additive plant height QTL with an effect equivalent to that of *Semi-dwarf 1* in a large rice F_2 population. *Crop J.* 9, 1428–1437. doi: 10.1016/j.cj.2020.11.011

Zhang, C., Yang, Y., Chen, S., Liu, X., Zhu, J., Zhou, L., et al. (2021b). A rare *Waxy* allele coordinately improves rice eating and cooking quality and grain transparency. *J. Integr. Plant. Biol.* 63, 889–901. doi: 10.1111/jipb.13010

Zhang, C., Yuan, Z., Wang, Y., Sun, W., Tang, X., Sun, Y., et al. (2020a). Genetic dissection of seed dormancy in rice (*Oryza sativa* L.) by using two mapping populations derived from common parents. *Rice* 13:52. doi: 10.1186/s12284-020-00413-4

Zhang, Z., Gao, S., and Chu, C. (2020c). Improvement of nutrient use efficiency in rice: current toolbox and future perspectives. *Theor. Appl. Genet.* 133, 1365–1384. doi: 10.1007/s00122-019-03527-6

Zhang, K., Su, J., Xu, M., Zhou, Z., Zhu, X., Ma, X., et al. (2020b). A common wild rice-derived *BOC1* allele reduces callus browning in indica rice transformation. *Nat. Commun.* 11:443. doi: 10.1038/s41467-019-14265-0

Zhang, G. (2021). Target chromosome-segment substitution: a way to breeding by design in rice. *Crop J.* 9, 658–668. doi: 10.1016/j.cj.2021.03.001

Zhang, H., Zhao, Q., Sun, Z. Z., Zhang, C. Q., Feng, Q., Tang, S. Z., et al. (2011). Development and high-throughput genotyping of substitution lines carring the chromosome segments of indica 9311 in the background of japonica Nipponbare. *J. Genet. Genomics* 38, 603–611. doi: 10.1016/j.jgg.2011.11.004

Zhang, W., Tan, L., Sun, H., Zhao, X., Liu, F., Cai, H., et al. (2019d). 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, B., Liu, H., Qi, F., Zhang, Z., Li, Q., Han, Z., et al. (2019a). Genetic interactions among *Ghd7*, *Ghd8*, *OsPRR37* and *Hd1* contribute to large variation in heading date in rice. *Rice* 12:48. doi: 10.1186/s12284-019-0314-x

Zhang, B., Shang, L., Ruan, B., Zhang, A., Yang, S., Jiang, H., et al. (2019b). Development of three sets of high-throughput genotyped rice chromosome segment substitution lines and QTL mapping for eleven traits. *Rice* 12:33. doi: 10.1186/s12284-019-0293-y

Zhang, H., Zhou, L., Xu, H., Wang, L., Liu, H., Zhang, C., et al. (2019c). The *qSAC3* locus from indica rice effectively increases amylose content under a variety of conditions. *BMC Plant Biol.* 19:275. doi: 10.1186/s12870-019-1860-5

Zhang, X., Zhou, S., Fu, Y., Su, Z., Wang, X., and Sun, C. (2006). Identification of a drought tolerant introgression line derived from Dongxiang common wild rice (*O. rufipogon* Griff.). *Plant Mol. Biol.* 62, 247–259. doi: 10.1007/s11103-006-9018-x

Zhang, Y., Tan, L., Zhu, Z., Yuan, L., Xie, D., and Sun, C. (2015). *TOND1* confers tolerance to nitrogen deficiency in rice. *Plant J.* 81, 367–376. doi: 10.1111/tpj.12736

Zhang, Y., Zhou, J., Li, J., Yang, Y., Xu, P., and Tao, D. (2018). Mapping of *S56(t)* responsible for interspecific hybrid sterility between *Oryza sativa* and *Oryza glumaepatula*. *Breed. Sci.* 68, 242–247. doi: 10.1270/jsbbs.17116

Zhang, Y., Zhou, J., Xu, P., Li, J., Deng, X., Deng, W., et al. (2022). A genetic resource for rice improvement: introgression library of agronomic traits for all AA genome *Oryza* species. *Front. Plant Sci.* 13:856514. doi: 10.3389/fpls.2022.856514

Zhang, Z., Li, J., Pan, Y., Li, J., Zhou, L., Shi, H., et al. (2017). Natural variation in *CTB4a* enhances rice adaptation to cold habitats. *Nat. Commun.* 8:14788. doi: 10.1038/ncomms14788

Zhao, C., Zhou, L., Yu, X., Zhao, Q., and Wang, C. (2012). Identification of tiller angle quantitative trait loci based on chromosome segment substituted lines and high-density physical map in rice. *Chin. Bull. Bot.* 47, 594–601. doi: 10.3724/sp.j.1259.2012.00594

Zhao, D. S., Li, Q. F., Zhang, C. Q., Zhang, C., Yang, Q. Q., Pan, L. X., et al. (2018). *GS9* acts as a transcriptional activator to regulate rice grain shape and appearance quality. *Nat. Commun.* 9:1240. doi: 10.1038/s41467-018-03616-y

Zhao, D.-S., Liu, J.-Y., Ding, A.-Q., Zhang, T., Ren, X.-Y., Zhang, L., et al. (2021). Improving grain appearance of erect-panicle japonica rice cultivars by introgression of the null *gs9* allele. *J. Integr. Agric.* 20, 2032–2042. doi: 10.1016/S2095-3119(21)63659-6

Zhao, J., Wang, S., Qin, J., Sun, C., and Liu, F. (2020). The lipid transfer protein OsLTPL159 is involved in cold tolerance at the early seedling stage in rice. *Plant Biotechnol. J.* 18, 756–769. doi: 10.1111/pbi.13243

Zhao, L., Lei, J., Huang, Y., Zhu, S., Chen, H., Huang, R., et al. (2016). Mapping quantitative trait loci for heat tolerance at anthesis in rice using chromosomal segment substitution lines. *Breed. Sci.* 66, 358–366. doi: 10.1270/jsbbs.15084

Zhao, L., Zhou, H., Lu, L., Liu, L., Li, X., Lin, Y., et al. (2009). Identification of quantitative trait loci controlling rice mature seed culturability using chromosomal segment substitution lines. *Plant Cell Rep.* 28, 247–256. doi: 10.1007/s00299-008-0641-7

Zhao, Z. G., Zhu, S. S., Zhang, Y. H., Bian, X. F., Wang, Y., Jiang, L., et al. (2011). Molecular analysis of an additional case of hybrid sterility in rice (*Oryza sativa* L.). *Planta* 233, 485–494. doi: 10.1007/s00425-010-1313-8

Zhou, H., Xia, D., Zhao, D., Li, Y., Li, P., Wu, B., et al. (2021). The origin of Wx^{la} provides new insights into the improvement of grain quality in rice. *J. Integr. Plant Biol.* 63, 878–888. doi: 10.1111/jipb.13011

Zhou, H., Yang, W., Ma, S., Luan, X., Zhu, H., Wang, A., et al. (2020). Unconditional and conditional analysis of epistasis between tillering QTLs based on single segment substitution lines in rice. *Sci. Rep.* 10:15912. doi: 10.1038/s41598-020-73047-7

Zhou, P. H., Tan, Y. F., He, Y. Q., Xu, C. G., and Zhang, Q. (2003). Simultaneous improvement for four quality traits of Zhenshan 97, an elite parent of hybrid rice, by molecular marker-assisted selection. *Theor. Appl. Genet.* 106, 326–331. doi: 10.1007/s00122-002-1023-0

Zhou, S. X., Tian, F., Zhu, Z. F., Fu, Y. C., Wang, X. K., and Sun, C. Q. (2006). Identification of quantitative trait loci controlling drought tolerance at seedling stage in Chinese Dongxiang common wild rice (*Oryza rufipogon* Griff.). *Yi Chuan Xue Bao* 33, 551–558. doi: 10.1016/s0379-4172(06)60084-x

Zhou, Y., Dong, G., Tao, Y., Chen, C., Yang, B., Wu, Y., et al. (2016). Mapping quantitative trait loci associated with toot traits using sequencing-based genotyping chromosome segment substitution lines derived from 9311 and Nipponbare in rice (*Oryza sativa* L.). *PLoS One* 11:e0151796. doi: 10.1371/journal.pone.0151796

Zhou, Y., Tao, Y., Yuan, Y., Zhang, Y., Miao, J., Zhang, R., et al. (2018). Characterisation of a novel quantitative trait locus, *GN4-1*, for grain number and yield in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 131, 637–648. doi: 10.1007/s00122-017-3025-y

Zhou, Y., Xie, Y., Cai, J., Liu, C., Zhu, H., Jiang, R., et al. (2017). Substitution mapping of QTLs controlling seed dormancy using single segment substitution lines derived from multiple cultivated rice donors in seven cropping seasons. *Theor. Appl. Genet.* 130, 1191–1205. doi: 10.1007/s00122-017-2881-9

Zhu, H., Li, Y., Liang, J., Luan, X., Xu, P., Wang, S., et al. (2018). Analysis of QTLs on heading date based on single segment substitution lines in rice (*Oryza Sativa* L.). *Sci. Rep.* 8:13232. doi: 10.1038/s41598-018-31377-7

Zhu, W., Lin, J., Yang, D., Zhao, L., Zhang, Y., Zhu, Z., et al. (2009). Development of chromosome segment substitution lines derived from backcross between two sequenced rice cultivars, indica recipient 93-11 and japonica donor Nipponbare. *Plant Mol. Biol. Rep.* 27, 126–131. doi: 10.1007/s11105-008-0054-3

Zhu, Z., Tan, L., Fu, Y., Liu, F., Cai, H., Xie, D., et al. (2013). Genetic control of inflorescence architecture during rice domestication. *Nat. Commun.* 4:2200. doi: 10.1038/ncomms3200

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

ORIGINAL RESEARCH
published: 29 April 2022
doi: 10.3389/fpls.2022.890568

Development of Wide-Compatible *Indica* Lines by Pyramiding Multiple Neutral Alleles of *Indica–Japonica* Hybrid Sterility Loci

Jie Guo[1,2†], Yun Li[1†], Liang Xiong[1†], Tingxian Yan[1], Jinsong Zou[1], Ziju Dai[1], Guang Tang[1], Kangli Sun[1], Xin Luan[1], Weifeng Yang[1], Quanya Tan[1], Haitao Zhu[1], Ruizhen Zeng[1], Shaokui Wang[1] and Guiquan Zhang[1]*

[1] Guangdong Provincial Key Laboratory of Plant Molecular Breeding, State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangzhou, China, [2] Rice Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, China

Since the development of *indica* hybrid rice in the 1970s, great success has been achieved in hybrid rice production in China and around the world. The utilization of inter-subspecific *indica–japonica* hybrid rice has always been considered due to its stronger heterosis characteristics. However, *indica–japonica* hybrids face a serious problem of sterility, which hinders the exploitation of their heterosis. In the past decades, the genetic basis of *indica–japonica* hybrid sterility has been well studied. It was found that in sterile *indica–japonica* hybrids, female sterility was mainly controlled by the *S5* locus and male sterility by the *Sa*, *Sb*, *Sc*, *Sd*, and *Se* loci. In this study, we developed wide-compatible *indica* lines (WCILs) by pyramiding multiple neutral (n) alleles of the hybrid sterility loci. First, we identified S^n alleles of the loci in single-segment substitution lines (SSSLs) in the genetic background of *indica* Huajingxian 74 (HJX74). Then, the S^n alleles of *S5*, *Sb*, *Sc*, *Sd*, and *Se* loci in SSSLs were pyramided in the HJX74 genetic background. The WCILs carrying S^n alleles at the *S5*, *Sb*, *Sc*, *Sd*, and *Se* loci showed wide compatibility with *indica* and *japonica* rice varieties. Therefore, the WCILs will be used to develop inter-subspecific *indica–japonica* hybrid rice with normal fertility.

Keywords: hybrid rice, heterosis, hybrid sterility, neutral allele, breeding by design

OPEN ACCESS

Edited by:
Kenneth Lee McNally, International Rice Research Institute (IRRI), Philippines

Reviewed by:
Sibin Yu, Huazhong Agricultural University, China
Longbiao Guo, China National Rice Research Institute (CAAS), China

***Correspondence:**
Guiquan Zhang
gqzhang@scau.edu.cn

† These authors have contributed equally to this work

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

Received: *06 March 2022*
Accepted: *30 March 2022*
Published: *29 April 2022*

Citation:
Guo J, Li Y, Xiong L, Yan T, Zou J, Dai Z, Tang G, Sun K, Luan X, Yang W, Tan Q, Zhu H, Zeng R, Wang S and Zhang G (2022) Development of Wide-Compatible Indica Lines by Pyramiding Multiple Neutral Alleles of Indica–Japonica Hybrid Sterility Loci. Front. Plant Sci. 13:890568. doi: 10.3389/fpls.2022.890568

INTRODUCTION

Asian cultivated rice (*Oryza sativa* L.) is the staple food for more than half of the world's population (Fukagawa and Ziska, 2019). The breeding of high-yielding varieties is essential for maintaining global food security (Peng et al., 2008; Khush, 2013). Since the 1970s, *indica* hybrid rice has been successfully developed in China and around the world (Yuan and Virmani, 1988; Cheng et al., 2007). However, the heterosis of intra-subspecific hybrid rice is limited, resulting in a yield plateau

for production of hybrid rice (Peng et al., 2004; Cheng et al., 2007). There is great heterosis in inter-subspecific hybrids, and exploiting this heterosis has long been considered a promising approach to further increase the yield potential of rice (Khush, 2013; Zhang et al., 2021). However, the severe sterility associated with *indica–japonica* hybrid hinders the utilization of heterosis (Ikehashi and Araki, 1986; Ouyang and Zhang, 2018; Zhang, 2020, 2022).

The sterility of hybrids produced by crossing *indica* and *japonica* rice varieties can be attributed to female or embryo sac sterility and male or pollen sterility. The female sterility in hybrid is mainly controlled by the *S5* locus, which was mapped on chromosome 6 (Ikehashi and Araki, 1986; Yanagihara et al., 1995; Ji et al., 2005; Qiu et al., 2005). The male sterility in hybrid is mainly controlled by *Sa*, *Sb*, *Sc*, *Sd*, and *Se* loci (Zhang and Lu, 1989, 1993; Zhang et al., 1993, 1994). Using molecular markers, *Sa* was found to be located on chromosome 1 (Zhuang et al., 1999; Su and Liu, 2003), *Sb* on chromosome 5 (Zhuang et al., 2002; Li et al., 2006), *Sc* on chromosome 3 (Zhang and Zhang, 2001; Yang et al., 2004), *Sd* on chromosome 1 (Li et al., 2008), and *Se* on chromosome 12 (Zhu et al., 2008). The *S5*, *Sa*, and *Sc* genes were then cloned and functionally analyzed (Chen et al., 2008; Long et al., 2008; Yang et al., 2012; Shen et al., 2017). The genetic model of hybrid sterility is the one-locus sporo-gametophytic interaction model (Ikehashi and Araki, 1986; Zhang and Lu, 1993; Zhang, 2020). In this genetic model, it is assumed that *indica* varieties have S^i allele, and *japonica* varieties have S^j allele at the loci. At the *S5* locus, the interaction between S^i and S^j causes the abortion of female gametes carrying S^j allele (Ikehashi and Araki, 1986). At the *Sa*, *Sb*, *Sc*, *Sd*, and *Se* loci, the interaction between S^i and S^j causes the abortion of male gametes carrying S^j allele (Zhang and Lu, 1993). At these loci, some varieties carry S^n, a neutral allele, and the allelic interaction between S^i/S^n and S^j/S^n cannot cause the abortion of any gamete (Ikehashi and Araki, 1986; Zhang and Lu, 1993; Yang et al., 2012; Shen et al., 2017; Xie et al., 2017). The understanding of the genetic and molecular mechanisms of sterility in *indica–japonica* hybrids has laid the foundation for overcoming hybrid sterility.

With the development of molecular breeding technology, the concept of "breeding by design" was proposed (Peleman and van der Voort, 2003). To implement the strategy of rice breeding by design, a library of single-segment substitution lines (SSSLs) in rice was constructed by using 43 accessions from seven species of AA genome as donors of chromosome substitution segments in the genetic background of Huajingxian 74 (HJX74), an elite *indica* variety from south China. A total of 2,360 HJX74-SSSLs have been included in the library, which contains rich genetic resources for rice breeding techniques (Zhang et al., 2004; Xi et al., 2006; He et al., 2017; Zhao et al., 2019; Zhang, 2021). The HJX74-SSSL library was used as a platform for designing new rice cultivars, and several cytoplasmic male sterility (CMS), maintainer, and restorer lines were developed (Dai et al., 2015, 2016; Luan et al., 2019). Therefore, target chromosome-segment substitution is a way to breeding by design in rice (Zhang, 2021).

With the understanding of the genetic and molecular mechanisms of *indica–japonica* hybrid sterility and the development of molecular breeding techniques, the breeding strategies for developing inter-subspecific *indica–japonica* hybrid rice were proposed (Zhang, 2020, 2022). One strategy for overcoming the hybrid sterility of *indica–japonica* rice is to develop *indica*-compatible *japonica* lines (ICJLs) (Zhang and Lu, 1999; Zhang, 2020). Recently, the ICJLs were developed by pyramiding S^i allele at the *Sb*, *Sc*, *Sd,* and *Se* loci and S^n allele at the *S5* locus in *japonica* genetic background by marker-assisted selection (MAS). The ICJLs are compatible with *indica* but incompatible with *japonica* in pollen fertility and spikelet fertility (Guo et al., 2016). Another strategy for overcoming the hybrid sterility of *indica–japonica* rice is to develop wide-compatible *indica* lines (WCILs) (Zhang, 2020, 2022). Herein, we report the development of WCILs using the HJX74-SSSL library. By pyramiding S^n allele at the *S5*, *Sb*, *Sc*, *Sd*, and *Se* loci in the HJX74 genetic background, the obtained WCILs were compatible with both *indica* and *japonica* rice in pollen fertility and spikelet fertility. The breeding of WCILs provides a technique to develop inter-subspecific *indica–japonica* hybrid rice.

MATERIALS AND METHODS

Plant Materials and Field Trials

Seven SSSLs carrying the *Sc* gene for hybrid male sterility in their chromosome substitution segments and seven SSSLs carrying the *S5* gene for hybrid female sterility in their chromosome substitution segments were selected from the HXJ74-SSSL library (**Supplementary Table 1**). A set of *indica* and *japonica* varieties were used as testers to test the hybrid fertility. The genotypes of *Sa*, *Sb*, *Sc*, *Sd*, and *Se* loci for hybrid male sterility and *S5* locus for hybrid female sterility have been identified in some of the testers. It was found that at these six loci, the *indica* variety Guang-lu-ai 4 (GLA4) carried the S^i alleles, while the *japonica* variety Taichung 65 (T65) carried the S^j alleles (Zhang et al., 1994; Guo et al., 2016). All the study samples were planted from 2008 to 2019 at the farm of South China Agricultural University, Guangzhou (23°07′N, 113°15′E). These plants were planted in two cropping seasons each year, with the first cropping season (FCS) running from late February to mid-July and the second cropping season (SCS) running from late July to mid-November. Seeds were sown in seedbeds, and seedlings were transplanted into the field. Field management, including irrigation, fertilization, and pest control, followed normal agricultural practices.

Genotyping by Molecular Markers

The SSR markers were selected on the rice microsatellite maps (McCouch et al., 2002; Zhang et al., 2007). The functional markers of the *S5* gene were selected to identify the genotypes at the *S5* loci (Sundaram et al., 2010; Du et al., 2011; Yang et al., 2012; Guo et al., 2016). Markers linked with the *Sa, Sb, Sc, Sd, Se,* and *S5* loci were selected from the published studies (Yang et al., 2004, 2012; Li et al., 2006, 2008; Chen et al., 2008; Long et al., 2008; Zhu et al., 2008). New molecular markers were developed in this study (**Supplementary Table 2**). The PCR products were separated into 6% non-denaturing polyacrylamide gels (Panaud et al., 1996; Li et al., 2006).

Phenotyping of Fertility and Agronomic Traits

To check pollen fertility, nine mature flowers were collected from the upper third of panicles during the flowering stage and fixed in FAA solution. Pollens were stained with the 1% I_2-KI solution containing 0.1% (w/v) iodine and 1% (w/v) potassium iodide. Pollens were divided into normal pollens and sterile pollens, which were further divided into stained abortive pollens (stained but small size) and empty abortive pollens (small size and empty) (Zhang and Lu, 1989). Three panicles per plant and 10–12 plants per line were used to examine the spikelet fertility, and 20–40 plants per line were used to investigate the agronomic traits.

Statistical Analysis

For statistical analysis, the percentage data were converted to the square root of the arcsine values. Student's *t*-test was used to compare the data between the two groups. The Dunnett *t*-test was used to compare multiple groups with the control group. The least significance range (LSR) was used for the multiple range test among the multiple groups. The chi-square (χ^2) test was performed to detect the distorted segregation of three genotypes in F_2 populations according to the Mendelian ratio of 1:2:1. SPSS statistics 23.0 and Origin Pro 9.0 were used for data analysis and charting[1].

RESULTS

Genotypes of *Sa*, *Sb*, *Sc*, *Sd*, *Se*, and *S5* Loci in Huajingxian 74

To identify the genotypes of *Sa*, *Sb*, *Sc*, *Sd*, *Se*, and *S5* loci associated with hybrid sterility, HJX74 was test crossed with T65, a *japonica* variety with S^j alleles at these six loci, and GLA4, an *indica* variety with S^i alleles at these six loci (Zhang et al., 1994; Guo et al., 2016). The F_1 hybrids obtained from the cross of T65/GLA4 showed severe sterility, where the pollen fertility was only 20.51% and spikelet fertility was only 5.89%. In contrast, the F_1 hybrid of the HJX74/GLA4 cross showed normal pollen fertility and spikelet fertility of 92.39% and 93.39%, respectively. In the F_1 hybrids obtained from the cross of HJX74/T65, the pollen fertility was 72.89% and the spikelet fertility was 61.58%, which were significantly higher than those of T65/GLA4 and significantly lower than those of HJX74/GLA4 hybrids (**Figure 1A**). The results showed that the hybrid of HJX74/T65 exhibited partial pollen sterility and partial spikelet sterility.

The molecular markers linked to the *Sa*, *Sb*, *Sc*, *Sd*, *Se*, and *S5* loci were used to investigate genotype segregation in the F_2 populations obtained from the crosses of HJX74/GLA4 and HJX74/T65. At the *Sb*, *Sd*, and *Se* loci, the genotype segregation of F_2 populations from both crosses fit the Mendelian ratio of 1:2:1. At the *Sa*, *Sc*, and *S5* loci, distorted segregation of the genotypes was detected in the F_2 population of HJX74/T65 but not in the genotypes of HJX74/GLA4. At the *Sa* locus, the genotype ratios of Sa^{HJX74}/Sa^{HJX74}, Sa^{HJX74}/Sa^{T65}, and Sa^{T65}/Sa^{T65} were 68:100:33, which significantly distorted from the Mendelian ratio of 1:2:1. At the *Sc* locus, the genotype ratios of Sc^{HJX74}/Sc^{HJX74}, Sc^{HJX74}/Sc^{T65}, and Sc^{T65}/Sc^{T65} were 69:70:3, which significantly

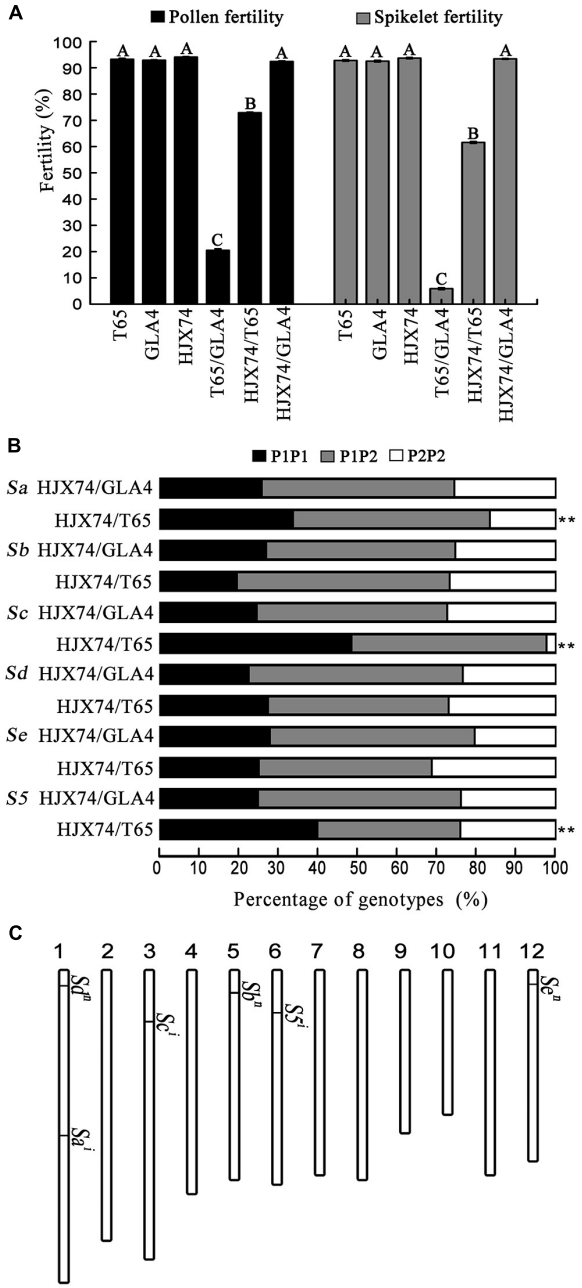

FIGURE 1 | Genotypes of hybrid sterility loci *Sa*, *Sb*, *Sc*, *Sd*, *Se*, and *S5* in HJX74. **(A)** Pollen and spikelet fertility of three parents and their F_1 hybrids. **(B)** Ratios of genotypes at *Sa*, *Sb*, *Sc*, *Sd*, *Se*, and *S5* loci in the F_2 populations from the crosses between HJX74 and testers Taichung 65 (T65) and Guang-lu-ai 4 (GLA4). **(C)** Chromosome location and genotypes of the genes at the *Sa*, *Sb*, *Sc*, *Sd*, *Se*, and *S5* loci in HJX74. Vertical bars represent rice chromosomes. P1P1, Genotypes of HJX74; P1P2, Heterozygous genotype; P2P2, Genotype of testers T65 (S^j/S^j) or GLA4 (S^i/S^i). Capital letters indicate statistical differences at the 0.01 probability level.

[1]https://www.originlab.com

distorted from the Mendelian ratio. Distorted segregation was also detected at the *S5* locus, where the genotype ratios of $S5^{HJX74}/S5^{HJX74}$, $S5^{HJX74}/S5^{T65}$, and $S5^{T65}/S5^{T65}$ were found to be 72:65:43 (**Figure 1B**). In addition, HJX74 was tested using a group of *indica* and *japonica* testers. The results showed that distorted segregation was detected only at the *Sa*, *Sc*, and *S5* loci in the crosses of HJX74/*japonica* testers (**Supplementary Table 3**).

These results indicated that HJX74 carried S^i/S^i at the *Sa*, *Sc*, and *S5* loci and S^n/S^n at the *Sb*, *Sd*, and *Se* loci (**Figure 1C**). At the *Sa* and *Sc* loci, the allele interaction between S^i of HJX74 and S^j of *japonica* testers caused the abortion of male gametes carrying S^j in hybrids, resulting in the significant reduction of plants with S^j/S^j in the F_2 populations. At the *S5* locus, the allele interaction between $S5^i$ of HJX74 and $S5^j$ of *japonica* testers caused the abortion of female gametes carrying $S5^j$ in hybrids, resulting in the significant reduction of plants with $S5^j/S5^j$ in the F_2 populations. At the *Sb*, *Sd*, and *Se* loci, allele interaction between S^n of HJX74 and S^j of *japonica* testers or S^i of *indica* testers could not cause the abortion of any gamete in hybrids, and genotype segregation in the F_2 populations fit the Mendelian ratio of 1:2:1 (**Supplementary Table 3**). In addition, compared with the *Sc* locus, the *Sa* locus showed weak distorted segregation, where $\chi^2_{(1:2:1)} = 34.00–62.27$ in the five segregation populations of the *Sc* locus, while $\chi^2_{(1:2:1)} = 9.36–12.19$ in the three segregation populations of the *Sa* locus (**Supplementary Table 3**). The results showed that the hybrid male sterility caused by the interaction between S^i and S^j at the *Sa* locus was weaker than that at the *Sc* locus.

Genotypes of the *Sc* Locus in the Substitution Segments of Single-Segment Substitution Lines

To screen the Sc^n gene, seven SSSLs carrying the *Sc* locus on the substitution segments obtained from different donors were selected from the HJX74-SSSL library (**Supplementary Table 1**). The pollen fertility of F_1 hybrids from the crosses between the SSSLs and HJX74 was over 90% (**Figure 2A**). The SSSLs were then tested with three *indica* testers and three *japonica* testers. Four SSSLs (01-03, 05-03, 06-03, and 14-03) and HJX74 showed significantly higher pollen fertility in their F_1 hybrids with *indica* testers than those obtained with *japonica* testers. In contrast, the other three SSSLs (11-03, 22-03, and 27-03) did not show a significant difference in the pollen fertility of F_1 hybrids between the crosses with *indica* and *japonica* testers (**Figure 2B**). Two SSSLs (06-03 and 27-03) were then selected to detect the segregation of *Sc* genotypes in F_2 populations obtained from the crosses with T65. In the F_2 population of the T65/27-03 cross, the *Sc* genotypes of T65/T65, T65/27-03, and 27-03/27-03 segregated in the ratios of 40:83:53, which fit the Mendelian ratio of 1:2:1. In contrast, in the F_2 population of the T65/06-03 cross, the genotype ratios of T65/T65, T65/06-03, and 06-03/06-03 were 23:64:69, which significantly distorted from the Mendelian ratio (**Figure 2C**). These results indicated that at the *Sc* locus, SSSLs 11-03, 22-03, and 27-03 carried the S^n allele, while 01-03, 05-03, 06-03, and 14-03 carried the S^i allele.

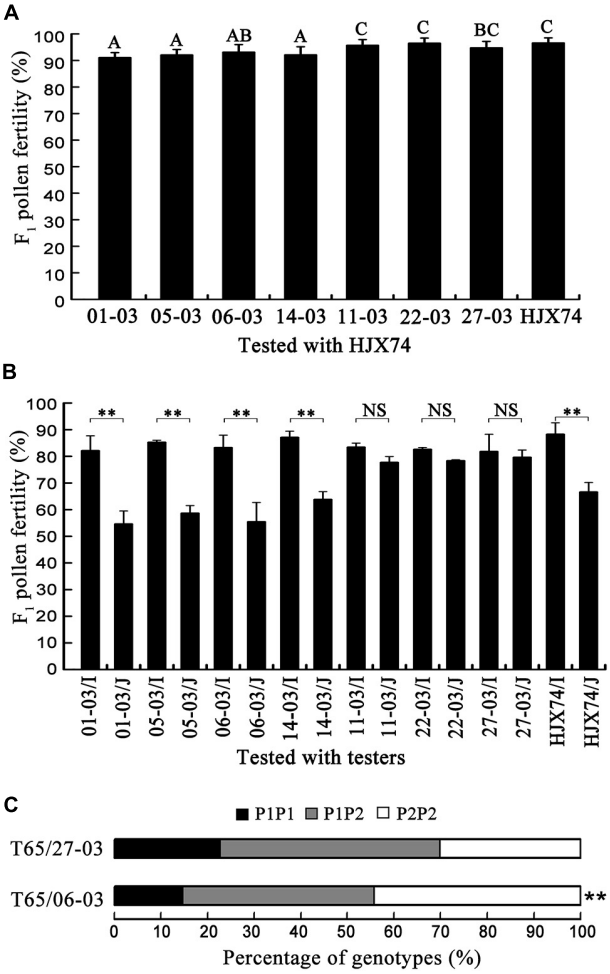

FIGURE 2 | Genotypes and genetic effects of the *Sc* gene on chromosomal substitution segments of SSSLs with the HJX74 genetic background. **(A)** Pollen fertility of F_1 hybrids from crosses between SSSLs and HJX74. **(B)** Pollen fertility of F_1 hybrids from crosses between SSSLs and testers. I, *indica* testers; J, *japonica* testers. **(C)** Segregation ratios of *Sc* genotypes in F_2 populations from crosses between SSSLs and T65. P1P1, Genotype of T65 (S^i/S^i); P1P2, Heterozygous genotypes; P2P2, Genotypes of SSSLs. **, Significant difference at 0.01 probability level. NS, No significance.

Genotypes of the *S5* Locus in the Substitution Segments of Single-Segment Substitution Lines

To screen the $S5^n$ gene, seven SSSLs carrying the *S5* locus in the substitution segments obtained from different donors were selected from the HJX74-SSSL library (**Supplementary Table 1**). The genotypes of the *S5* locus in the SSSLs were detected by functional markers. The results showed that in the substitution segments, three SSSLs (04-06, 13-06, and 14-06) carried $S5^i$, one SSSL (10-06) carried $S5^j$, and the other three SSSLs (21-06, 23-06, and 27-06) carried $S5^n$ (**Supplementary Table 4**).

Five genotypes of the *S5* locus were obtained from the F_1 hybrids crossed by seven SSSLs (**Supplementary Table 5**). The pollen fertility of hybrids was normal in all crosses, ranging from

93.04 to 94.42%. The spikelet fertility of $S5^i/S5^i$, $S5^i/S5^n$, $S5^n/S5^j$, and $S5^n/S5^n$ genotypes was normal (from 89.84% to 91.28%), but that of $S5^i/S5^j$ genotype from the crosses between 10-06 carrying $S5^j/S5^j$ and SSSLs carrying $S5^i/S5^i$ was only 68.34%, which was significantly lower than the spikelet fertility of the other four genotypes (**Figure 3A** and **Supplementary Table 5**). The segregation of *S5* genotypes in F_2 populations obtained from three heterozygous genotypes, $S5^n/S5^j$, $S5^n/S5^i$, and $S5^j/S5^i$, was detected by using the functional markers of the *S5* gene. Distorted segregation was detected in the $S5^j/S5^i$ segregation population produced from the crosses between 10-06 carrying $S5^j/S5^j$ and SSSLs carrying $S5^i/S5^i$, but was not detected in the segregation populations of $S5^n/S5^j$ from 21-06/10-06 and of $S5^n/S5^i$ from 21-06/13-06 (**Figure 3B**).

The three SSSLs with $S5^n$, 21-06, 23-06, and 27-06, were tested for their wide compatibility by crossing with *indica* and *japonica* testers. The F_1 hybrids from all crosses showed high spikelet fertility, from 80.28% to 95.05%. As a control, the spikelet fertility of F_1 hybrids in HJX74/*japonica* testers was 70.13% (**Figure 3C**). In addition, distorted segregation of the *S5* locus was detected in the F_2 population of T65/04-06, but was not detected in the F_2 populations of the other three crosses, that is, GLA4/04-06, T65/23-06, and GLA4/23-06 (**Figure 3D**). These results showed that the three SSSLs (21-06, 23-06, and 27-06) were compatible with *indica* testers and *japonica* testers in spikelet fertility as a result of their carrying $S5^n$ locus.

Pyramiding of S^n Alleles at the *Sc* and *S5* Loci in the Huajingxian 74 Genetic Background

Three SSSLs (11-03, 22-03, and 27-03) with the Sc^n gene and three SSSLs (21-06, 23-06, and 27-06) with the $S5^n$ gene were selected to pyramid the two S^n genes in the HJX74 genetic background. Three SSSLs with Sc^n were crossed with three SSSLs with $S5^n$, respectively. In the segregating populations, the plants carrying Sc^n and $S5^n$ loci were selected. Nine pyramiding lines were developed, which carried Sc^n and $S5^n$ loci from different donors and Sb^n, Sd^n, and Se^n in the HJX74 genetic background (**Figure 4A** and **Supplementary Table 6**). Therefore, the nine pyramiding lines thus obtained were WCILs.

In the nine WCILs, the plant type was similar to HJX74 (**Figure 4B**). In addition, no significant difference between HJX74 and WCILs was found in the majority of the investigated traits, including heading date, plant height, width of flag leaf, length of flag leaf, grain length, grain width, and grain weight (**Supplementary Table 7**).

Compatibility of Wide-Compatible *indica* Lines

To evaluate the compatibility of nine WCILs, the WCILs were test crossed with six *indica* testers and five *japonica* testers (**Supplementary Tables 8–11**). When tested with *indica* tester group, F_1 hybrids of nine WCILs showed normal pollen fertility and spikelet fertility, with no significant difference when compared to HJX74. When tested with the *japonica* tester group, nine WCILs showed significantly higher F_1 pollen fertility and spikelet fertility when compared to HJX74 (**Figure 5**). These results indicated that the WCILs showed wide compatibility, producing high pollen fertility and spikelet fertility in their F_1 hybrids with both *indica* and *japonica* rice varieties.

Although the WCILs showed significantly higher F_1 pollen fertility when tested with *japonica* testers, the F_1 pollen fertility was still lower when tested with *indica* testers (**Figures 5C,D**). To identify the problem, the genotype segregation at the *Sa, Sb, Sc, Sd, Se*, and *S5* loci in F_2 populations was examined with molecular markers linked with these loci. No distorted segregation at the *Sb*, *Sc*, *Sd*, *Se*, and *S5* loci was found in all the detected F_2 populations, further confirming the fact that the alleles of the *Sb*, *Sc*, *Sd*, *Se*, and *S5* loci in WCILs were S^n. However, significantly distorted segregation was detected at the *Sa* locus in all four populations (**Supplementary Table 12**). These results verified that WCILs carried the Sa^i gene in the HJX74 genetic background, and the interaction between Sa^i from WCILs and Sa^j from *japonica* testers caused some male gametes with Sa^j to become abortive in F_1 hybrids obtained from the crosses of WCILs with *japonica* testers.

Three *indica* lines, GLA4 carrying the genotype of Sa^i, Sb^i, Sc^i, Sd^i, Se^i, and $S5^i$, HJX74 carrying the genotype of Sa^i, Sb^n, Sc^i, Sd^n, Se^n, and $S5^i$, and WCIL 2223 carrying the genotype of Sa^i, Sb^n, Sc^n, Sd^n, Se^n, and $S5^n$, were selected to test their compatibility with eight *japonica* varieties of different ecotypes. In the F_1 hybrids of GLA4 with eight *japonica* varieties, pollen fertility was 13.24–90.68% with an average of 46.94%, and spikelet fertility was 5.89–92.95% with an average of 44.40%. In the F_1 hybrids of HJX74 with eight *japonica* varieties, pollen fertility was 75.19–95.74% with an average of 85.37%, and spikelet fertility was 58.34–93.90% with an average of 74.69%. On comparison of data, pollen fertility was 82.35–95.79% with an average of 88.82%, and spikelet fertility was 89.19–94.29% with an average of 91.73% in the F_1 hybrids of WCIL 2223 with the eight *japonica* varieties (**Supplementary Table 13**). The results showed that WCIL 2223 had higher and wider compatibility with *japonica* varieties than GLA4 and HJX74. The pollen fertility and spikelet fertility in F_1 hybrids of WCIL with various *japonica* varieties were normal or near normal.

DISCUSSION

Sterility or Compatibility of Hybrids Between *indica* and *japonica* Subspecies Is a Complex Trait

In the past decades, the genetic basis of *indica–japonica* hybrid sterility has been understood. In *indica–japonica* hybrid sterility, the *S5* locus was found to be responsible for female sterility, and the *Sa*, *Sb*, *Sc*, *Sd*, and *Se* loci were responsible for male sterility. Following the tri-allele pattern and the one-locus sporo-gametophytic interaction model, the allele interaction between S^i and S^j leads to the abortion of male or female gametes carrying S^j, whereas the allele interaction between S^n and S^i or S^j does not lead to the abortion of any gamete (Zhang, 2020, 2022). Thus, the sterility or compatibility of hybrids between *indica*

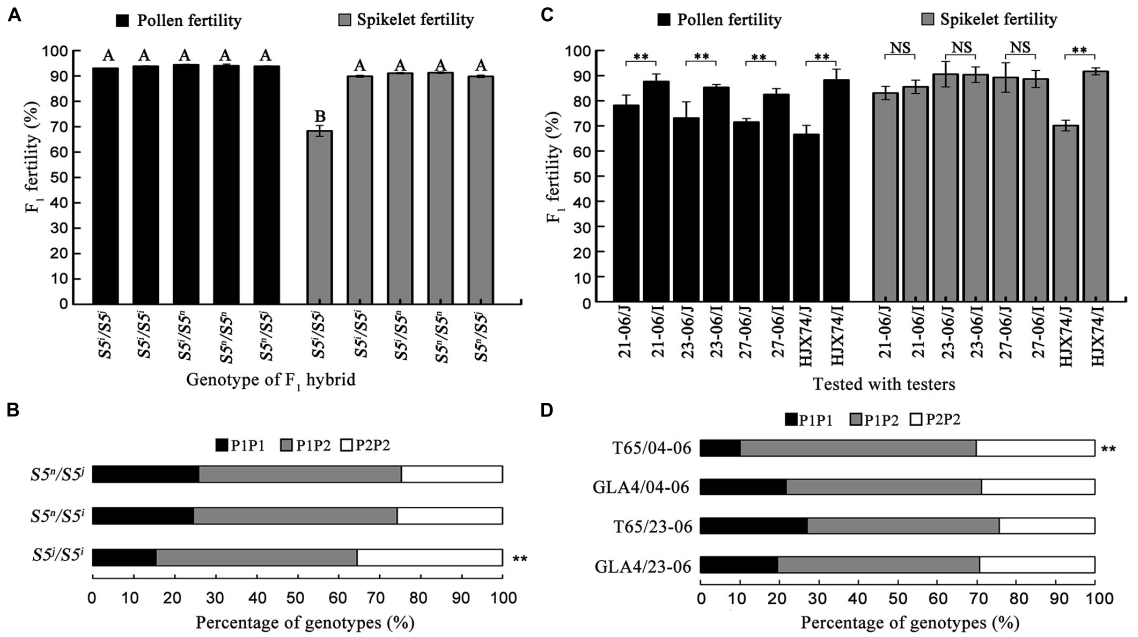

FIGURE 3 | Genotypes and genetic effects of the *S5* gene on chromosomal substitution segments of SSSLs with the HJX74 genetic background. **(A)** Pollen fertility and spikelet fertility in F_1 hybrids from the crosses between SSSLs and different *S5* genotypes. **(B)** Segregation ratios of *S5* genotypes in F_2 populations from different crosses. **(C)** Pollen fertility and spikelet fertility in F_1 hybrids from crosses between SSSLs (21-06, 23-06, and 27-06) and testers. I, *indica* testers; J, *japonica* testers. Capital letters indicate statistical differences at the 0.01 probability level. **(D)** Segregation ratios of *S5* genotypes in F_2 populations from crosses between SSSLs and testers. P1P1, Genotype of tester; P1P2, Heterozygous genotypes; P2P2, Genotypes of SSSLs. **, Significant difference at 0.01 probability level. NS, No significance.

and *japonica* subspecies is a complex trait that is controlled by multiple genes. Due to the diversity of *indica* and *japonica* rice varieties, the genotypes of hybrid sterility vary greatly among different varieties, particularly modern varieties, resulting in different crossing combinations with different degrees of hybrid sterility. In addition, the effects of alleles obtained from different donors are quantitatively different, resulting in the continuous variation of fertility at a single locus (Zhang et al., 1993, 1994). The molecular basis of allele diversity has been revealed by the cloned genes of *S5* (Chen et al., 2008; Yang et al., 2012), *Sa* (Long et al., 2008; Xie et al., 2017), and *Sc* (Shen et al., 2017). In this study, we found that HJX74, the recipient of SSSLs, carried the S^n allele at the *Sb*, *Sd*, and *Se* loci but the S^i allele at the *S5*, *Sa*, and *Sc* loci (**Figure 1**). In addition, the effect of Sa^i was weaker than that of Sc^i in HJX74 (**Supplementary Table 3**). The identification of genotypes that lead to hybrid sterility provided a prerequisite for improving the compatibility of HJX74.

Hybrid Sterility in *indica–japonica* Rice Can Be Overcome by Developing ICJLs and Wide-Compatible *indica* Lines

Based on the tri-allele pattern and the one-locus sporo-gametophytic interaction model, the *indica–japonica* hybrid sterility can be overcome by developing ICJLs and WCILs (Zhang and Lu, 1999; Zhang, 2020, 2022). ICJLs can be developed by transferring the S^i allele from *indica* to *japonica* rice. In hybrids between *indica* varieties having S^i allele and ICJLs having S^i allele in *japonica* genetic background, the S^i/S^i genotype cannot cause the abortion of any gamete. In a previous study, we transferred the S^i allele from *indica* donors to the *japonica* T65 variety to develop ICJLs, which carry the S^i allele at hybrid sterility loci in the *japonica* genetic background. The result was that ICJLs were compatible with *indica* but incompatible with *japonica* rice (Guo et al., 2016). In another method, WCILs can be developed by transferring the S^n allele from donors to *indica* rice. In hybrids between WCILs having S^n allele in *indica* genetic background and *japonica* varieties having S^j allele, the S^n/S^j genotype cannot cause the abortion of any gamete. In this study, we pyramided the S^n allele of SSSLs to develop WCILs, which carry S^n allele in *indica* HJX74 genetic background. The result was that WCILs showed wide compatibility, which was compatible with both *indica* and *japonica* rice varieties (**Figures 4**, **5**). These results showed that the breeding of ICJLs and WCILs is practicable and that the *indica–japonica* hybrid sterility could be overcome by using ICJLs and WCILs.

The Single-Segment Substitution Line Library Is a Powerful Platform for Developing Wide-Compatible *indica* Lines

The development of WCILs requires pyramiding S^n alleles of multiple hybrid sterility loci to improve compatibility. The breeding of WCILs is a challenging task because it is a time-consuming and laborious technique. First, the S^n alleles of the *S5*,

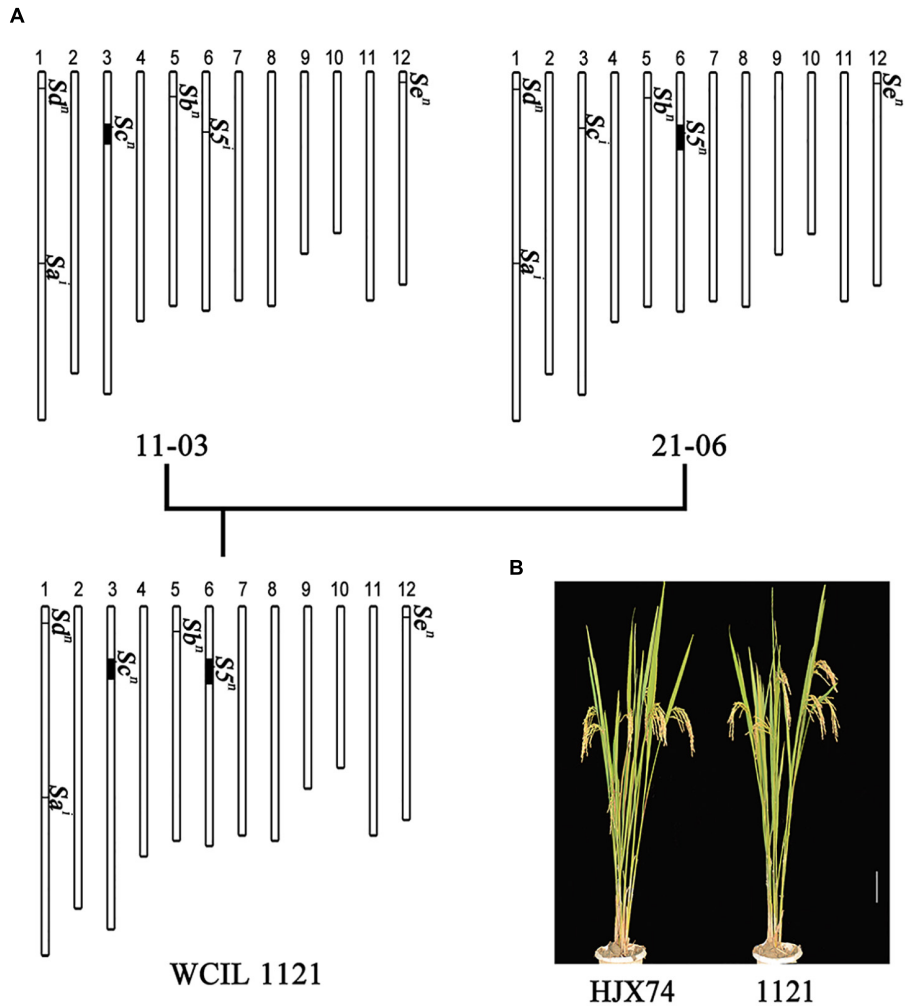

FIGURE 4 | Pyramiding of S^n genes at the *Sc* and *S5* loci in the HJX74 genetic background. **(A)** Development of WCIL 1121 by pyramiding of Sc^n in the substitution segment of SSSL 11-03 and $S5^n$ in the substitution segment of SSSL 21-06. Scale bar, 2 cm. Vertical bars represent rice chromosomes. Deep parts represent the substitution segments from donors and light parts represent the genetic background of HJX74. **(B)** Plant types of WCIL 1121 and HJX74. Scale bar, 10 cm.

Sa, *Sb*, *Sc*, *Sd*, and *Se* loci need to be identified and selected from a wide range of genetic resources. Second, S^n alleles of multiple loci need to be pyramided in an *indica* genetic background by MAS. In addition, WCILs need to have improved traits to be used as parents of *indica–japonica* hybrid rice. Over the past two decades, we have constructed a HJX74-SSSL library, which is used as a platform for rice design (Zhang, 2021). Using this platform, a series of CMS, maintainer, and restorer lines were developed (Dai et al., 2015, 2016; Luan et al., 2019). In this study, we identified S^n alleles at the *S5*, *Sb*, *Sc*, *Sd*, and *Se* loci from the HJX74-SSSL library. Since HJX74, the recipient of SSSLs, carried the S^n alleles at *Sb*, *Sd*, and *Se* loci, but the S^i alleles at *S5*, *Sa*, and *Sc* loci, the SSSLs carrying $S5^n$ or Sc^n alleles were selected from the HJX74-SSSL library (**Figures 2**, **3**). The Sc^n and $S5^n$ of the SSSLs were then pyramided in the HJX74 genetic background. Nine WCILs carrying S^n alleles at the *S5*, *Sb*, *Sc*, *Sd*, and *Se* loci in the HJX74 genetic background were developed (**Figures 4**, **5**). The results show that the HJX74-SSSL library is a powerful platform for developing WCILs possessing the complex trait of wide compatibility.

Wide-Compatible *indica* Lines Will Be Used to Develop *indica–japonica* Hybrid Rice

It is believed that inter-subspecific hybrids have stronger heterosis than intra-subspecific hybrids (Fu et al., 2014; Birchler, 2015). Therefore, the exploitation of inter-subspecific heterosis for the production of improved rice varieties has long been considered (Cheng et al., 2007; Zhang, 2020). The main obstacle in utilizing inter-subspecific heterosis in rice is the *indica–japonica* hybrid sterility. In this study, WCILs were developed using the HJX74-SSSL platform. The WCILs had compatibility with a wide range of *japonica* varieties (**Figure 5** and **Supplementary Table 13**). Therefore, the development of WCILs is an effective approach to overcoming the problem

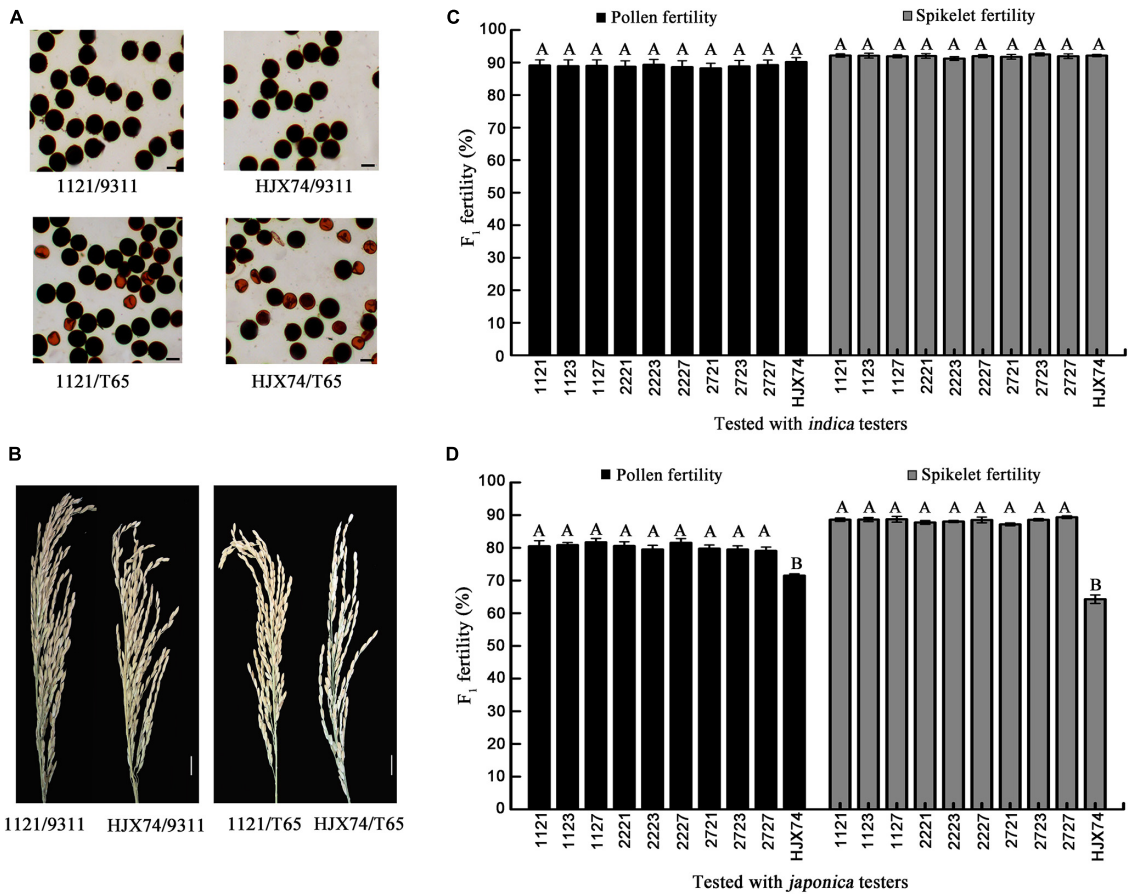

FIGURE 5 | Compatibility of WCILs. **(A)** Pollen grains stained by I_2-KI solution in F_1 hybrids of four crosses. Scale bar, 30 μm. **(B)** Spikelet fertility of the panicles in F_1 hybrids. Scale bar, 2 cm. **(C)** Pollen fertility and spikelet fertility of the F_1 hybrids from the crosses between WCILs (HJX74 as control) and *indica* testers (See **Supplementary Tables 8–9**). **(D)** Pollen fertility and spikelet fertility of the F_1 hybrids from the crosses between WCILs (HJX74 as control) and *japonica* testers (See **Supplementary Tables 10–11**). The information about nine WCILs (1121, 1123, 1127, 2221, 2223, 2227, 2721, 2723, and 2727) is given in **Supplementary Table 6**. HJX74 is a recipient of SSSLs. The 9311 is an *indica* tester. T65 is a *japonica* tester. Capital letters indicate statistical differences at the 0.01 probability level.

of *indica–japonica* hybrid sterility in breeding practice. By further improving their fertility restoration ability, WCILs can be improved to produce wide-compatible *indica* restorer lines (WCIRLs). Using the HJX74-SSSL platform, a series of WCIRLs is being developed and will be used to develop *indica–japonica* hybrid rice by crossing with *japonica* male sterile lines. Therefore, it is expected that *indica–japonica* hybrid rice will be the rice of next generation (Zhang, 2022).

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

GZ designed and supervised the work, analyzed the data and wrote the manuscript. JG, YL, and LX performed most of the experiments and compiled the experimental data. TY, JZ, ZD, GT, KS, XL, WY, and QT conducted a part of the experiments. HZ, RZ, and SW prepared the experimental materials and supervised some experiments. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the major program of transgenic new variety breeding in China (grant no. 2009ZX08009005) and the National Natural Science Foundation of China (grant nos. 91435207 and 31801449).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.890568/full#supplementary-material

REFERENCES

Birchler, J. A. (2015). Heterosis: the genetic basis of hybrid vigour. *Nat. Plants* 1:15020. doi: 10.1038/nplants.2015.20

Chen, J., Ding, J., Ouyang, Y., Du, H., Yang, J., Cheng, K., et al. (2008). A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of indica-japonica hybrids in rice. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11436–11441. doi: 10.1073/pnas.0804761105

Cheng, S., Cao, L., Zhuang, J., Chen, S., Zhan, X., Fan, Y., et al. (2007). Super hybrid rice breeding in China: achievements and prospects. *J. Integr. Plant Biol.* 49, 805–810. doi: 10.1111/j.1744-7909.2007.00514.x

Dai, Z., Lu, Q., Luan, X., Cai, J., Zhu, H., Liu, Z., et al. (2015). Development of a platform for breeding by design of CMS lines based on an SSSL library in rice (*Oryza sativa* L.). *Euphytica* 205, 63–72. doi: 10.1007/s10681-015-1384-5

Dai, Z., Lu, Q., Luan, X., Ouyang, L., Guo, J., Liang, J., et al. (2016). Development of a platform for breeding by design of CMS restorer lines based on an SSSL library in rice (*Oryza sativa* L.). *Breed. Sci.* 66, 768–775. doi: 10.1270/jsbbs.16044

Du, H., Ouyang, Y., Zhang, C., and Zhang, Q. (2011). Complex evolution of S5, a major reproductive barrier regulator, in the cultivated rice *Oryza sativa* and its wild relatives. *New Phytol.* 191, 275–287. doi: 10.1111/j.1469-8137.2011.03691.x

Fu, D., Xiao, M., Hayward, A., Fu, Y., Liu, G., Jiang, G., et al. (2014). Utilization of crop heterosis: a review. *Euphytica* 197, 161–173. doi: 10.1007/s10681-014-1103-7

Fukagawa, N. K., and Ziska, L. H. (2019). Rice: importance for global nutrition. *J. Nutr. Sci. Vitaminol.* 65, S2–S3. doi: 10.3177/jnsv.65.S2

Guo, J., Xu, X., Li, W., Zhu, W., Zhu, H., Liu, Z., et al. (2016). Overcoming inter-subspecific hybrid sterility in rice by developing indica-compatible japonica lines. *Sci. Rep.* 6:26878. doi: 10.1038/srep.26878

He, N., Wu, R., Pan, X., Peng, L., Sun, K., Zou, T., et al. (2017). Development and trait evaluation of chromosome single-segment substitution lines of *O. meridionalis* in the background of *O. sativa*. *Euphytica* 213:281. doi: 10.1007/s10681-017-2072-4

Ikehashi, H., and Araki, H. (1986). "Genetics of F1 sterility in remote crosses in rice," in *Proceedings of the 1st Rice Genetics Symposium: Rice Genetics* (Manila: International Rice Research Institute), 119–130. doi: 10.1142/9789812814265_0011

Ji, Q., Lu, J., Chao, Q., Gu, M., and Xu, M. (2005). Delimiting a rice wide-compatibility gene S5n to a 50 kb region. *Theor. Appl. Genet.* 111, 1495–1503. doi: 10.1007/s00122-005-0078-0

Khush, G. S. (2013). Strategies for increasing the yield potential of cereals: case of rice as an example. *Plant Breed.* 132, 433–436. doi: 10.1111/pbr.1991

Li, W., Zeng, R., Zhang, Z., Ding, X., and Zhang, G. (2006). Fine mapping of locus S-b for F1 pollen sterility in rice (*Oryza sativa* L.). *Chin. Sci. Bull.* 51, 675–680. doi: 10.1007/s11434-006-0675-6

Li, W., Zeng, R., Zhang, Z., Ding, X., and Zhang, G. (2008). Identification and fine mapping of S-d, a new locus conferring the partial pollen sterility of intersubspecific F1 hybrids in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 116, 915–922. doi: 10.1007/s00122-008-0723-5

Long, Y., Zhao, L., Niu, B., Su, J., Wu, H., Chen, Y., et al. (2008). Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18871–18876. doi: 10.1073/pnas.0810108105

Luan, X., Dai, Z., Yang, W., Tan, Q., Lu, Q., Guo, J., et al. (2019). Breeding by design of CMS lines on the platform of SSSL library in rice. *Mol. Breed.* 39:126. doi: 10.1007/s11032-019-1028-x

McCouch, S. R., Teytelman, L., Xu, Y., Lobos, K. B., Clare, K., Walton, M., et al. (2002). Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res.* 9, 199–207. doi: 10.1093/dnares/9.6.199

Ouyang, Y., and Zhang, Q. (2018). The molecular and evolutionary basis of reproductive isolation in plants. *J. Genet. Genomics* 45, 613–620. doi: 10.1016/j.jgg.2018.10.004

Panaud, O., Chen, X., and McCouch, S. R. (1996). Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* 252, 597–607. doi: 10.1007/bf02172406

Peleman, J. D., and van der Voort, J. R. (2003). Breeding by design. *Trends Plant Sci.* 8, 330–334. doi: 10.1016/S1360-1385(03)00134-1

Peng, S., Khush, G. S., Virk, P., Tang, Q., and Zou, Y. (2008). Progress in ideotype breeding to increase rice yield potential. *Field Crops Res.* 108, 32–38. doi: 10.1016/j.fcr.2008.04.001

Peng, S., Laza, R. C., Visperas, R. M., Khush, G. S., Virk, P., and Zhu, D. (2004). "Rice: progress in breaking the yield ceiling "new directions for a diverse planet"," in *Proceedings of the 4th International Crop Science Congress*, Brisbane, Qld.

Qiu, S. Q., Liu, K., Jiang, J. X., Song, X., Xu, C. G., Li, X. H., et al. (2005). Delimitation of the rice wide compatibility gene S5n to a 40-kb DNA fragment. *Theor. Appl. Genet.* 111, 1080–1086. doi: 10.1007/s00122-005-0033-0

Shen, R., Wang, L., Liu, X., Wu, J., Jin, W., Zhao, X., et al. (2017). Genomic structural variation-mediated allelic suppression causes hybrid male sterility in rice. *Nat. Commun.* 8:1310. doi: 10.1038/s41467-017-01400-y

Su, J., and Liu, Y. (2003). Fine mapping and cloning of the gene S-a for F1 pollen sterility in cultivated rice (*Oryza sativa* L.). *Mol. Plant Breed.* 1, 757–758.

Sundaram, R. M., Sakthivel, K., Hariprasad, A. S., Ramesha, M. S., Viraktamath, B. C., Neeraja, C. N., et al. (2010). Development and validation of a PCR-based functional marker system for the major wide-compatible gene locus S5 in rice. *Mol. Breed.* 26, 719–727. doi: 10.1007/s11032-010-9482-5

Xi, Z., He, F., Zeng, R., Zhang, Z., Ding, X., Li, W., et al. (2006). Development of a wide population of chromosome single-segment substitution lines in the genetic background of an elite cultivar of rice (*Oryza sativa* L.). *Genome* 49, 476–484. doi: 10.1139/g06-005

Xie, Y., Niu, B., Long, Y., Li, G., Tang, J., Zhang, Y., et al. (2017). Suppression or knockout of SaF/SaM overcomes the Sa-mediated hybrid male sterility in rice. *J. Integr. Plant Biol.* 59, 669–679. doi: 10.1111/jipb.12564

Yanagihara, S., Mccouch, S. R., Ishikawa, K., Ogi, Y., Maruyama, K., and Ikehashi, H. (1995). Molecular analysis of the inheritance of the S-5 locus, conferring wide compatibility in indica/japonica hybrids of rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 90, 182–188. doi: 10.1007/BF00222200

Yang, C., Chen, Z., Zhuang, C., Mei, M., and Liu, Y. (2004). Genetic and physical fine-mapping of the Sc locus conferring indica-japonica hybrid sterility in rice (*Oryza sativa* L.). *Chin. Sci. Bull.* 49, 1718–1721. doi: 10.1007/BF03184305

Yang, J., Zhao, X., Cheng, K., Du, H., Ouyang, Y., Chen, J., et al. (2012). A killer-protector system regulates both hybrid sterility and segregation distortion in rice. *Science* 337, 1336–1340. doi: 10.1126/science.1223702

Yuan, L., and Virmani, S. (1988). "Status of hybrid rice research and development," in *Hybrid Rice* (Manila: Internationa Rice Research Institute), 7–24.

Zhang, G. (2020). Prospects of utilization of inter-subspecific heterosis between indica and japonica rice. *J. Integr. Agr.* 19, 1–10. doi: 10.1016/S2095-3119(19)62843-1

Zhang, G. (2021). Target chromosome-segment substitution: a way to breeding by design in rice. *Crop J.* 9, 658–668. doi: 10.1016/j.cj.2021.03.001

Zhang, G. (2022). The next generation of rice: inter-subspecific indica-japonica hybrid rice. *Front. Plant Sci.* 13:857896. doi: 10.3389/fpls.2022.857896

Zhang, G., and Lu, Y. (1989). Genetic studies of the hybrid sterility in cultivated rice (*Oryza sativa*). I. Diallel analysis of the hybrid sterility among isogenic F1 sterile lines. *Chin. J. Rice Sci.* 3, 97–101.

Zhang, G., and Lu, Y. (1993). Genetic studies of the hybrid sterility in cultivated rice (*Oryza sativa*). II. A genic model for F1 pollen sterility. *Acta Genet. Sin.* 20, 222–228.

Zhang, G., and Lu, Y. (1999). Breeding of the indica-compatible japonica lines and their use in the breeding of super-high-yield hybrid rice. *Hybrid Rice* 14, 3–5.

Zhang, G., Lu, Y., Liu, G., Yang, J., and Zhang, H. (1993). Genetic studies on the hybrid sterility in cultivated rice (*Oryza sativa*). III. Allele differentiation of F1 pollen sterility in different types of varieties. *Acta Genet. Sin.* 20, 541–551.

Zhang, G., Lu, Y., Zhang, H., Yang, J., and Liu, G. (1994). Genetic studies on the hybrid sterility in cultivated rice (*Oryza sativa*). IV. Genotypes for F1 pollen sterility. *Acta Genet. Sin.* 21, 34–41.

Zhang, G., Zeng, R., Zhang, Z., Ding, X., Li, W., Liu, G., et al. (2004). The construction of a library of single segment substitution lines in rice (*Oryza sativa* L.). *Rice Genet. Newsl.* 21, 85–87.

Zhang, S., Huang, X., and Han, B. (2021). Understanding the genetic basis of rice heterosis: advances and prospects. *Crop J.* 9, 688–692. doi: 10.1016/j.cj.2021.03.011

Zhang, Z., Deng, Y., Tan, J., Hu, S., Yu, J., and Xue, Q. (2007). A genome-wide microsatellite polymorphism database for the indica and japonica rice. *DNA Res.* 14, 37–45. doi: 10.1093/dnares/dsm005

Zhang, Z., and Zhang, G. (2001). Fine mapping of the S-c locus and marker-assisted selection using PCR markers in rice. *Acta Agron. Sin.* 27, 704–709.

Zhao, H., Sun, L., Xiong, T., Wang, Z., Liao, Y., Zou, T., et al. (2019). Genetic characterization of the chromosome single-segment substitution lines of *O. glumaepatula* and *O. barthii* and identification of QTLs for yield-related traits. *Mol. Breed.* 39:51. doi: 10.1007/s11032-019-0960-0

Zhu, W., Li, W., Ding, X., Zhang, Z., Zeng, R., Zhu, H., et al. (2008). Preliminary identification of F1 pollen sterility gene S-e in *Oryza sativa*. *J. South China Agri. Univ.* 29, 1–5.

Zhuang, C., Mei, M., Zhang, G., and Lu, Y. (2002). Chromosome mapping of the S-b locus for F1 pollen sterility in cultivated rice (*Oryza sativa* L.) with RAPD markers. *Acta Genet. Sin.* 29, 700–705. doi: 10.1088/1009-1963/11/5/313

Zhuang, C., Zhang, G., Mei, M., and Lu, Y. (1999). Molecular mapping of the S-a locus for F1 pollen sterility in cultivated rice (*Oryza sativa* L.). *Acta Genet. Sin.* 26, 213–218.

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.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 13 June 2022
doi: 10.3389/fpls.2022.787703

Phenotypic Variation and the Impact of Admixture in the *Oryza rufipogon* Species Complex (*ORSC*)

Georgia C. Eizenga[1]‡, HyunJung Kim[2†‡], Janelle K. H. Jung[2‡], Anthony J. Greenberg[3], Jeremy D. Edwards[1], Maria Elizabeth B. Naredo[4†], Maria Celeste N. Banaticla-Hilario[4], Sandra E. Harrington[2], Yuxin Shi[2], Jennifer A. Kimball[2†], Lisa A. Harper[2], Kenneth L. McNally[4] and Susan R. McCouch[2]**

[1] Dale Bumpers National Rice Research Center, USDA-ARS, Stuttgart, AR, United States, [2] Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, United States, [3] Bayesic Research LLC., Ithaca, NY, United States, [4] International Rice Research Institute, Los Baños, Philippines

Crop wild relatives represent valuable reservoirs of variation for breeding, but their populations are threatened in natural habitats, are sparsely represented in genebanks, and most are poorly characterized. The focus of this study is the *Oryza rufipogon* species complex (*ORSC*), wild progenitor of Asian rice (*Oryza sativa* L.). The *ORSC* comprises perennial, annual and intermediate forms which were historically designated as *O. rufipogon, O. nivara*, and *O. sativa* f. *spontanea* (or *Oryza* spp., an annual form of mixed *O. rufipogon/O. nivara* and *O. sativa* ancestry), respectively, based on non-standardized morphological, geographical, and/or ecologically-based species definitions and boundaries. Here, a collection of 240 diverse *ORSC* accessions, characterized by genotyping-by-sequencing (113,739 SNPs), was phenotyped for 44 traits associated with plant, panicle, and seed morphology in the screenhouse at the International Rice Research Institute, Philippines. These traits included heritable phenotypes often recorded as characterization data by genebanks. Over 100 of these *ORSC* accessions were also phenotyped in the greenhouse for 18 traits in Stuttgart, Arkansas, and 16 traits in Ithaca, New York, United States. We implemented a Bayesian Gaussian mixture model to infer accession groups from a subset of these phenotypic data and ascertained three phenotype-based group assignments. We used concordance between the genotypic subpopulations and these phenotype-based groups to identify a suite of phenotypic traits that could reliably differentiate the *ORSC* populations, whether measured in tropical or temperate regions. The traits provide insight into plant morphology, life history (perenniality versus annuality) and mating habit (self- versus cross-pollinated), and are largely consistent with genebank species designations. One phenotypic group contains predominantly *O. rufipogon* accessions characterized as perennial and largely out-crossing and one contains predominantly *O. nivara* accessions characterized as annual and largely inbreeding. From these groups, 42 "core" *O. rufipogon* and 25 "core" *O. nivara* accessions were identified for domestication studies. The third group, comprising 20% of our collection, has the most accessions identified as *Oryza* spp. (51.2%) and levels of *O. sativa* admixture accounting for more than 50% of the genome. This third group is potentially useful as a

OPEN ACCESS

Edited by:
Anna Maria Mastrangelo, Research Centre for Cereal and Industrial Crops, Council for Agricultural and Economics Research (CREA), Italy

Reviewed by:
Ryo Ishikawa, Kobe University, Japan
Robert Henry, The University of Queensland, Australia

***Correspondence:**
Georgia C. Eizenga
georgia.eizenga@usda.gov
orcid.org/0000-0003-0330-9225
Susan R. McCouch
srm4@cornell.edu
orcid.org/0000-0001-9246-3106

†Present addresses:
HyunJung Kim, LG Chemical Ltd., Seoul, South Korea
Maria Elizabeth B. Naredo, Philippine Coconut Authority – Albay Research Center, Guinobatan, Philippines
Jennifer A. Kimball, University of Minnesota-Twin Cities, St. Paul, MN, United States

‡*These authors share first authorship*

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

Received: *01 October 2021*
Accepted: *13 April 2022*
Published: *13 June 2022*

Citation:
Eizenga GC, Kim HJ, Jung JKH, Greenberg AJ, Edwards JD, Naredo MEB, Banaticla-Hilario MCN, Harrington SE, Shi Y, Kimball JA, Harper LA, McNally KL and McCouch SR (2022) Phenotypic Variation and the Impact of Admixture in the Oryza rufipogon Species Complex (ORSC). Front. Plant Sci. 13:787703. doi: 10.3389/fpls.2022.787703

"pre-breeding" pool for breeders attempting to incorporate novel variation into elite breeding lines.

Keywords: rice, Bayesian Gaussian mixture models, *Oryza rufipogon*, *Oryza sativa*, *Oryza nivara*, genebank accessions, crop wild relatives, *Oryza rufipogon* species complex

INTRODUCTION

Wild relatives of domesticated crop species are some of the greatest sources of untapped genetic variation available to plant breeders as they confront the challenges of a changing climate. Yet populations of crop wild relatives are underrepresented in genebanks, threatened in their natural environments, and poorly characterized in both *ex situ* and *in situ* collections. The *Oryza rufipogon* species complex (*ORSC*), previously referred to as *O. rufipogon* Griff. or *O. perennis* Moench, is the wild progenitor of Asian rice (*Oryza sativa* L.) and is comprised of perennial, annual, and intermediate ecotypes (Sano et al., 1980; Oka, 1988). As summarized in **Supplementary Table 1**, the perennial ecotype exhibits vigorous vegetative growth, is largely out-crossing, and is found in areas which are continuously wet while the annual ecotype shows less vigorous vegetative growth, is primarily in-breeding, and is found in areas that are seasonally wet (Oka and Morishima, 1967; Morishima et al., 1984; Vaughan et al., 2008). In accordance with these differences in life history, the perennial form is capable of reproducing clonally via stolons, typically has low seed productivity, is late flowering and photoperiod sensitive, while the annual form lacks stolons, has high seed productivity, is frequently early-flowering and photoperiod insensitive (Morishima et al., 1984; Barbier, 1989a,b; Barbier et al., 1991; Morishima, 2001; Vaughan et al., 2003, 2008; Cai et al., 2004). An extensive survey of both annual and perennial ecotypes of the *ORSC* led Sharma and Shastry (1965) to treat the annual form as a separate species, designated *O. nivara*, as originally suggested by Chatterjee (1948), to distinguish it from the perennial *O. rufipogon*. A weedy, intermediate form is classified as *O. sativa* f. *spontanea* Roschev (Roschevicz, 1931; Oka, 1988) and sometimes erroneously designated *O. spontanea* (Oka, 1974; Sano et al., 1980; Vaughan, 1989; Hill, 2010). Weedy rice is considered an annual of mixed *O. rufipogon/O. nivara* and *O. sativa* ancestry (Morishima et al., 1961; Chang, 1976; Vaughan et al., 2001; Sharma, 2003).

To this day, wide differences in the nomenclature and trait-based definitions assigned to the wild ancestor of *O. sativa* persist in rice genebank databases and research labs around the world. According to the online database of the International Rice Germplasm Collection (IRGC)[1] the list of available *ORSC* accessions includes 846 *O. rufipogon*, 1,455 *O. nivara*, and 1,134 accessions listed as either '*Oryza* spp.' or '*Oryza* hybr.'. Many of the *Oryza* spp. or *Oryza* hybr. were previously designated as *O. spontanea (O. sativa* f. *spontanea)* or referred to as hybrids between *O. sativa* and *O. rufipogon* and/or *O. nivara*. The online wild strain database of the Japanese National Institute of Genetics, *Oryzabase* (Kurata and Yamazaki, 2006; Yamazaki et al., 2010), acknowledges confusion in the wild rice taxonomy and nomenclature, including a distinction between *O. nivara* and *O. rufipogon* based on ecology and life habit, but clearly states its use of the *sensu lato* classification of *O. rufipogon* as a single species with a continuous range of annual, perennial, and intermediate types as held by core collectors Morishima and colleagues (Morishima et al., 1961). There are 682 accessions in *Oryzabase* designated as '*O. rufipogon*' with additional, but incomplete information on former species designations (i.e., '*O. perennis', 'O. perennis (O. nivara),'* or *'O. sativa f. spontanea*,'), and life habit designations, such as annual or perennial[2].

While the species designations originally assigned to accessions in the IRGC have persisted largely unchanged through the decades, there has been a paradigm shift in biology from morphology-based to genetic identity-based species definitions. In the case of the *ORSC*, studies using molecular markers (isozymes, RFLPs, SSRs, SINEs and other indels, and SNPs) present conflicting and often contradictory results that collectively fail to support the existence of two, well-differentiated species. Global studies of genomic diversity in the *ORSC* document three to eight genetically distinct groups (Londo et al., 2006; Zheng and Ge, 2010; Huang et al., 2012; Banaticla-Hilario et al., 2013a,b; Liu et al., 2015; Kim et al., 2016) more closely associated with geography than with life-habit. Pronounced genetic differentiation has been reported between *O. rufipogon* and *O. nivara* only in studies involving local collections of germplasm from South Asia (Samal et al., 2018) and Southeast Asia (Kuroda et al., 2006; Banaticla-Hilario et al., 2013a), where both species occur sympatrically but remain differentiated due to differences in flowering time and ecological adaptation. Genetic differentiation has not been documented in China or Oceania where native populations of *O. nivara* are largely absent (Vaughan, 1994; Liu et al., 2015).

Several studies based on genome-wide SNPs have examined relationships among *ORSC* accessions and among *ORSC* and *O. sativa* subpopulations. A study by Huang et al. (2012) examined a panel of 446 *ORSC* accessions and 1,083 *O. sativa* cultivars that had been genotyped with ∼5M SNPs and identified three major groups of *ORSC* accessions using a neighbor-joining tree. These groups were significantly correlated with geographic distribution and were referred to as *OrI, OrII* and *OrIII*. Subsequently, *OrI* was divided into two sub-clades, where *OrIa* included *ORSC* accessions and rice belonging to the *O. sativa-indica* subpopulation and *OrIb* included *ORSC* and *O. sativa-aus* rices; *OrIII* was also divided into two sub-clades, where *OrIIIa* included *ORSC* accessions from Southern China

[1] https://gringlobal.irri.org/gringlobal/search

[2] https://shigen.nig.ac.jp/rice/oryzabase/strain/wildCore/list

as well as *O. sativa-japonica* rices, and *OrIIIb* included *ORSC* accessions that clustered independently of *O. sativa*.

The *ORSC* dataset generated by Huang et al. (2012) was subsequently re-analyzed by three independent research groups: Civán et al. (2015), Wang et al. (2017), and Choi and Purugganan (2018). In each case the authors reached very different conclusions about the subpopulation structure of the *ORSC* and relationships with the cultivated gene pools of *O. sativa.* These subsequent studies revealed evidence of substantial gene flow from domesticated groups into wild populations and suggested that introgression rather than phylogenetic relationship was a primary driver of the *ORSC* groups reported by Huang et al. (2012). Use of larger K values (population number) and local ancestry estimation techniques helped to resolve the heavily *O. sativa*-admixed wild groups as distinct from independent wild populations. The analysis by Wang et al. (2017) combined the sequence data of 435 *ORSC* accessions (Huang et al., 2012) with sequencing data from 203 *O. sativa* accessions included in the Rice Minicore collection (Wang et al., 2016). Subsequent STRUCTURE analysis ($K = 9$) ascertained six *ORSC* subpopulations: four unique *ORSC* subpopulations that were correlated with geographic distribution and two that clustered with *O. sativa* subpopulations. Of these, *Or-A* accessions had the broadest range with the highest proportion from Oceania, *Or-B* accessions were almost exclusively found in China, *Or-C* had the highest proportion from West India and Sri Lanka, and *Or-D* was mainly from the SE Asia, Bangladesh, and East India, *Or-E* accessions clustered with *O. sativa-aus* accessions, and *Or-F* clustered with *O. sativa-indica* (Wang et al., 2017). Of note, about 42% of these *ORSC* accessions were deemed to be substantially admixed, thus could not be assigned to a single ancestry group.

To further disentangle the population dynamics of the *ORSC*, a different collection of 286 accessions from the IRGC was genotyped using genotyping-by-sequencing (113,739 SNPs) and evaluated at 25 polymorphic sites in the chloroplast genome (Kim et al., 2016). Six wild subpopulation groups were identified; three were closely related to *O. sativa-indica, -aus,* and *-japonica,* respectively, while the other three subpopulations were genetically divergent, had unique chloroplast haplotypes, and were located at the geographical extremes of the species range.

Another study examined local differentiation between *O. rufipogon* and *O. nivara* using a collection of 52 pairs of sympatric *O. rufipogon* and *O. nivara* accessions that were phenotyped for 32 traits (Banaticla-Hilario et al., 2013a) and genotyped with 29 SSR markers (Banaticla-Hilario et al., 2013b). The genotypic neighbor-joining tree showed no clear-cut genetic separation of *O. nivara* and *O. rufipogon* accessions though, species separation was apparent at the local scale. Nonetheless, Bayesian clustering was able to differentiate the two species in sympatric population pairs across the entire range of their distribution (Banaticla-Hilario et al., 2013a). Furthermore, this study identified a cluster of *O. nivara* accessions from Nepal that diverged significantly from the other population groups (similar to the findings reported by Kim et al., 2016), as well as two or three other genetically and geographically distinct subpopulations of *O. nivara,* suggesting that reproductive barriers in this inbreeding group may tend to intensify under sympatric conditions, and that *O. nivara* may have originated more than once from its perennial ancestor. Local differences in flowering time or in floral and panicle structure associated with the mating system would impact adaptation and have been known to drive reproductive isolation in closely related species (Grillo et al., 2009; Liu et al., 2015). Statistical analysis of the phenotypic data showed 27 of the 32 traits were significantly different between *O. rufipogon* and *O. nivara* with spikelet width, anther length, culm length and photoperiod explaining 25.4 percent of the variation between the two species. These studies also noted that Southeast Asian accessions appeared to be more recently diverged and/or had more interspecific gene flow compared to those from South Asia.

To augment our understanding of diversity and population structure in the *ORSC* and to facilitate the selection of materials for use in plant breeding, we phenotyped the collection genotyped by Kim et al. (2016). Phenotyping was performed at three locations (Los Baños, Philippines, Arkansas, United States and New York, United States) for heritable traits, with emphasis on those commonly used to characterize rice accessions in genebanks (Bioversity International et al., 2007). Several of the traits were associated with differences in life habit and/or mating system and served to substantiate the distinction between *O. nivara* and *O. rufipogon* as distinct ecotypes within the *ORSC*. A significant proportion of *ORSC* accessions carried unexpected combinations of phenotypes and were associated with admixture from *O. sativa*.

To enable a more robust interpretation of the phenotypic variation observed, we also merged genotypic datasets for the 286 *ORSC* accessions analyzed by Kim et al. (2016) and the 446 *ORSC* accessions analyzed by Huang et al. (2012), Civán et al. (2015), Wang et al. (2017), and Choi and Purugganan (2018) and analyzed the extent of unique or shared variation in each of the studies. Based on these analyses, we identify a set of 'core accessions' representing *O. rufipogon* and *O. nivara* along with a large group of *ORSC* accessions that represent unique sources of naturally occurring, highly admixed pre-breeding material for use in plant breeding. This work adds value to the *ORSC* genetic resources currently conserved in the IRGC and enhances opportunities for expanded utilization in research and plant improvement.

MATERIALS AND METHODS

Plant Materials and Phenotypic Data Collection

For this study, 240 *ORSC* accessions from the International Rice Germplasm Collection (IRGC) previously genotyped by Kim et al. (2016), were selected for phenotypic evaluation (**Supplementary Table 2**). This collection included 102 *ORSC* accessions from Southeast Asia, 79 accessions from South Asia, 55 accessions from East Asia and four accessions from Australasia. Of these, 222 accessions were evaluated in a screenhouse at the IRRI, Los Baños, Philippines and 130 were evaluated under

greenhouse conditions in the United States, with 112 accessions phenotyped in more than one location. At IRRI two plants were grown of each accession with one plant per pot and each pot was considered a single replication. Seed of each accession was planted between June 20 and July 4, 2008, and the plants were phenotyped for the 44 morphological traits related to the vegetative, reproductive and harvest growth stages (**Supplementary Table 3**). Twenty-two of these 44 traits were measured five times on an individual plant to improve the accuracy of the measurement (**Supplementary Table 4**).

In the United States, 104 accessions were phenotyped at both Cornell University (CU) in Ithaca, New York and the Dale Bumpers National Rice Research Center (DB) near Stuttgart, Arkansas. Three plants were grown of each accession with one plant per pot and each pot was considered a single replication. At CU, 104 accessions were planted between August 31 and September 8, 2008, and each plant was phenotyped for 16 morphological traits. At DB, 129 *ORSC* accessions were grown in the greenhouse with three plants per accession, one plant per pot, each pot was considered one replication, and plants were phenotyped for 18 morphological traits. In Arkansas, each accession was phenotyped in two different years, thus a total of six plants were characterized for most accessions. The accessions were planted August 24 to September 6, 2007 (86 accessions); August 1 to September 12, 2008 (124 accessions) and August 12 to 13, 2009 (41 accessions). **Supplementary Table 2** lists the location where each accession was grown.

Supplementary Table 3 summarizes the 57 phenotypic traits characterized across the three locations and indicates the location(s) where the trait data were collected. Additional details about the collection of phenotypic data at IRRI are included in Bioversity International et al. (2007). Also listed are the corresponding Planteome acronyms and the trait ontology or crop ontology terms (Cooper et al., 2018). At IRRI, for the 22 quantitative traits with five measurements per plant, the within-plant variances were very small, thus a mean was calculated for subsequent analyses. The replicated trait data collected for each accession by location was used to calculate a best linear unbiased estimator (BLUE) for each trait and accession by location as described in the next section. The BLUEs are listed by location in **Supplementary Table 4**.

Data Quality

Using the replicated trait measurements taken at each location and a relationship matrix constructed from the GBS genotypes (Kim et al., 2016), we fit multi-trait Bayesian hierarchical models (Greenberg et al., 2011) to the data from each site separately. The model is:

$$\begin{aligned} y_{i\cdot} &\sim t_{\nu_e,d}(\mu^{acc}_{j[i]\cdot};\ \Sigma^e) \\ \mu^{acc}_{j\cdot} &\sim \mathrm{N}_d(\mu + u_{j\cdot}\Gamma;\ \Sigma^s) \\ \gamma_{j\cdot} &\sim \mathrm{N}_d(0_d;\ \Sigma^a) \end{aligned}$$

Location parameters of the form $\boldsymbol{y}_{\mathrm{i}\cdot}$ are row-vectors of the corresponding parameter matrices that have data points as rows and traits as columns. The overall intercept μ was modeled with a high-variance ($\boldsymbol{I}\sigma_0^2 = 10^6$) Gaussian prior. The errors were modeled using a multivariate Student-*t* distribution with three degrees of freedom to dampen the effects of outliers (Greenberg et al., 2010; Greenberg et al., 2011). We computed the parameters using a C++ program based on an openly-available library[3] that implements methods described by Greenberg et al. (2011).

The marker effects enter through the eigenvectors $\boldsymbol{U}$ of the relationship matrix, weighted by square roots of their eigenvalues. We estimated the relationship matrix from all non-singleton SNPs in each data set using the van Raden method (VanRaden, 2008). We used all the eigenvectors that correspond to non-zero eigenvalues. If this regression is performed using a Gaussian prior on the coefficients $\gamma_{j\cdot}$, it is the Bayesian analog of the (matrix-variate) mixed effect model (Kang et al., 2010; Hoffman, 2013). The accession means $\mu^{acc}_{j\cdot}$ are analogous to BLUEs. Modes of their posterior distributions were used in the subsequent analyses.

The covariance matrices Σ^x were modeled using Wishart distributions with weakly-informative Wishart priors with two degrees of freedom for Σ^e and Σ^s and four degrees of freedom for Σ^a (Gelman et al., 2004; Greenberg et al., 2011). This implies a uniform prior on narrow-sense (marker) heritability. Estimated model parameters can then be used to calculate marker heritability:

$$h^2_{M,p} = \frac{(U\Gamma)^T_{\cdot p}(U\Gamma)_{\cdot p}/(N_{acc}-1)}{(U\Gamma)^T_{\cdot p}(U\Gamma)_{\cdot p}/(N_{acc}-1) + \Sigma^s_{p,p} + \Sigma^e_{p,p}},$$

where N_{acc} is the number of accessions, and the rest of the notation is as described above. We used the variance of genome-estimated breeding values (GEBV).

$$(U\Gamma)^T_{\cdot p}(U\Gamma)_{\cdot p}/(N_{acc}-1)$$

rather than values from the marker covariance matrix $\Sigma^a_{p,p}$ because the latter reflects the additive covariance only when the prior on the principal component regression (see above) is Gaussian. The $\Sigma^s_{p,p}$ values estimate the portion of total genetic variance not explained by genotyped markers. They include the non-additive effects and additive effects not tagged by SNPs. The sum of the GEBV and background variance, divided by total phenotypic variance, is broad-sense heritability (Lynch and Walsh, 1998).

Mixture Model

Using estimates of phenotypic values for each accession, we fit a Gaussian mixture model (Robert, 1996) using a variational Bayes approach (McGrory and Titterington, 2007). The model is

$$\mu^{acc}_{j\cdot} \sim \frac{1}{C}\sum_{m=1}^{N_M} N_d(\mu_{m\cdot};\ \Sigma_{A,m})^{z_{jm}}$$

[3]https://github.com/tonymugen/MuGen

$$z_{jm} \sim \text{Bernoulli}(\pi_m)$$
$$\pi_m \sim \text{Dirichlet}(\alpha)$$
$$\mu_{m\cdot} \sim \text{N}_{N_M,d}(0; \lambda_0 \Sigma_{A,m})$$
$$\Sigma_{A,m} \sim \text{W}(\Sigma_0; \upsilon_0)$$

where the notation is the same as for the hierarchical model above and $\mu_{m\cdot}$ is a row vector of group means. Each group has a separate mean and covariance matrix. The prior population size α is set to a small value, shrinking groups with few members to zero. The number of groups must be set *a priori*.

We perform inference by running the model multiple times from different starting points and picking the best fit, as assessed using the deviance information criterion (DIC, McGrory and Titterington, 2007). Small DIC values indicate better fit. Our implementation of the model is available as the R package MuGaMix[4]. We use two methods to assess how many groups are statistically supported: the DIC and the number of groups with at least two accessions. We include the pipeline (implemented in R) that runs the mixture model, processes the results, and generates all the plots and summary tables presented in this manuscript in the **Supplementary Material File 1**.

Merged Population Structure Analysis

For population structure analysis, the SNP datasets from Kim et al. (2016) and Huang et al. (2012) were merged into a single dataset. First, the Kim et al. (2016) SNPs from *ORSC* and *O. sativa* accessions were imputed using Beagle 5.0 (Browning et al., 2018). Next the intersection between the SNPs in the imputed Kim et al. (2016) and the (already imputed) SNPs in the Huang et al. (2012) was determined and a merged file was generated using BCFtools version 1.14 (Danecek et al., 2021). Population structure analysis was done using fastSTRUCTURE version 1.0 (Raj et al., 2014) with *K* values ranging from 2 to 9.

RFMix Analysis

The local ancestry of the *ORSC* individuals was determined using RFMix (Maples et al., 2013). The training panel included ten individuals from each subpopulation (W1-W6) with the highest global ancestry as reported by Kim et al. (2016) and included individuals from each of the five *O. sativa* subpopulations, *indica* (IND) (10 individuals), *aus*, (AUS) (9 individuals), *temperate japonica* (TEJ) (9 individuals), *tropical japonica* (TRJ) (10 individuals), and *aromatic* (ARO) (6 individuals). Of the 60 *ORSC* individuals in the training panel, 49 were phenotyped in the current study. Prior to RFMix analysis, the genotypic data were imputed and phased using Beagle 5.0 (Browning et al., 2018).

RESULTS

Evaluation of Phenotypic Data

Broad-sense and marker heritability distributions for each trait and location are presented in **Supplementary Table 5**. Heritabilities range from 0.98 to 0.33 and are generally moderate to high. Markers account for at least 33% of total genetic variance across all traits and locations, but it is the generally high H^2 that gives confidence in within-site reproducibility of our phenotypic measurements.

Fourteen traits were measured at more than one site. We tested among-site reproducibility of these measurements by estimating correlations among phenotypes evaluated on the same accessions and generally see positive correlations (**Supplementary Figure 2** and **Supplementary Table 6**), again suggesting that phenotypic measurements are reproducible across sites despite the diverse climatic conditions. One notable exception is panicle number (PNNB). Values of this trait are positively correlated between Cornell and Dale Bumpers (temperate and sub-tropical sites, respectively), but both are negatively correlated with IRRI measurements (tropical site), suggesting a genotype by environment interaction.

Inferring Phenotypic Groups

Using estimates of 32 phenotypic values (binary traits were excluded) for each accession in the IRRI data set (n = 222), we fit a Gaussian mixture model (Robert, 1996) using a variational Bayes approach (McGrory and Titterington, 2007). To infer the underlying number of groups that is statistically supported and biologically meaningful, we ran analyses assuming two to 15 groups *a priori* and assessed model fit (**Supplementary Figure 3**) using the deviance information criterion and the number of non-empty groups. Both metrics lend statistical support for the division of our accessions into as many as 10 phenotypic groups. However, to make biologically relevant inferences we need a parsimonious classification that is not only statistically significant, but also robust to model uncertainty, reproducible, and interpretable given other biological information. The variational Bayes approach we used is fast but produces only point estimates of parameters. To test robustness of these estimates to initial conditions, we started the model fitting process from five independent sets of initial values. Model fit metrics were consistent across initial value sets (**Supplementary Figure 3**).

Moving beyond model fit estimates, we tested stability of group membership across model specifications and starting value sets. To anchor these inferences, we started by noting taxonomic classification of our accessions, as found in the IRRI-GRIN-Global database. Based on traditional classification methods, accessions in this database are listed as either *O. rufipogon*, *O. nivara*, or, if the taxonomic assignment is uncertain, as *Oryza* spp. or *Oryza* hybr. We combined the latter two categories into one, calling it *Oryza* spp. We tracked each accession, starting from the species designations, across phenotypic groups inferred given different *a priori* group numbers (N_G) using a Sankey plot (**Supplementary Figure 4**; Kennedy and Sankey, 1898; implemented in the ggsankey R package)[5]. A line, which is like a ribbon in such plots, represents a single accession, while boxes arranged along the *y*-axis are groups. Groups derived from models assuming a different number of groups are shown along

[4]https://github.com/tonymugen/MuGaMix

[5]https://github.com/davidsjoberg/ggsankey

the *x*-axis. We can thus follow a given accession as its line connects to each group across models with different N_G. If we set N_G = 3, the model infers one large group comprising most accessions. The second largest group is composed predominantly of accessions classified as *Oryza* spp. Increasing the *a priori* N_G to four, we can assign most *O. rufipogon* and *O. nivara* accessions to separate groups. If we further increase N_G, distinct groups corresponding to the traditional species designations are maintained, with an additional group that comprises most *Oryza* spp. accessions. However, additional small groups inferred with higher *a priori* N_G do not appear to maintain coherent membership across model runs (**Supplementary Figure 4**). Thus, while there is statistical support for finer classification, these groupings are not robust to model specification. We therefore focused our analyses on inferences assuming four *a priori* groups.

Given that models with N_G = 4 appear to be the most parsimonious, while still capturing biologically meaningful stratification of the *ORSC*, we next tested these models for robustness. We started model fitting from 20 sets of initial conditions and averaged posterior probabilities that a given accession belongs to a particular group across runs. We see a set of accessions that reproducibly belong to the P1 or P4 groups (**Figure 1A**), corresponding to *O. rufipogon* and *O. nivara*, respectively. The P3 group has the least reproducible membership. This is because for some initial conditions, a small number of accessions are assigned to it, while for others these lines are placed in P2 (see **Supplementary Material 1** for a detailed analysis of these shifts). This may be due to the approximate nature of the variational Bayes inference. Therefore, we added the posterior probabilities of P2 and P3 assignment in subsequent analyses and called the resulting group P2/P3.

The resulting P1 phenotype group contains 106 accessions, of which 82 (78%) are classified as perennial *O. rufipogon* species. In contrast, P4 has 73 accessions with 43 (59%) classified as *O. nivara*. About half of the P2/P3 group are designated as *Oryza* spp. in the IRGC, with the remaining accessions about equally divided between *O. nivara* and *O. rufipogon* (**Supplementary Figure 5** and **Supplementary Table 2**).

To further test the reproducibility of phenotypic group inference, we used data collected at DB and CU. These data sets include different, but overlapping, phenotype sets that were measured on fewer than half the accessions represented in the IRRI data (**Supplementary Tables 2, 3**). Despite this, we can still identify a similar set of three groups (**Figure 1B**). This time, setting N_G = 3 *a priori* is sufficient to separate *O. rufipogon* and *O. nivara* from each other and from a group that largely corresponds to the *Oryza* spp. designation from IRRI-GRIN-Global. Taking these observations together with the above statistical robustness analyses, we see that a core set of accessions can be reproducibly assigned to a phenotypic group, while inference for others is less certain.

Subsets of Traits Defining Phenotypic Groups

Having established a robust phenotype-driven grouping of *ORSC* accessions, we next wanted to define a minimal combination of phenotypic traits that is sufficient to reproduce these divisions. Our mixture model allows both phenotype means and covariance to vary among groups. We started by ranking traits according to how well their values correlate with our groups (**Figure 2A**). We used a multivariate regression with group identity probabilities as the response variable and trait values as predictors. Since the last group membership is fully determined by the other three, only P1 and the combined P2 and P3 membership probabilities were used. We estimated regression coefficients and their covariance and used them to calculate a Hotelling T^2 statistic (Hotelling, 1931) of each trait-group association. This is a multivariate version of the square of Student's *t* statistic widely used to determine statistical significance of individual regression coefficients. However, in this case we were not interested in the significance of associations. Instead, we ranked traits by their T^2 values (**Figure 2A**). Two traits, culm length (CULT) and number of empty spikelets per panicle (UNFILLED), stand out for their association with phenotypic groups. Their Hotelling T^2 statistics (0.71 for CULT and 0.70 for UNFILLED) are over twice the value of the next seven highest trait-group associations for which T^2 ranges from 0.17 to 0.28 (**Table 1** and **Supplementary Table 7**). Accessions belonging to the P1 group are on average taller, with P4 the shortest and P2/P3 intermediate. P2/P3 accessions typically have the most empty spikelets per panicle, while P4 accessions have the least (**Supplementary Figure 6**).

To find pairs of traits that change correlations across groups, we first estimated within-group associations and then used a permutation test (randomly assigning accessions to groups) to assess statistical evidence for across-group variation (**Figure 2B**). The most significant correlation changes occur between the number of empty spikelets (UNFILLED) and either panicle length (PNLG) or culm diameter (CUDI). These correlations are positive (0.54 and 0.48, respectively) only in P4 (**Supplementary Table 7**).

To test if subsets of traits can reproduce our phenotypic groups, we took 11 traits that belong to pairs that change correlations with probability at least 0.001 (**Table 1**). This set also includes the top two traits (CULT and UNFILLED) whose values correlate best with group assignments. To expand the trait subset, we added seven phenotypes with $T^2 > 0.17$. Hotelling values showed a marked drop below this cut-off. These two sets have four traits in common [CULT, UNFILLED, CUDI, and anther length (ANTLT)]. We also used the union of the two sets, containing 16 traits, or half the number in the full data set. Although covariance matrices include both correlations and variances, all traits with appreciable among-group variance changes are already accounted for by using correlation and mean differences (**Supplementary Figures 6, 7**).

We used three data subsets (nine traits associated with phenotypic groups by value only, 11 that switch correlations, and 16 that were associated either by value or correlation) to re-fit our mixture model, setting the *a priori* number of groups at three. We ran our analyses 20 times and averaged posterior probabilities of accessions belonging to each group. Using all three data sets, we consistently see three groups that separate *O. rufipogon* from *O. nivara*. The nine- and 16-trait data sets assign most P2/P3 accessions to P1 (**Figure 2C**). The 11 traits most likely to change

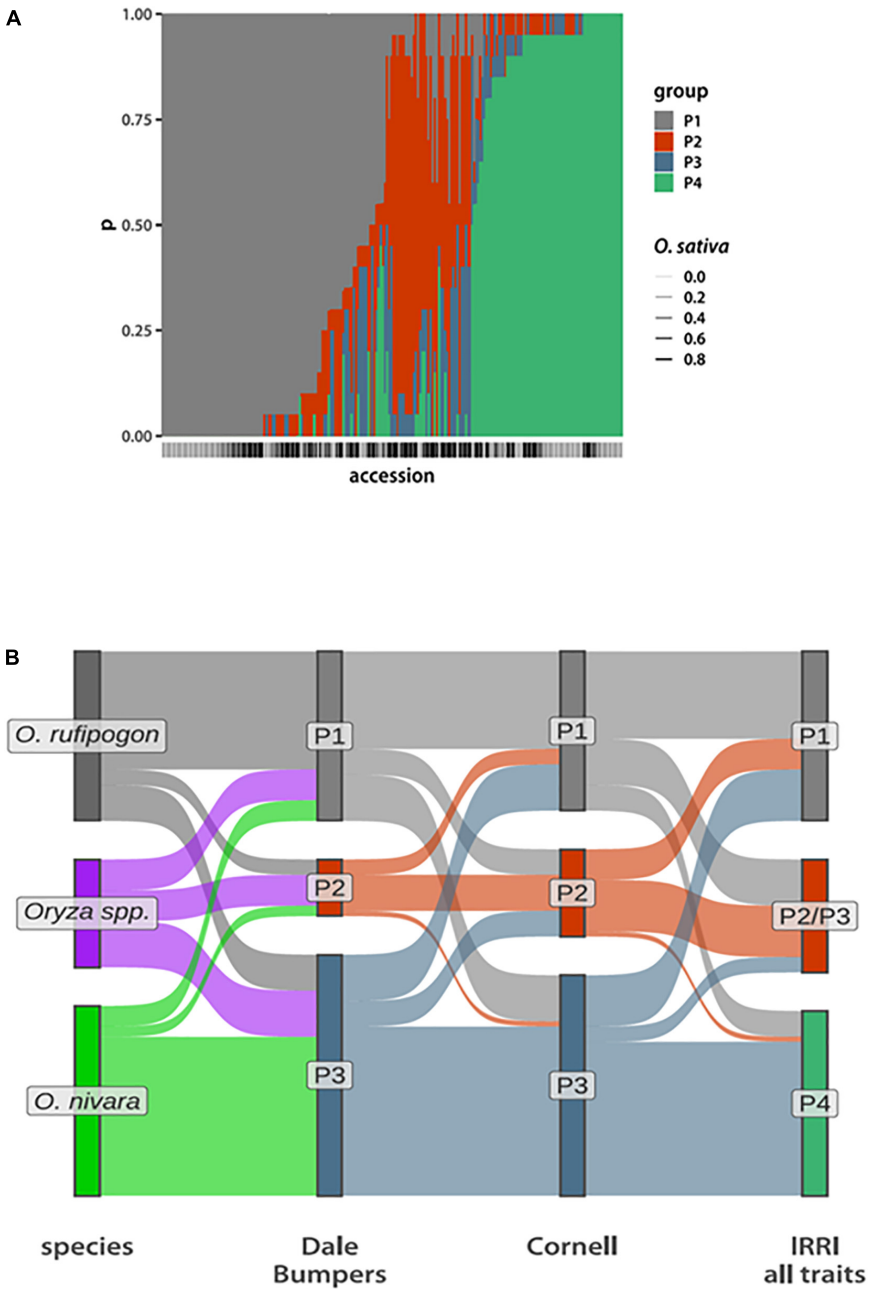

FIGURE 1 | Phenotypic group composition. **(A)** Bar plot depicting posterior probabilities that a given accession (along the *x*-axis) belongs to each of the four phenotypic groups. The bars are stacked, with the total probability summing to 1. The bars below the x-axis indicate the *O. sativa* genome fraction, with the darker bars having a higher *O. sativa* fraction. (The actual values are in **Supplementary Table 2**) **(B)** Sankey plots comparing compositions of phenotypic groups inferred from data collected at Dale Bumpers and Cornell assuming N_G = 3 compared to species designations and the four groups inferred from data collected at IRRI. Each accession is a single "ribbon" running from left to right and colored according to the grouping to the left.

correlations across groups appear to be sufficient to recapitulate the grouping inferred from the whole data set. The two traits, CULT and UNFILLED, most closely aligned with groups by value and are included in this collection.

Our approach identified a minimally sufficient set of traits that captures most of the phenotypic group structure in the IRRI data. This does not mean that other sets of traits could not achieve similar results but an exhaustive search for all possible combinations that could achieve the same result was not undertaken considering time and the computational expense. Instead, we pursued two approaches using the Cornell and Dale Bumpers data sets to find alternative sets of traits that could be used to re-infer the three major phenotypic groups identified using the full data set. First, we used the process outlined above

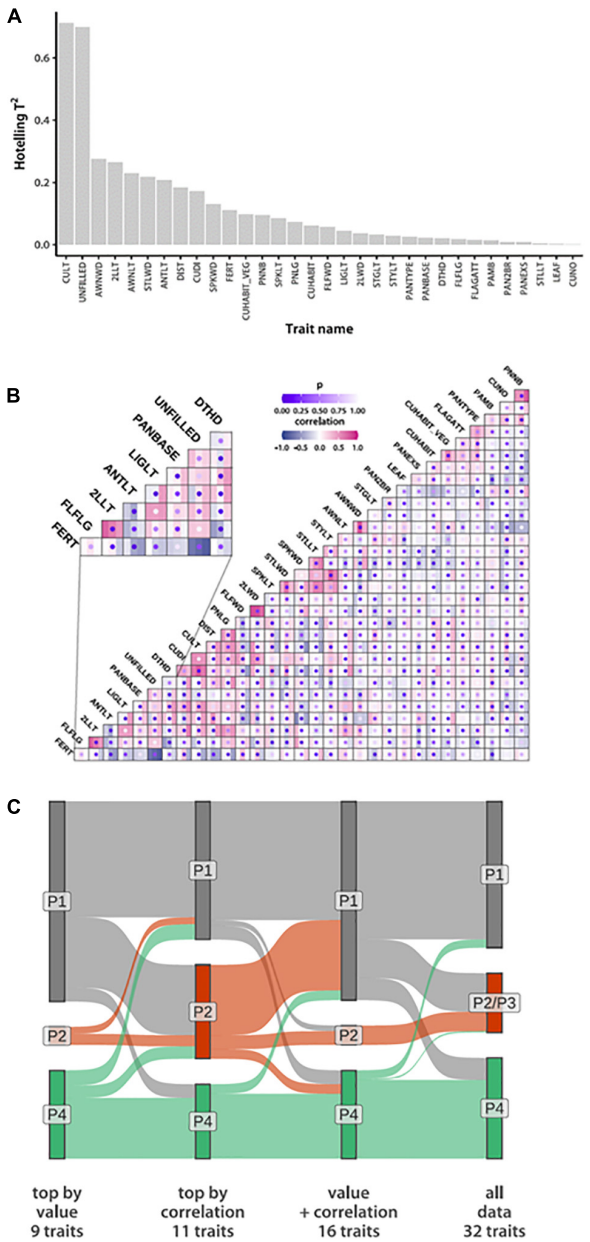

FIGURE 2 | Phenotypic identifiers of group divergence based on the IRRI data. **(A)** Traits sorted by the strength of the association of their values to N_G = 4 group assignments based on Hotelling T^2 values **(B)** Among-trait correlation estimates within each phenotypic group. Squares mark trait pairs, with the three bars within each square colored according to correlation magnitude in the three phenotypic groups. Dots represent probabilities of observing the correlation differences by chance (lower probabilities reflect high confidence of correlation difference). The inset shows a segment of the plot in more detail. (Trait acronyms and actual values are in **Supplementary Table 7**) **(C)** Sankey plot comparing accession membership across groups inferred from trait subsets listed in **Table 1**.

to identify traits associated with the phenotypic groups that were inferred from IRRI data by value or that change correlations with other traits. **Supplementary Figures 8, 9** visualize the Dale Bumpers data analyses and **Supplementary Figures 10, 11** visualize the Cornell data analyses with the actual values in **Supplementary Table 7**. This approach did not identify additional trait subsets that could reliably identify phenotypic groups (**Supplementary Figures 12A,C**). Second, we identified traits evaluated at Dale Bumpers and Cornell that are the same or similar to the 11 phenotypes (**Supplementary Table 3**) that defined the grouping patterns in the IRRI data. Re-running our mixture model with these analogous traits reproduced the original groups, albeit at lower resolution (**Supplementary Figures 12B,D**). Using fewer than 11 traits thus appears to significantly degrade group assignment precision and reproducibility.

Relationships Between Phenotypic Groups and Genetic Subpopulations

While our phenotypic groups largely correspond to IRRI-GRIN-Global species designations, we wanted to know how the groups related to previously characterized genotypic subpopulations. The accessions used in this study were genotyped by Kim et al. (2016) who reported six *ORSC* subpopulations (W1-W6). Based on RFMix (Maples et al., 2013) analysis, we identified introgressions among *ORSC* subpopulations and between the *ORSC* and five *O. sativa* subpopulations (IND, AUS, TRJ, TEJ, ARO). We partitioned accessions with more than half of their genome introgressed from *O. sativa* into a separate "admixture with *O. sativa*" group (ADM/OSAT). By means of a Sankey plot, we related the species designation, phenotypic group and genetic subpopulation for the *ORSC* accessions with complete information (**Figure 3A** and **Supplementary Table 2**). Across the *ORSC* subpopulations, the majority of the W1, all W3, and about half of the W6 accessions (as per Kim et al., 2016) clustered in the P1 group, taxonomically designated as *O. rufipogon*. Conversely, all of the W5 accessions and most of the W2 clustered in the P4 group, taxonomically designated as *O. nivara*. Finally, accessions with significant *O. sativa* introgression fell almost exclusively into the P2/P3 group and are classified as *Oryza* spp. The P2/P3 group identified using the reduced set of 11 traits also includes most of the W6 subpopulation (**Supplementary Figure 13**). Eight accessions taxonomically classified as *O. rufipogon* in the IRGC were genetically designated as W2 or W4 by Kim et al. (2016) and clustered with the P4 group in this study. These accessions had less than 20% of their genome from *O. sativa* and were re-classified as *O. nivara* in our analyses (including in the results reported above) and are identified in **Supplementary Table 2**.

Surveying the individual introgressions across the genome, there were no obvious patterns identifying particular genomic regions underlying the phenotypic group differentiation (**Figure 3B**). This result was not surprising, given that these groups are distinguished by multiple polygenic traits. We do see that those accessions from the P2/P3 group were more likely to harbor *O. sativa* introgressions (**Figures 3B,C**) and the median portion of the genome coming from *O. sativa* is also larger in P2/P3 than in P1 or P4 accessions (**Supplementary Figures 14B, 15**). However, P1 and to a lesser extent P4 harbor a few accessions with significant contributions

TABLE 1 | Phenotypic traits identified by trait value (9 traits), correlation switching (11 traits) and the union of both sets of traits (16 traits) based on the analysis of the phenotypic trait data collected at IRRI.

Descriptive trait name	Trait acronym (IRRI)	Top traits by value (9)	Top traits by correlation (11)	Union of value & correlation (16)
Anther length	ANTLT	X	X	X
Awn length	AWNLT	X		X
Awn width	AWNWD	X		X
Culm diameter	CUDI	X	X	X
Culm length	CULT	X	X	X
Days to 50% heading	DTHD		X	X
Distance of nearest spikelet to the panicle base	DIST	X		X
Flag leaf lamina width	FLFWD		X	X
Ligule length	LIGLT		X	X
No. empty spikelets per panicle	UNFILLED	X	X	X
Panicle fertility	FERT		X	X
Panicle length	PNLG		X	X
Penultimate (2nd) leaf length	2LLT	X		X
Spikelet length	SPKLT		X	X
Spikelet width	SPKWD		X	X
Sterile lemma width	STLWD	X		X

*These were used to construct the groups shown in the Sankey plot (**Figure 2C**).*

from *O. sativa* or *ORSC* subpopulations typically associated with a different phenotypic group, suggesting that similar phenotypic syndromes can be achieved through divergent genotypes.

Most *O. sativa* introgressions come from the *indica* subpopulation (**Figure 3** and **Supplementary Figure 14B**). Only accessions belonging to the P2/P3 group show appreciable contributions from *O. sativa japonica*. Introgressions from *aus* are relatively more prevalent in accessions from both P2/P3 and P4.

Genomic Analysis of Merged *ORSC* Datasets and Geographic Distribution

To compare the population structure of the collection of *ORSC* analyzed by Kim et al. (2016) (n = 286) with that generated by Huang et al. (2012) and subsequently analyzed by Wang et al. (2017) (n = 435), a merged SNP dataset was constructed and analyzed using fastSTRUCTURE (Raj et al., 2014). The combined dataset contained 55,213 SNPs shared between them. Structure analysis of the merged data resulted in six subpopulations that match nearly perfectly with the six subpopulations reported by Wang et al. (2017), (**Figure 4A**). The six subpopulations, designated W1 to W6 by Kim et al. (2016) have a clear connection to those designated *Or-A* to *Or-F* by Wang et al. (2017) as illustrated by shared merged groups. The main differences are due to the thresholds used to classify accessions designated as *admixed*, and the different size and composition of the two collections. Where Wang et al. (2017) used a threshold of > 80% ancestry to classify accessions as belonging to one of the wild groups, Kim et al. (2016) used a threshold of >60%. As a result, roughly half of the accessions designated as W1 by Kim et al. (2016) were identified as *admixed* in the combined analysis. It is also noteworthy that Kim et al. (2016) had a smaller proportion of wild accessions from China (W6), and a larger proportion from Papua, New Guinea (W3) and from Nepal (W5), which impacted the subpopulation structure that emerged in the analysis. Nonetheless, there is a clear correspondence between genetic groups in the two studies, with W1 and W3 (predominantly *O. rufipogon* as described by Kim et al., 2016) corresponding to *Or-A* (Wang et al., 2017), W6 (Chinese *O. rufipogon*) to *Or-B,* and W2 (predominantly *O. nivara*) to *Or-C*. *W4* (*aus*-like wild ancestor as described by Kim et al., 2016) is split between *Or-D* (wild) and *Or-E* (feral), and interestingly, the accessions clustered as *Or-E* by Wang et al. (2017) were recognized as a separate group by Kim et al. (2016) when a higher K value (K = 8) was used in the analysis.

The groups identified by Wang et al. (2017) also show reasonable correspondence to the original three groups proposed by Huang et al. (2012), (**Figure 4B**), though Huang et al. (2012) did not explicitly call out the large proportion of *ORSC* accessions that are highly admixed with *O. sativa* as both Wang et al. (2017) and Kim et al. (2016) did. Given the different compositions of the two germplasm collections, this merged genomic analysis allows us to interpret our findings about the three phenotypic groups in a larger context.

There is support for geographic subpopulation structure whereby accessions collected from the Malay Archipelago (Malaysia, Indonesia, Philippines, Papua New Guinea), which represents the southeastern extreme of the geographic distribution of the *ORSC,* are phenotypically classified as P1 in this study; they are also predominantly classified as *O. rufipogon* in the IRGC (**Supplementary Figure 16**). Accessions from India, Nepal, and Sri Lanka, representing the northwestern extreme of the geographic distribution, are predominantly classified as P4 in this study and *O. nivara* in the IRGC. Accessions collected from China were genetically classified as *O. rufipogon* with substantial admixture from *O. sativa* (Kim et al., 2016), but they were phenotypically more similar to annual *O. nivara* in this study. Accessions collected across the coastal regions of South and Southeast Asia are recognized phenotypically as P2/P3 and are

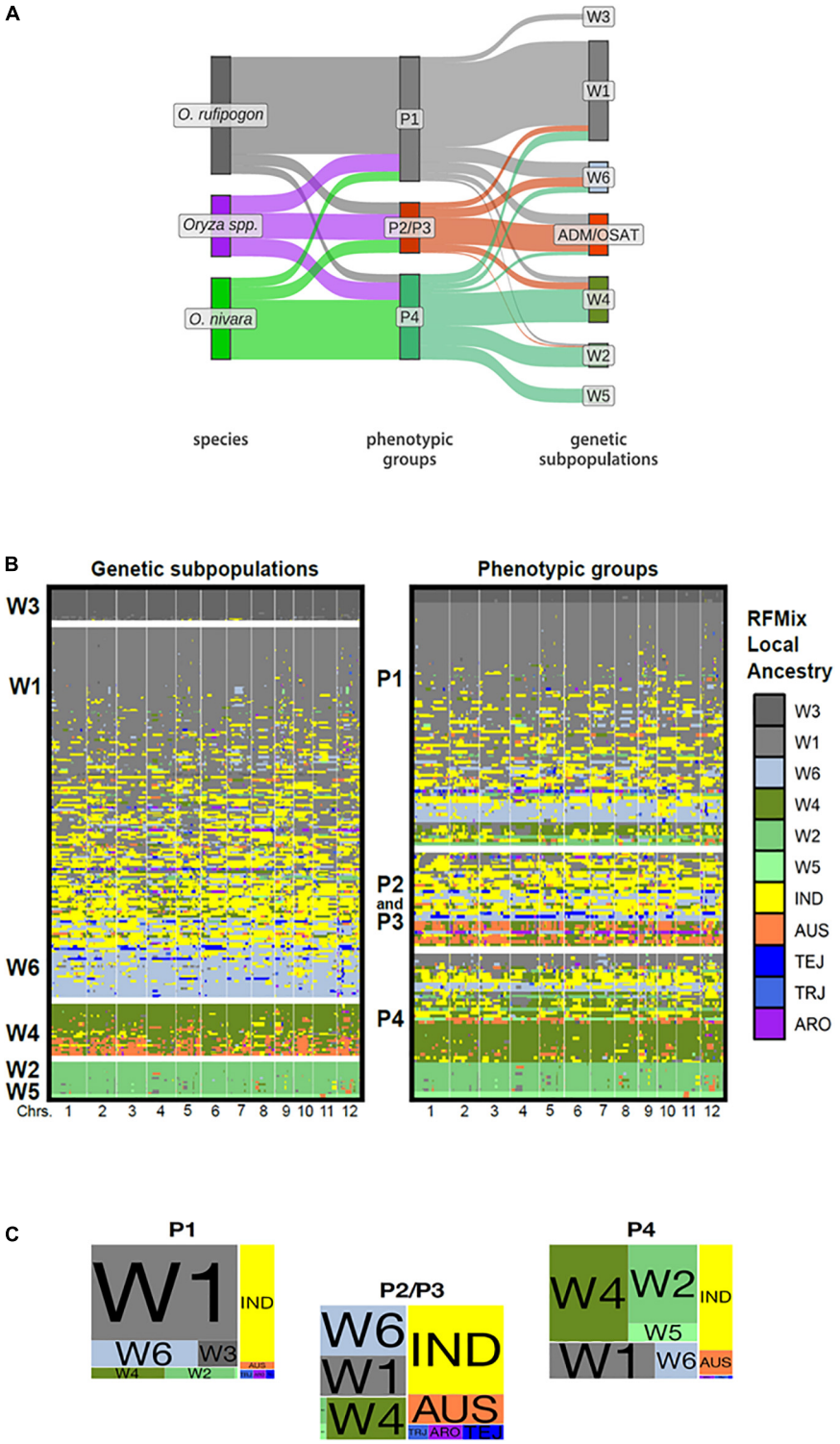

FIGURE 3 | Genotypic composition, local ancestry and phenotypic groups of *ORSC* accessions. **(A)** Sankey plot depicting correspondence between traditional species designations, our phenotypic groups, and Kim et al. (2016) genetic subpopulations using 32 phenotypic traits. (ADM/OSAT are accessions with a high proportion of admixture with *O. sativa*) **(B)** Graphical genotypes display the local ancestry of each *ORSC* individual as determined by RFMix trained on six *ORSC* subpopulations (W1-W6) and five *O. sativa* subpopulations: *indica* (IND), *aus*, (AUS), *temperate japonica* (TEJ), *tropical japonica* (TRJ), and *aromatic* (ARO). The accessions are arranged according to global genotypic ancestry (left) and phenotypic group assignment (right). Each individual's genotype is represented by two adjoining rows corresponding to phased haplotypes with colors indicating RFMix population assignments of chromosome segments **(C)** Proportion of the *ORSC* subpopulations (left) and *O. sativa* subpopulations (right) within each of the three phenotypic classes.

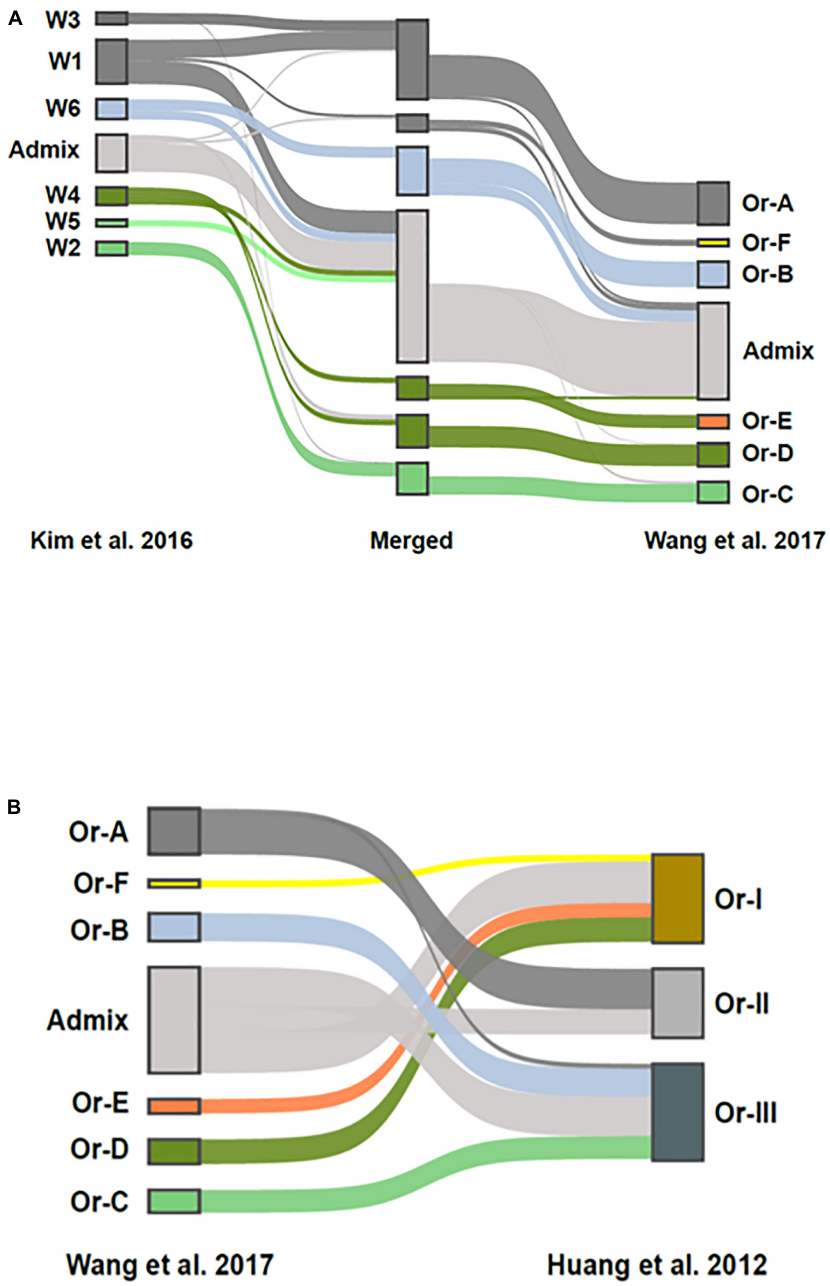

FIGURE 4 | Comparison of *ORSC* genotypic subpopulations reported by Kim et al. (2016) based on 286 *ORSC* accessions to those reported by Wang et al. (2017) and Huang et al. (2012) based on 446 *ORSC* accessions. **(A)** Population structure assignments in a merged SNP dataset consisting of accessions from two *ORSC* collections. On the left side are the accessions from Kim et al. (2016), on the right are accessions from the analysis by Wang et al. (2017). In the center is a merged set with group assignments at $K = 7$. Ribbons connect the group assignments of the merged accessions with their group assignments in the respective *ORSC* collections **(B)** Comparison of the three subpopulations originally identified by Huang et al. (2012) to the six subpopulations and admixed group identified by Wang et al. (2017) when reanalyzing the same genotypic data. The inclusion of *aus*-like (*Or-E*) and *indica*-like (*Or-F*) accessions in *Or-I* and *japonica*-like (*Or-B*) accessions in *Or-III* as reported by Huang et al. (2012) is confirmed.

typically the most highly admixed with *O. sativa*. A great diversity of phenotypic variation is found in mainland Southeast Asia where *ORSC* accessions manifesting both P1 and P4 phenotypic syndromes co-exist, often sympatrically or in contiguous environments. These geographical distributions of *O. rufipogon* and *O. nivara* are clearly shown in **Supplementary Figure 16B** where only the locations of the "core" accessions, those for which the species, phenotypic group and genotypic subpopulation agreed, are shown. The fact that cultivated forms of rice, predominantly representing the *indica* subpopulation, overlap

with wild rice habitats in many parts of Asia creates endless possibilities for cross fertilization and the emergence of novel phenotypes. In addition to their inherent value and interest, these naturally occurring interspecific populations represent valuable, ecologically viable pools of variation for plant breeders interested in identifying novel forms of climate resilience for crop improvement.

DISCUSSION

Trait Significance

The traits measured in this study emphasize highly heritable, easily measured morphological and physiological traits that can be readily observed in populations of crop wild relatives grown under controlled conditions. Many of the same traits are evaluated by genebanks and widely used for identification, tracking and management of *ex situ* germplasm collections. Of the traits measured here, four were significantly associated with phenotypic groups and with life history; these included days to heading, plant height, percent filled spikelets per panicle (seeds/panicle), and spikelet/seed size (related to grain weight). These traits provided information that was valuable for differentiating annual and perennial life habits and were similar to traits evaluated by Grillo et al. (2009) in two *O. rufipogon*/*O. nivara* mapping populations. We also measured several traits that were potentially informative about mating habit, including anther length, style length and stigma length, but only anther length discriminated the groups based on both Hotelling T^2 value and correlation estimates (**Table 1**). These traits related to life history and mating habit support our understanding of the biological significance of our phenotypic groups.

Traits That Differentiate Phenotypic Groups

The single most informative trait for distinguishing *ORSC* phenotypic groups based on the Hotelling T^2 values across all three datasets was seed width (HULGRWD) (T^2 value = 1.15) measured at Dale Bumpers, which served as a proxy for spikelet width (SPKWD) measured at IRRI, and seed volume, (T^2 value = 1.70) calculated from seed width and length. Spikelet width and seed width were highly correlated (r = 0.62) and both width and volume were informative for distinguishing groups. Seed traits were not measured at Cornell and therefore could not be used for further confirmation.

Comparing this analysis to the principle component analysis (PCA) of phenotypic data collected at IRRI on a different set of 116 *ORSC* accessions, Banaticla-Hilario et al. (2013a) also reported SPKWD as one of the most important characters separating *O. nivara* and *O. rufipogon* accessions. The other important traits included anther length (ANTLT), days to heading (DTHD) and culm length (CULT). Of note, spikelet fertility was not included in the PCA because it was highly correlated with anther length and spikelet width, leading to the conclusion that anther length and spikelet width serve as a proxy for fertility when plants are grown under controlled screenhouse conditions. Our data on DTHD collected at IRRI differentiated P1 from the other groups. The trend is the same in the Cornell and the Dale Bumpers data, but with more overlap between groups. This may be a consequence of evaluating DTHD during the short days of fall at higher latitudes. Across our three environments, unfilled grain number (UNFILGRNB) in the IRRI and Dale Bumpers data and seed weight (Seed_S2) collected from bagged and unbagged panicles at Cornell provided useful proxies for spikelet fertility. Lastly, culm length (CULT) which was measured from the soil surface to the panicle base at IRRI and plant height (PTHT) measured from the soil surface to the tip (end) of the panicle at Dale Bumpers and Cornell were prominent differentiators across all three studies. This comparison confirms that many traits distinguishing the phenotypic groups can be identified across different environments and panels of *ORSC* accessions. These observations are confirmed by earlier studies summarized in **Supplementary Table 1** in which *O. nivara* was characterized by having wider seeds, higher seed production (fertility), shorter culm length (plant height), and earlier flowering (due to photoperiod insensitivity) compared to *O. rufipogon* which was characterized by relatively narrow seeds, low seed production, longer culm length (taller plant) and later flowering (due to photoperiod sensitivity) (Morishima et al., 1961; Oka and Morishima, 1967; Barbier, 1989a).

Considering the high information content of SPKWD (spikelet/seed width) as a distinguishing feature of the phenotypic groups identified in this study and its utility for differentiating the *O. rufipogon* and *O. nivara* species, we recommend including seed width and length in the suite of traits collected on *ORSC* accessions in genebanks around the world. With the availability of scanners and appropriate software, this would be an inexpensive investment that would help improve the classification of *ORSC* accessions, making collections of these crop wild relatives more valuable to users.

"Core" *Oryza rufipogon* and *Oryza nivara* Accessions and Admixture With *Oryza sativa*

Of the *ORSC* accessions phenotyped in this study, approximately 41% (∼98 accessions, including the 49 extracted for use as training sets) had less than 5% introgression from *O. sativa* and can be considered non-admixed and wild. Most of these (n = 91) clustered in either P1 (n = 48) where most were classified as *O. rufipogon* or in P4 (n = 43) where most were classified as *O. nivara,* while a few (n = 7) clustered in P2/P3.

Based on analyses of both phenotypic and genotypic data using different model parameters, we identified 42 "core" accessions that are consistently classified as perennial *O. rufipogon* and 25 "core" accessions consistently classified as annual *O. nivara*. These "core" accessions all have <20% admixture from *O. sativa* and intersect with the group carrying <5% *O. sativa* introgression. This collection of 67 accessions provides a useful subset of *ORSC* materials for studying the genetic basis of the phenotypic divergence that distinguishes the two ecotypes of wild rice in Asia.

ORSC accessions that are highly admixed with *O. sativa* are also of interest. In this study, 25% of accessions were classified as admixed because they had >40% of the genome introgressed from *O. sativa*. In the previous study by Wang et al. (2017) using a different panel of *ORSC* accessions and a different threshold for determining admixture, 42% of wild rice samples were reported to be substantially admixed with >20% of the genome showing introgression from *O. sativa*. In both studies, *indica* and *aus* introgressions accounted for the largest proportion, and *japonica* introgressions were rare. Admixed accessions found in P1 and P4 in this study had an average of 19.8% and 19.0% of their genomes comprised of introgressions from *O. sativa,* respectively, while levels of admixture in P2/P3 averaged 52.6%; indeed, two P2/P3 accessions carried ∼85% *O. sativa* DNA (**Figure 3C** and **Supplementary Figure 15**). Nonetheless, the genome-wide level of admixture alone, is not a good predictor of phenotype, given that 13 accessions classified phenotypically into P1 carry > 40% *O. sativa* DNA, and 9 accessions classified phenotypically as P4 carry > 40% *O. sativa* DNA. We therefore infer that it is the particular distribution of *O. sativa* introgressions across the genome and the subpopulation ancestry of each introgression, in combination with the *ORSC* genetic background that determines the phenotypic outcome. Our evaluation of this *ORSC* collection provides the foundation for future research to examine the genotype-phenotype relationships for specific traits of interest.

SINE-Codes and Chloroplast Markers as Predictors of Genotypic, Species and Phenotypic Groups

In rice, SINE codes have been associated with life history (annual, intermediate or perennial growth habit) and found useful for distinguishing *O. nivara* from *O. rufipogon* (Motohashi et al., 1997; Cheng et al., 2003; Chen et al., 2004; Xu et al., 2007). Kim (2016) developed a SINE code and used it to classify most of the *ORSC* accessions used in this study (**Supplementary Table 2**). Using the SINE code, 74.5% of the accessions classified as P1 were identified as perennial (compared to 72.5% of accessions classified as *O. rufipogon*), and 62.5% of accessions classified as P4 were identified as annual or intermediate (compared to 61.7% classified as *O. nivara*). In P2/P3, 47.4% of accessions were classified as perennial and 52.6% as annual or intermediate using the SINE code (compared to 23.3% classified as *O. rufipogon*, 25.6% as *O. nivara* in P2/P3 with the remainder identified as *Oryza* spp.). Thus, the SINE-based classifications approximated the species designations in these groups, though they were not entirely concordant.

Kim et al. (2016) further analyzed 286 *ORSC* accessions for variation at 25 polymorphic sites in the chloroplast genome and identified unique chloroplast haplotypes associated with three wild subpopulation groups located at the geographical extremes of the species range. These may be of interest for investigating refuge populations of *ORSC* and elucidating phylogenetic relationships among wild populations of *Oryza* but chloroplast haplotype analysis alone was not robust enough to be useful as a predictor of genetic subpopulation, life habit or species in the *ORSC*. Nevertheless, the chloroplast data provided clear evidence of gene flow between annual and perennial populations, as well as between *ORSC* and *O. sativa* populations, and confirmed that it occurs through both pollen dissemination and seed dispersal (likely facilitated by human migration). Taken together, these data suggest that the population structure of the *ORSC* is the result of complex evolutionary pathways that intersect and loop back upon each other due to highly permeable 'species' boundaries and, as documented in this study, highly plastic phenotypic variation. Information about SINE-code and chloroplast haplotype variation (as determined by Kim et al., 2016) corresponding to each accession is summarized in **Supplementary Table 2**. These features, along with geographical and ecological information about where each accession was collected, disease and insect resistance, grain quality and additional use-data provided by the IRGC database add value to these genetic resources and enhance our understanding of where they come from and how they might be used in the future.

Emergent Phenotypes

The P2/P3 group is characterized by 43 accessions with trait combinations that form a coherent group, despite the fact that they do not match either of the *a priori* taxonomic species descriptions. This group is comprised of 29 (67%) admixed *ORSC* accessions with > 40% of the genome derived from *O. sativa* but it also contains accessions classified as annual *O. nivara* or perennial *O. rufipogon*. The fact that the P2/P3 accessions were collected in diverse ecological and geographic regions, especially from coastal regions of China, Southeast and South Asia (**Supplementary Figure 16**), and yet share a common suite of emergent phenotypes suggests that particular combinations of traits may evolve repeatedly and independently when annual, perennial and intermediate forms of common wild rice come together as sympatric swarms in unsupervised settings. The patterns of admixture observed in this group suggest that diverse *ORSC* populations have hybridized among themselves and with local populations of *O. sativa,* giving rise to forms of trait variation not observed in the parental populations. In rice, several lines of evidence suggest that annual and intermediate forms have emerged repeatedly from genetically and geographically diverse populations of perennial ancestors in response to changing patterns of temperature, rainfall and CO_2 (Grillo et al., 2009; Banaticla-Hilario et al., 2013a,b; Liu et al., 2015). The *ORSC* populations examined here provide opportunities to deepen our understanding of incipient speciation and to identify mechanisms by which annual life forms may evolve from perennial ancestors in response to changes in the environment. They also provide material for examining what kinds of selection pressure(s) are associated with a shift toward the annual, early-flowering, seed-bearing habit or conversely, serve to maintain the perennial, stoloniferous, late-flowering and vegetatively vigorous habit. During times of climate change, a strategy of balancing selection to maintain components of annual, intermediate and perennial life forms may be favored by evolution to allow for the emergence of new resilience mechanisms. We have a lot to learn from the existence of these dynamic *ORSC* populations. They have survived many waves of climate change in the past and are likely to hold secrets about survival strategies for the future.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

KM, MEN, MB-H, GCE, HJK, JJ, SH, JK, and LH phenotyped the plants. YS managed the data. AG, JE, GCE, HJK, JJ, and SRM analyzed the data. GCE, AG, JE, and SRM wrote the manuscript. SRM conceptualized the project. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded in part by the National Science Foundation with grants to SRM and GCE (NSF-PGRP nos. 0606461 and 1026555). Mention of a trademark or proprietary product does not constitute a guarantee or warranty of the product by the US Department of Agriculture and does not imply its approval to the exclusion of other products that also can be suitable. The USDA is an equal opportunity provider and employer.

ACKNOWLEDGMENTS

The excellent technical assistance of Daniel E. Wood in collecting the phenotypic data and the assistance of Teresa A. Hancock with organizing and conducting the preliminary analyses of the data collected at the Dale Bumpers National Rice Research Center, is acknowledged. The assistance of Aaron K. Jackson with the manuscript references is acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.787703/full#supplementary-material

REFERENCES

Banaticla-Hilario, M. C. N., Sosef, M. S. M., McNally, K. L., Hamilton, N. R. S., and van den Berg, R. G. (2013a). Ecogeographic variation in the morphology of two Asian wild rice species, *Oryza nivara* and *Oryza rufipogon*. *Int. J. Plant Sci.* 174, 896–909. doi: 10.1086/670370

Banaticla-Hilario, M. C., van den Berg, R. G., Hamilton, N. R., and McNally, K. L. (2013b). Local differentiation amidst extensive allele sharing in *Oryza nivara* and *O. rufipogon*. *Ecol. Evol.* 3, 3047–3062. doi: 10.1002/ece3.689

Barbier, P. (1989a). Genetic variation and ecotypic differentiation in the wild rice species *Oryza rufipogon*. I. Population differentiation in life-history traits and isozymic loci. *Jpn. J. Genet.* 64, 259–271. doi: 10.1266/jjg.64.259

Barbier, P. (1989b). Genetic variation and ecotypic differentiation in the wild rice species *Oryza rufipogon* II. Influence of the mating system and life-history traits on the genetic structure of populations. *Jpn. J. Genet.* 64, 273–285. doi: 10.1266/jjg.64.273

Barbier, P., Morishima, H., and Ishihama, A. (1991). Phylogenetic relationships of annual and perennial wild rice: probing by direct DNA sequencing. *Theor. Appl. Genet.* 81, 693–702. doi: 10.1007/bf00226739

Bioversity International, IRRI, and Africa Rice Center (2007). *Descriptors for Wild and Cultivated Rice (Oryza spp.)*. Rome: Bioversity International.

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* 103, 338–348. doi: 10.1016/j.ajhg.2018.07.015

Cai, H. W., Wang, X. K., and Morishima, H. (2004). Comparison of population genetic structures of common wild rice (*Oryza rufipogon* Griff.), as revealed by analyses of quantitative traits, allozymes, and RFLPs. *Heredity (Edinb.)* 92, 409–417. doi: 10.1038/sj.hdy.6800435

Chang, T. T. (1976). The origin, evolution, cultivation, dissemination, and diversification of Asian and African rices. *Euphytica* 25, 425–441. doi: 10.1007/BF00041576

Chatterjee, D. (1948). A modified key and enumeration of the species of *Oryza* Linn. *Indian J. Agric. Sci.* 18, 185–192.

Chen, L. J., Lee, D. S., Song, Z. P., Suh, H. S., and Lu, B. R. (2004). Gene flow from cultivated rice (*Oryza sativa*) to its weedy and wild relatives. *Ann. Bot.* 93, 67–73. doi: 10.1093/aob/mch006

Cheng, C., Motohashi, R., Tsuchimoto, S., Fukuta, Y., Ohtsubo, H., and Ohtsubo, E. (2003). Polyphyletic origin of cultivated rice: based on the interspersion pattern of SINEs. *Mol. Biol. Evol.* 20, 67–75. doi: 10.1093/molbev/msg004

Choi, J. Y., and Purugganan, M. D. (2018). Multiple origin but single domestication led to *Oryza sativa*. *G3* 8, 797–803. doi: 10.1534/g3.117.300334

Civán, P., Craig, H., Cox, C. J., and Brown, T. A. (2015). Three geographically separate domestications of Asian rice. *Nat. Plants* 1:15164. doi: 10.1038/nplants.2015.164

Cooper, L., Meier, A., Laporte, M. A., Elser, J. L., Mungall, C., Sinn, B. T., et al. (2018). The Planteome database: an integrated resource for reference ontologies, plant genomics and phenomics. *Nucleic Acids Res.* 46, D1168–D1180. doi: 10.1093/nar/gkx1152

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

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*. Boca Raton, FL: CRC Press.

Greenberg, A. J., Hackett, S. R., Harshman, L. G., and Clark, A. G. (2010). A hierarchical Bayesian model for a novel sparse partial diallel crossing design. *Genetics* 185, 361–373. doi: 10.1534/genetics.110.115055

Greenberg, A. J., Hackett, S. R., Harshman, L. G., and Clark, A. G. (2011). Environmental and genetic perturbations reveal different networks of metabolic regulation. *Mol. Syst. Biol.* 7:563. doi: 10.1038/msb.2011.96

Grillo, M. A., Li, C., Fowlkes, A. M., Briggeman, T. M., Zhou, A., Schemske, D. W., et al. (2009). Genetic architecture for the adaptive origin of annual wild rice, *Oryza nivara*. *Evolution* 63, 870–883. doi: 10.1111/j.1558-5646.2008.00602.x

Hill, R. (2010). The cultivation of perennial rice, an early phase in Southeast Asian agriculture? *J. Hist. Geogr.* 36, 215–223. doi: 10.1016/j.jhg.2009.09.001

Hoffman, G. E. (2013). Correcting for population structure and kinship using the linear mixed model: theory and extensions. *PLoS One* 8:e75707. doi: 10.1371/journal.pone.0075707

Hotelling, H. (1931). The generalization of student's ratio. *Ann. Math. Stat.* 2, 360–378,319.

Huang, X., Kurata, N., Wei, X., Wang, Z.-X., Wang, A., Zhao, Q., et al. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490, 497–501. doi: 10.1038/nature11532

Jung, J. K. H. (2016). *Scientific Capacity Building: Enhancing our Understanding of Wild Rice Germplasm and Human Resource Development*. Doctoral dissertation. Ithaca, NY: Cornell University.

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-Y., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42, 348–354. doi: 10.1038/ng.548

Kennedy, A. B. W., and Sankey, H. R. (1898). The thermal efficiency of steam engines. Report of the committee appointed to the council upon the subject of the definition of a standard or standards of thermal efficiency for steam engines: with an introductory note. *Minutes Proc. Inst. Civil Eng.* 134, 278–312. doi: 10.1680/imotp.1898.19100

Kim, H. (2016). *Genetic Characterization of the O. rufipogon Species Complex and Construction of Interspecific Pre-Breeding Resources for the Rice Improvement*. Doctoral dissertation. Ithaca, NY: Cornell University.

Kim, H., Jung, J., Singh, N., Greenberg, A., Doyle, J. J., Tyagi, W., et al. (2016). Population dynamics among six major groups of the *Oryza rufipogon* species complex, wild relative of cultivated Asian rice. *Rice (N. Y.)* 9:56. doi: 10.1186/s12284-016-0119-0

Kurata, N., and Yamazaki, Y. (2006). Oryzabase. An integrated biological and genome information database for rice. *Plant Physiol.* 140, 12–17. doi: 10.1104/pp.105.063008

Kuroda, Y., Sato, Y.-I., Bounphanousay, C., Kono, Y., and Tanaka, K. (2006). Genetic structure of three *Oryza* AA genome species (*O. rufipogon*, *O. nivara* and *O. sativa*) as assessed by SSR analysis on the Vientiane Plain of Laos. *Conserv. Genet.* 8, 149–158. doi: 10.1007/s10592-006-9156-3

Liu, R., Zheng, X. M., Zhou, L., Zhou, H. F., and Ge, S. (2015). Population genetic structure of *Oryza rufipogon* and *Oryza nivara*: implications for the origin of *O. nivara*. *Mol. Ecol.* 24, 5211–5228. doi: 10.1111/mec.13375

Londo, J. P., Chiang, Y.-C., Hung, K.-H., Chiang, T.-Y., and Schaal, B. A. (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9578–9583. doi: 10.1073/pnas.0603152103

Lynch, M., and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer Associates, Inc.

Maples, B. K., Gravel, S., Kenny, E. E., and Bustamante, C. D. (2013). RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *Am. J. Hum. Genet.* 93, 278–288. doi: 10.1016/j.ajhg.2013.06.020

McGrory, C. A., and Titterington, D. M. (2007). Variational approximations in Bayesian model selection for finite mixture distributions. *Comput. Stat. Data Anal.* 51, 5352–5367. doi: 10.1016/j.csda.2006.07.020

Morishima, H. (2001). "Evolution and domestication of rice," in *Proceedings of the 4th International Rice Genetics Symposium Rice Genetics IV*, eds G. S. Khush, D. S. Brar, and B. Hardy (Los Baños, PH: International Rice Research Institute), 63–77. doi: 10.1142/9789812814296_0004

Morishima, H., Oka, H., and Chang, W. (1961). Directions of differentiation in populations of wild rice, *Oryza perennis* and *O. sativa F. spontanea*. *Evolution (N. Y.)* 15, 326–339. doi: 10.1111/j.1558-5646.1961.tb03158.x

Morishima, H., Sano, Y., and Oka, H. I. (1984). Differentiation of perennial and annual types due to habitat conditions in the wild rice *Oryza perennis*. *Plant Syst. Evol.* 144, 119–135. doi: 10.1007/BF00986670

Motohashi, R., Mochizuki, K., Ohtsubo, H., and Ohtsubo, E. (1997). Structures and distribution of p-SINE1 members in rice genomes. *Theor. Appl. Genet.* 95, 359–368. doi: 10.1007/s001220050571

Oka, H. (1974). Analysis of genes controlling F_1 sterility in rice by the use of isogenic lines. *Genetics* 77, 521–534. doi: 10.1093/genetics/77.3.521

Oka, H. I. (1988). *Origin of Cultivated Rice*. Tokyo: Japan Scientific Societies.

Oka, H. I., and Morishima, H. (1967). Variations in the breeding system of a wild rice, *Oryza perennis*. *Evolution* 21, 249–258. doi: 10.1111/j.1558-5646.1967.tb00153.x

Raj, A., Stephens, M., and Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589. doi: 10.1534/genetics.114.164350

Robert, C. P. (1996). "Mixtures of distributions: inference and estimation," in *Markov Chain Monte Carlo in Practice*, eds W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (Boca Raton, FL: Chapman and Hall/CRC), 441–464.

Roschevicz, R. J. (1931). A contribution to the knowledge of rice. *Appl. Bot. Genet. Pl. Breed. Bull.* 27, 3–133.

Samal, R., Roy, P. S., Sahoo, A., Kar, M. K., Patra, B. C., Marndi, B. C., et al. (2018). Morphological and molecular dissection of wild rices from eastern India suggests distinct speciation between *O. rufipogon* and *O. nivara* populations. *Sci. Rep.* 8:2773. doi: 10.1038/s41598-018-20693-7

Sano, Y., Morishima, H., and Oka, H.-I. (1980). Intermediate perennial-annual populations of *Oryza perennis* found in Thailand and their evolutionary significance. *Bot. Mag.* 93, 291–305. doi: 10.1007/BF02488735

Sharma, S. (2003). "Species of genus *Oryza* and their interrelationships," in *Monograph on Genus Oryza*, J. S. Nanda and S. D. Sharma (Enfield, NH: Science Publishers).

Sharma, S., and Shastry, S. (1965). Taxonomic studies in Genus *Oryza* I. Asiatic types of sativa complex. *Indian J. Genet. Plant Breed.* 25, 245–259.

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980

Vaughan, D. A. (1989). The genus *Oryza* L.: current status of taxonomy. *IRRI Res. Pap. Ser.* 138, 1–21.

Vaughan, D. A. (1994). *The Wild Relatives of Rice: A Genetic Resources Handbook*. Los Banos, PH: International Rice Research Institute.

Vaughan, D. A., Lu, B.-R., and Tomooka, N. (2008). Was Asian rice (*Oryza sativa*) domesticated more than once? *Rice* 1, 16–24. doi: 10.1007/s12284-008-9000-0

Vaughan, D. A., Morishima, H., and Kadowaki, K. (2003). Diversity in the *Oryza* genus. *Curr. Opin. Plant Biol.* 6, 139–146. doi: 10.1016/s1369-5266(03)00009-8

Vaughan, L. K., Ottis, B. V., Prazak-Havey, A. M., Bormans, C. A., Sneller, C., Chandler, J. M., et al. (2001). Is all red rice found in commercial rice really *Oryza sativa*? *Weed Sci.* 49, 468–476.

Wang, H., Vieira, F. G., Crawford, J. E., Chu, C., and Nielsen, R. (2017). Asian wild rice is a hybrid swarm with extensive gene flow and feralization from domesticated rice. *Genome Res.* 27, 1029–1038. doi: 10.1101/gr.204800.116

Wang, H., Xu, X., Vieira, F. G., Xiao, Y., Li, Z., Wang, J., et al. (2016). The power of inbreeding: NGS-based GWAS of rice reveals convergent evolution during rice domestication. *Mol. Plant* 9, 975–985. doi: 10.1016/j.molp.2016.04.018

Xu, J. H., Cheng, C., Tsuchimoto, S., Ohtsubo, H., and Ohtsubo, E. (2007). Phylogenetic analysis of *Oryza rufipogon* strains and their relations to *Oryza sativa* strains by insertion polymorphism of rice SINEs. *Genes Genet. Syst.* 82, 217–229. doi: 10.1266/ggs.82.217

Yamazaki, Y., Sakaniwa, S., Tsuchiya, R., Nonomura, K.-I., and Kurata, N. (2010). Oryzabase: an integrated information resource for rice science. *Breed. Sci.* 60, 544–548. doi: 10.1270/jsbbs.60.544

Zheng, X. M., and Ge, S. (2010). Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Mol. Ecol.* 19, 2439–2454. doi: 10.1111/j.1365-294x.2010.04674.x

Conflict of Interest: AG is employed by Bayesic Research LLC.

The remaining 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

REVIEW
published: 27 June 2022
doi: 10.3389/fpls.2022.908342

Understanding the Nature of Hybrid Sterility and Divergence of Asian Cultivated Rice

*Yu Zhang[1], Jie Wang[1,2], Qiuhong Pu[1], Ying Yang[1], Yonggang Lv[1], Jiawu Zhou[1], Jing Li[1], Xianneng Deng[1], Min Wang[1,2] and Dayun Tao[1]**

[1] Yunnan Key Laboratory for Rice Genetic Improvement, Food Crops Research Institute, Yunnan Academy of Agricultural Sciences (YAAS), Kunming, China, [2] Institute of Plant Resources, Yunnan University, Kunming, China

Intraspecific hybrid sterility is a common form of postzygotic reproductive isolation in Asian cultivated rice, which is also the major obstacle to utilize the strong heterosis in the rice breeding program. Here, we review recent progress in classification and hybrid sterility in Asian cultivated rice. A genome-wide analysis of numerous wild relatives of rice and Asian cultivated rice has provided insights into the origin and differentiation of Asian cultivated rice, and divided Asian cultivated rice into five subgroups. More than 40 conserved and specific loci were identified to be responsible for the hybrid sterility between subgroup crosses by genetic mapping, which also contributed to the divergence of Asian cultivated rice. Most of the studies are focused on the sterile barriers between *indica* and *japonica* crosses, ignoring hybrid sterility among other subgroups, leading to neither a systematical understanding of the nature of hybrid sterility and subgroup divergence, nor effectively utilizing strong heterosis between the subgroups in Asian cultivated rice. Future studies will aim at identifying and characterizing genes for hybrid sterility and segregation distortion, comparing and understanding the molecular mechanism of hybrid sterility, and drawing a blueprint for intraspecific hybrid sterility loci derived from cross combinations among the five subgroups. These studies would provide scientific and accurate guidelines to overcome the intraspecific hybrid sterility according to the parent subgroup type identification, allowing the utilization of heterosis among subgroups, also helping us unlock the mysterious relationship between hybrid sterility and Asian cultivated rice divergence.

Keywords: ***Oryza sativa*** **L., hybrid sterility, subgroup, divergence, heterosis**

OPEN ACCESS

Edited by:
Yongzhong Xing,
Huazhong Agricultural University, China

Reviewed by:
Wen Yao,
Henan Agricultural University, China
Xianjin Qiu,
Yangtze University, China

***Correspondence:**
Dayun Tao
taody12@aliyun.com

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

Received: *30 March 2022*
Accepted: *30 May 2022*
Published: *27 June 2022*

Citation:
Zhang Y, Wang J, Pu Q, Yang Y, Lv Y, Zhou J, Li J, Deng X, Wang M and Tao D (2022) Understanding the Nature of Hybrid Sterility and Divergence of Asian Cultivated Rice. Front. Plant Sci. 13:908342. doi: 10.3389/fpls.2022.908342

INTRODUCTION

Biological species form when the gene flow between two divergent populations has mostly or completely limited, and the striking characteristic of species is that species can actually or potentially interbreed, but are prevented from producing the fertile progeny with other species (Paterson, 1985). Reproductive isolation is critical for the formation and maintenance of species between different populations (Oka, 1988). It could occur at different developmental stages, including prezygotic and postzygotic reproductive isolation, the former prevents mating or formation of hybrid and the latter reduces the fitness of hybrid (Seehausen et al., 2014). Hybrid sterility is a major form of postzygotic reproductive isolation in rice and has been a good example

of genetic studies of population speciation as an intrinsic genetic factor (McDermott and Noor, 2010; Ouyang et al., 2010; Ouyang and Zhang, 2013).

Genus *Oryza* contains 23 wild species and two cultivated rice species with 11 genome types. Of the 11 genome types, 6 are diploid, including AA, BB, CC, EE, FF, and GG, and five are polyploidy, including BBCC, CCDD, HHJJ, HHKK, and KKLL (Khush, 1997; Stein et al., 2018). Six wild species of *Oryza nivara*, *Oryza rufipogon*, *Oryza barthii*, *Oryza glumaepatula*, *Oryza longistaminata*, and *Oryza meridionalis*, as well as the two cultivated species of *Oryza sativa* and *Oryza glaberrima* were classified into the AA genome (Khush, 1997). Thus, AA genome species were considered as the natural germplasm pool for rice genetic improvement (Vaughan, 1994). However, hybrid sterility occurs widely in intraspecific and interspecific hybrids in rice, which is the major barrier to utilize heterosis and make genetic improvement difficult by restricting gene exchange in the Asian cultivated rice, African cultivated rice, and between the two cultivated species and the wild relatives (Ouyang and Zhang, 2013; Li et al., 2020).

This review summarized recent progress in understanding the genetic differentiation in Asian cultivated rice, expatiated on the genetic basis of hybrid sterility between different subpopulations, discussed the role of hybrid sterility on the population divergence, as well as tried to outline the future efforts to be done in intraspecific hybrid sterility study.

GENETIC DIFFERENTIATION OF THE ASIAN CULTIVATED RICE

Asian cultivated rice (*O. sativa* L.) is widely cultivated around the world and mainly in Asia. Abundant genetic diversity and significant genetic differentiation due to the different geographic or ecological adaptation, as well as domestication, had been reported (Huang et al., 2011). The two distinguishable groups were recognized as "*hsien*" and "*Keng*" in Chinese ancient records from the Han dynasty (Ting, 1949a,b,c). These two different groups of *O. sativa* subs. *japonica* (*Keng/Geng*) and *indica* (*hsien/Xian*) were first named based on their morphological and serological characteristics, together with hybrid fertility differences (Kato et al., 1928), consistent with the classification by the ancient Chinese people. Then, distinct classification results emerged in view of different criteria. The three types, A, B, and C, were used to classify *O. sativa* according to the different geographical distribution, later regarded as *japonica*, *javanica,* and *indica* (Matsuo, 1952; Morinaga, 1954). Further studies indicated that Asian cultivated rice should be classified into three types and five subgroups according to hybrid affinity ability and geographic distributions. Japanese rice belonged to the *japonica* type, aman and tjereh subgroup belonged to *indica* type because of the low affinity of the Japanese subgroup, whereas aus together with bulu subgroup could be reconsidered as neither *japonica* nor *indica*, because they showed considerable high hybrid fertility to both types, so they were called the intermediate type. Thus, *O. sativa* might be grouped into three types: *japonica*, *indica,* and intermediate types (Morinaga and Kuriyama, 1955, 1958). A fast and feasible criterion for distinguishing *Keng* (*japonica*) from *hsien* (*indica*) based on some classification traits was established, such as apiculus hair length, grain size, hull color at heading, phenol reaction, and other traits, some uncertain intermediate varieties were assessed followed by traits of leaf hair and 1–2 internode length of the main panicle. This method, called the "Cheng's index," was effective in practice (Cheng et al., 1984). Isozyme was also used to study the differentiation of *O. sativa* at the biochemical level. One thousand six hundred and eighty-eight Asian varieties were detected by 15 isozyme markers at eight loci, the result indicated that Asian cultivated rice was divided into six different groups: I, II, III, IV, V, and VI, respectively, corresponding *indica*, *aus*, *deepwater* rice in Bangladesh, *deepwater* rice in Northeast India, *basmati*, and *japonica* (Glaszmann, 1987). Reviewing the previous results, nearly all the researchers agreed with the opinion that Asian cultivated rice contained two major groups: *japonica* and *indica*, but the classification of intermediate groups was ambiguous. The limited number or geographic distribution of varieties, evaluated by a few traits and genetic markers, ignoring the role of origin, domestication, and selection on the divergence of *O. sativa*, resulted in the un-uniformed classification of *O. sativa*.

Progress in the development of a great number of molecular markers and genome sequencing provided a new insight into understanding the differentiation of *O. sativa* at the molecular level. Two hundred and thirty-four rice varieties representing a wide geographic range of *O. sativa* were genotyped using 169 nuclear SSRs and two chloroplast loci; the results indicated that five distinct groups were detected, corresponding to temperate *japonica*, tropical *japonica*, *indica*, *aus*, and *aromatic* rice (Garris et al., 2005). Twenty diverse varieties and landraces by genotyping 160,000 non-redundant SNPs not only fitted well with five groups of *O. sativa*, but also demonstrated genomic introgression from one varietal group into another under strong artificial selection for production or environment adaptation demands (McNally et al., 2009). The analysis of the genetic variation of 3,010 Asian cultivated rice accessions by genome-wide SNP supported the classification of the five major groups in *O. sativa*, but also revealed seven unreported and geographic distribution-related subgroups (Wang et al., 2018). Population genetics of 5,152 rice accessions analyzed by resequencing and genetic diversity analysis of 60 rice varieties using 190 SNPs had clearly demonstrated the genetic difference among temperate *japonica*, tropical *japonica*, *indica*, *aus*, and *aromatic* rice (Kishor et al., 2020; Yan et al., 2020). Taken together, the population structure and genetic diversity of the worldwide rice gave insights into the five distinct subgroups; this classification of *O. sativa* was widely recognized by rice scientists.

Basmati rice is the typical representative of the *aromatic* type, and the fragrance trait can be found in *basmati* (Dom-sufid from Iran), *indica* (KDM105 from Thailand), and tropical *japonica* (Azucena from Philippines) (Bradbury et al., 2005; Bourgis et al., 2008; Kovach et al., 2009; Myint et al., 2012); therefore, the word "*aromatic*" representing one of the five subgroups is easily confused. Here, we use *basmati* to be the *japonica*-clinous rice from Myanmar to South Asia and Iran so as to discriminate

aromatic rice from *indica*, tropical *japonica*, temperate *japonica*, and *aus*.

In addition, a high-throughput genome sequence, coupled with archaeobotanical, paleoclimate, and geographic data, revealed the origin and domestication history of the five subgroups of *O. sativa,* thus supporting multiple origins and at least one domestication in Asian cultivated rice, which was accepted by scientists (Zhao, 2011; Huang et al., 2012; Gross and Zhao, 2014; Civan et al., 2015; Travis et al., 2015; Stein et al., 2018; Choi et al., 2020; Gutaker et al., 2020; **Figure 1**), but this classification was not proved by reproductive isolation.

GENETIC AND MOLECULAR BASIS OF HYBRID STERILITY IN ASIAN CULTIVATED RICE

Main Features of Hybrid Sterility in Asian Cultivated Rice

Hybrid sterility is widely observed in rice, causing reduced gamete viability of hybrids between divergent subspecies. Hybrid sterility between *japonica* and *indica* was first reported when Asian cultivated rice was classified into the two subspecies using the criterion of hybrid fertility (Kato et al., 1928). Since then, the genetic basis of hybrid sterility had been investigated and scientists found some distinguished features in hybrid sterility: (1) Cytoplasm had almost no effect on hybrid fertility because reciprocal crosses showed no significant difference in male and female gametes fertility of F_1 plants (Kato, 1930; Oka, 1957); (2) F_1 hybrids between *indica* and *japonica* were usually sterile, the hybrid fertility varied dependent on the number of hybrid sterility genes and genetic distance between the two parents (Oka, 1988; Zhang et al., 1989, 1994); (3) Sterility could be explained by hybrid sterility genes instead of chromosome disturbance (Kato, 1930); (4) The F_1 and F_2 fertilities are not correlated (Oka and Doida, 1962); (5) Hybrid sterility between *japonica* and *indica* could be observed due to male gamete abortion, female gamete abortion, and embryo abortion (Teng et al., 1996; Zhu et al., 1996; Liu et al., 2004); (6) Hybrid gene interactions, including both allelic interaction at the same locus and non-allelic interactions at the different loci, were responsible for varying degrees of hybrid sterility in crosses between subgroups of *O. sativa* (Zhang and Lu, 1993; Zhang et al., 1993; Kubo et al., 2008).

Identification, Cloning, and Molecular Characterization of Hybrid Sterility Loci

By now, more than 40 loci responsible for hybrid sterility in the crosses between the different subgroups are identified by genetic mapping in rice (**Supplementary Table 1**). Surprisingly, three loci, *Ef1*, *Pf3,* and *Sf3*, were found in the cross between ZS97 and Minghui 63, which were famous *indica* parents in hybrid rice breeding. It is suggested that hybrid sterility also happened within the *indica* cross (**Supplementary Table 1**). But it was notable that ZS97 and Minghui 63 were not traditional varieties, they were the product of modern breeding incorporating other genome introgression fragments from other subgroups by multiple round crosses (McNally et al., 2009); therefore, cloning and gene sequence analysis of three loci will answer whether those loci were responsible for hybrid sterility in *indica* × *indica* cross or not. Moreover, there is no gene or QTL identified in the cross between *basmati* and other subgroups, further study will pay more attention to identifying, cloning, and characterizing the molecular mechanism of hybrid sterile genes from *basmati* rice.

Until now, six loci for intraspecific hybrid sterility were cloned and characterized. *Sa* was the first characterized hybrid male sterility loci between *indica* and *japonica*. SaM encoded a small ubiquitin-like modifier E3 ligase, and SaF encoded an F-box protein. Three divergent genetic components: SaF^+ (the *indica* allele), SaM^+ (the *indica* allele), and SaM^- (the *japonica* allele) consisted of a killer system to selectively abort the male gametes harboring SaM^- in the *indica*×*japonica* hybrid (Long et al., 2008). *S5* locus conferring female hybrid sterility consisted of three tightly linked genes, *ORF3*, *ORF4,* and *ORF5*, which acted in a killer–protector system. ORF3, ORF4, and ORF5 encoded an HSP70 heat shock protein, a transmembrane protein, and a putative aspartic protease, respectively. During female development, endoplasmic reticulum (ER) stress is caused by the action of *ORF5*+ (killer) and *ORF4*+ (partner), and *ORF3*+ (protector) could prevent this ER stress, but *ORF3*– could not prevent ER stress, triggering premature programmed cell death, and leading to embryo-sac abortion (Chen et al., 2008; Yang et al., 2012). A tetratricopeptide repeat domain-containing protein encoded by *S7* was required for hybrid female sterility in the crosses between *aus* and *indica* or *japonica* (Yu et al., 2016). Two tightly linked genes, *HSA1a* and *HSA1b*, were responsible for hybrid female sterility in a cross between temperate *japonica* and *indica*. *HSA1a* encoded an unknown function protein with a highly conserved plant-specific domain, whereas *HSA1b*, encoded an uncharacterized protein with the nucleotide-binding domain. Heterozygous *HSA1a* with homozygous *HSA1b* allele from *indica* conferred hybrid sterility due to selective abortion of $HSA1a\text{-}i^s$ gamete in the hybrid (Kubo et al., 2016a). Allelic suppression mediated by *Sc* contributed to hybrid male sterility between temperate *japonica* and *indica.* The *japonica* allele, *Sc-j*, encoded a DUF1618-domain protein; the *indica* allele, *Sc-i*, consisted of two or three tandem-duplicated *Sc-j* homolog. The selective abortion of *Sc-j* pollen leading to the transmission ratio distortion resulted from the expression of *Sc-j* suppressed by high expression of *Sc-i* in *Sc-j*/*Sc-i* hybrids (Shen et al., 2017). Hybrid pollen germination caused by reciprocal gene *DPL1*/*DPL2* was isolated from *japonica* and *aus* cross. DPL encoded a plant-specific small protein. The pollen from Nipponbare possessed $DPL1\text{-}N^+$/ $DPL2\text{-}N^-$ genotype, whereas pollen from Kasalath carried $DPL1\text{-}K^-$/ $DPL2\text{-}K^+$ genotype. Pollen harboring two nonfunctional alleles $DPL1\text{-}K^-$ and $DPL2\text{-}N^-$ failed to germinate and was not transmitted to the next generation (Mizuta et al., 2010). According to the above information, it is easy to find that the hybrid sterility of the different genetic materials was controlled by diverse genes or QTLs, and allelic interaction at each locus and non-allelic interaction at the different loci from two parents contribute to the intraspecific hybrid sterility. These genes or QTLs identified in the different populations will lay the foundation for elucidating the genetic and molecular mechanisms of hybrid sterility in Asian cultivated rice.

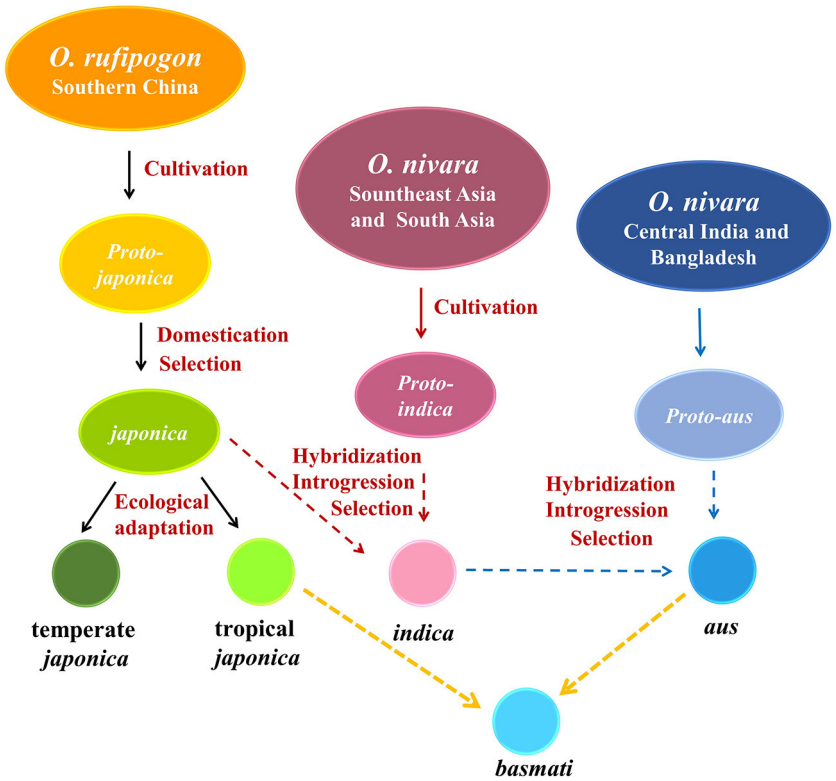

FIGURE 1 | Origin and domestication model of Asian cultivated rice. Archaeological and genetic evidence indicated that *japonica* rice originated from a *O. rufipogon* population around the middle area of the Pear River and began to be cultivated about 9,000 years BP in the Yangtze Valley in southern China; *japonica* spread to South Asia about 4,000 years BP leading to the introgression of domestication genes into *pro-indica* derived from local *O. nivara* in Southeast Asia and South Asia, and *indica* was then formed to be cultivated in the lower Ganges valley (Zhao, 2011; Huang et al., 2012; Gross and Zhao, 2014; Stein et al., 2018). Tropical *japonica* and temperate *japonica* differed genetically due to a global reduction in temperature and humidity; it was the result of strong pressure to adapt to different climatic conditions (Kovach et al., 2007; Gutaker et al., 2020). Recently, two additional subspecies of Asian cultivated rice, *aus* and *basmati*, were recognized to differentiate distinctly from *japonica* and *indica*. *O. nivara* in central India and Bangladesh as the gene pool crossed with *indica* and selected to adapt to rainfed conditions, including rainfed upland, rainfed lowland, and deep-water, gave rise to *aus* group, whereas *aromatic* was the hybrid of *aus* and tropical *japonica* and was popular in Pakistan, northern India (*basmati*), and Iran (sadri) (Khush et al., 2013; Civan et al., 2015; Travis et al., 2015; Choi et al., 2020).

Two genetic models were popular for explaining hybrid sterility. One model was the one-locus sporo-gametophytic interaction model (Kitamura, 1962), and another was a duplicate gametic lethal model interpreting two independent loci involved in the hybrid sterility (Oka, 1957, 1974). *S5*, *S7*, *Sa*, *Sc,* and *HSA1* fitted in with the one-locus sporo-gametophytic interaction model, and *DPL1/DPL2* followed the duplicate gametic lethal model (**Figure 2**; Long et al., 2008; Mizuta et al., 2010; Yang et al., 2012; Kubo et al., 2016a; Yu et al., 2016; Shen et al., 2017).

Epistatic Interaction Between the Different Hybrid Sterility Loci

Hybrid sterility is a complex quantitative trait controlled by multiple loci, hence, it is conceived that gene–gene interaction is related to the sterility mechanism. Epistatic interaction between *S35* and *S24* regulated male sterility in rice intraspecific hybrid. *S24* locus in the heterozygous plants caused abortion of male gametes from the *japonica* allele, independent of the *S35* genotype, whereas male gametes from *japonica* alleles at the *S35* were aborted when the homozygous genomic fragment harboring the *S24* locus was introgressed into the heterozygous *S35* background plants (Kubo et al., 2000, 2008). Further study confirmed that a locus *INK* tightly linked with *S24*, not *S24*, activated *S35* to abort the male gametes from the *japonica* alleles at the *S35* heterozygous locus. Homozygous *EFS* allele from *indica* can activate the *S24* functioning in the pollen sterility, but the heterozygous *EFS* allele could eliminate the harmful effect of *S24* on the male gametes in the hybrid (Kubo et al., 2016b). Another sample was that heterozygous *HSA1a* locus from *indica* × *japonica* cross caused the female gamete abortion in the homozygous *HSA1b* allele from *indica*, Conversely, the female gamete from *japonica* at heterozygous *HSA1a* locus was normally developed with homozygous *HSA1b-j* allele (Kubo et al., 2016a). Four *S5*-interacting QTLs, *qSIG3.1*, *qSIG3.2*, *qSIG6.1*, and *qSIG12.1* were identified to interact with the *S5* locus by genetic mapping (Rao et al., 2021). These evidences suggested that a genetic network with multiple genetic factors played a pivotal role in the intraspecific hybrid sterility in rice. However, to date, the gene–gene interactions among hybrid sterility loci are not clear; one reason might be this

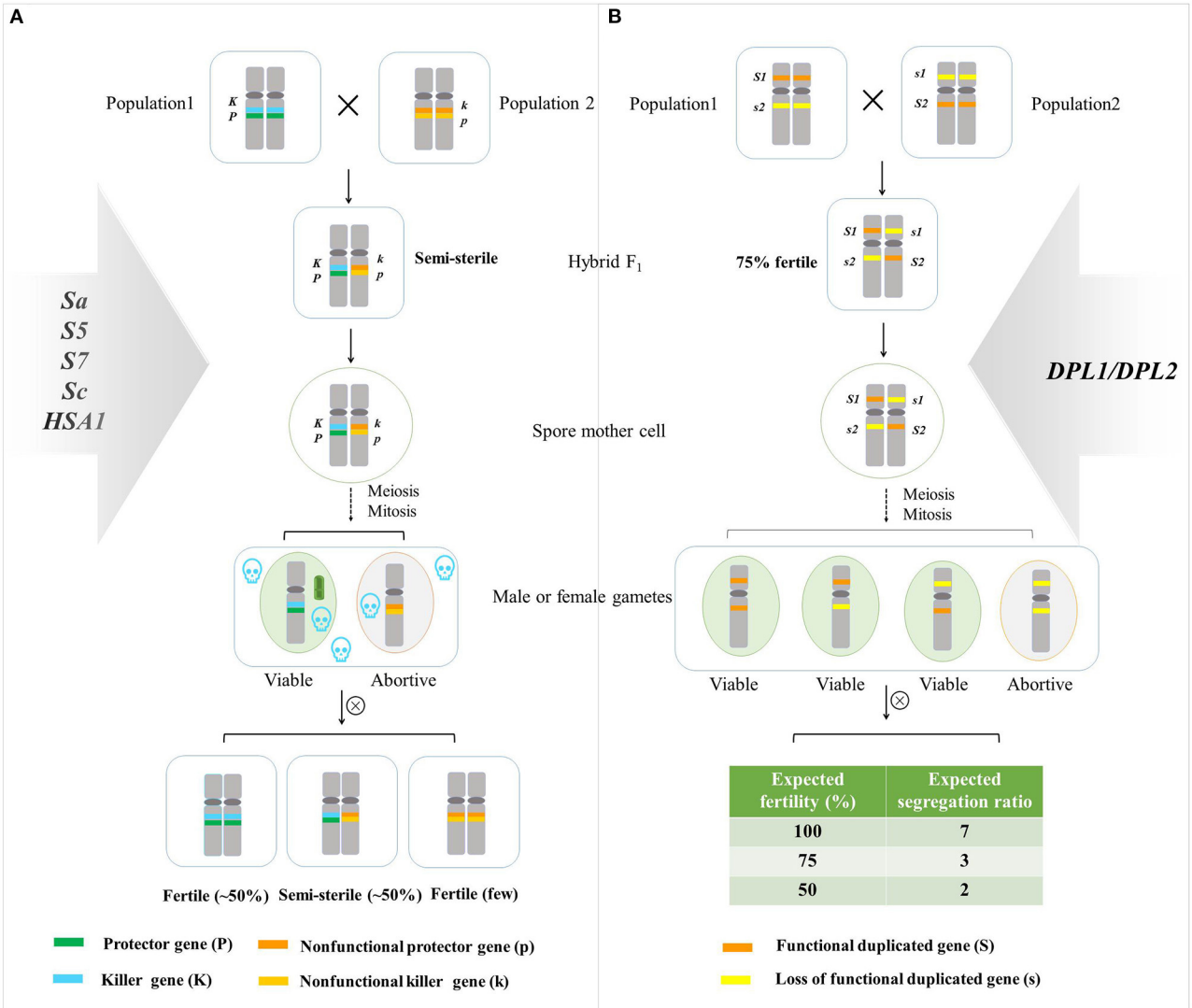

FIGURE 2 | A genetic model of hybrid sterility. **(A)** Modified one-locus sporophytic–gametophytic allelic interaction model (Ikehashi and Araki, 1986). A hybrid sterility system consisted of at least two tightly linked genes: protector gene (P) and killer gene (K) in one population, nonfunctional protector gene (p), and nonfunctional killer gene (k) in another population. *PP* and *KK* constitute a functional killer system that triggers a sterility signal in the sporophytic stage, which selectively eliminates gametes without the *PP* gene, PP could protect the gametes in a gametophytic manner, leading to semi-sterility in the hybrid. Thus, the percentage of hybrid sterility plants in F_1 hybrid accounted for 50%. The fertile plants and semi-sterile plants fit in with the ratio of 1:1 in the F_2 population. *S5*, *S7*, *Sa*, *Sc*, and *HSA1* fitted in with this model (Chen et al., 2008; Long et al., 2008; Yang et al., 2012; Kubo et al., 2016a; Yu et al., 2016; Shen et al., 2017). **(B)** Duplicate gametic lethal model interpreting two independent loci involved in the hybrid sterility. When the gametes carrying *S1S1/s2s2* or *s1s1/S2S2* developed normally, whereas two recessive alleles at both loci, *s1s1/s2s2*, conferred deleterious interaction, and about 25% of gametes were sterile (Oka, 1957, 1974). The ratio of the fertile plants, partial fertile plants, and semi-sterile plants conforms to the theoretical ratio of 7:3:2 in the progeny. *DPL1/DPL2* followed this model (Mizuta et al., 2010).

interaction depends on the special genetic background, and no effective investigation method of gene–gene interaction between the different loci might be another reason. In the future, improvements in the quantitative genetics methods coupled with a representative number of rice varieties could be needed to clarify this issue. In addition, in terms of the genetic model, there remain some substantial holes in explaining the epistasis effect on hybrid sterility, thus more comprehensive model will be established in light of the complex genetic basis of hybrid sterility.

HYBRID STERILITY LOCI AND DIVERGENCE OF SUBGROUPS IN ASIAN CULTIVATED RICE

Different Hybrid Sterility Loci Were Involved in the Divergence Between Two Subgroups

Twenty-two hybrid sterility loci had been described between *indica* and temperate *japonica* cultivars (**Figure 3**; **Supplementary Table 1**), including *Sa*, *Sc*, *Sd*, *S11*, *S33*, *S34*,

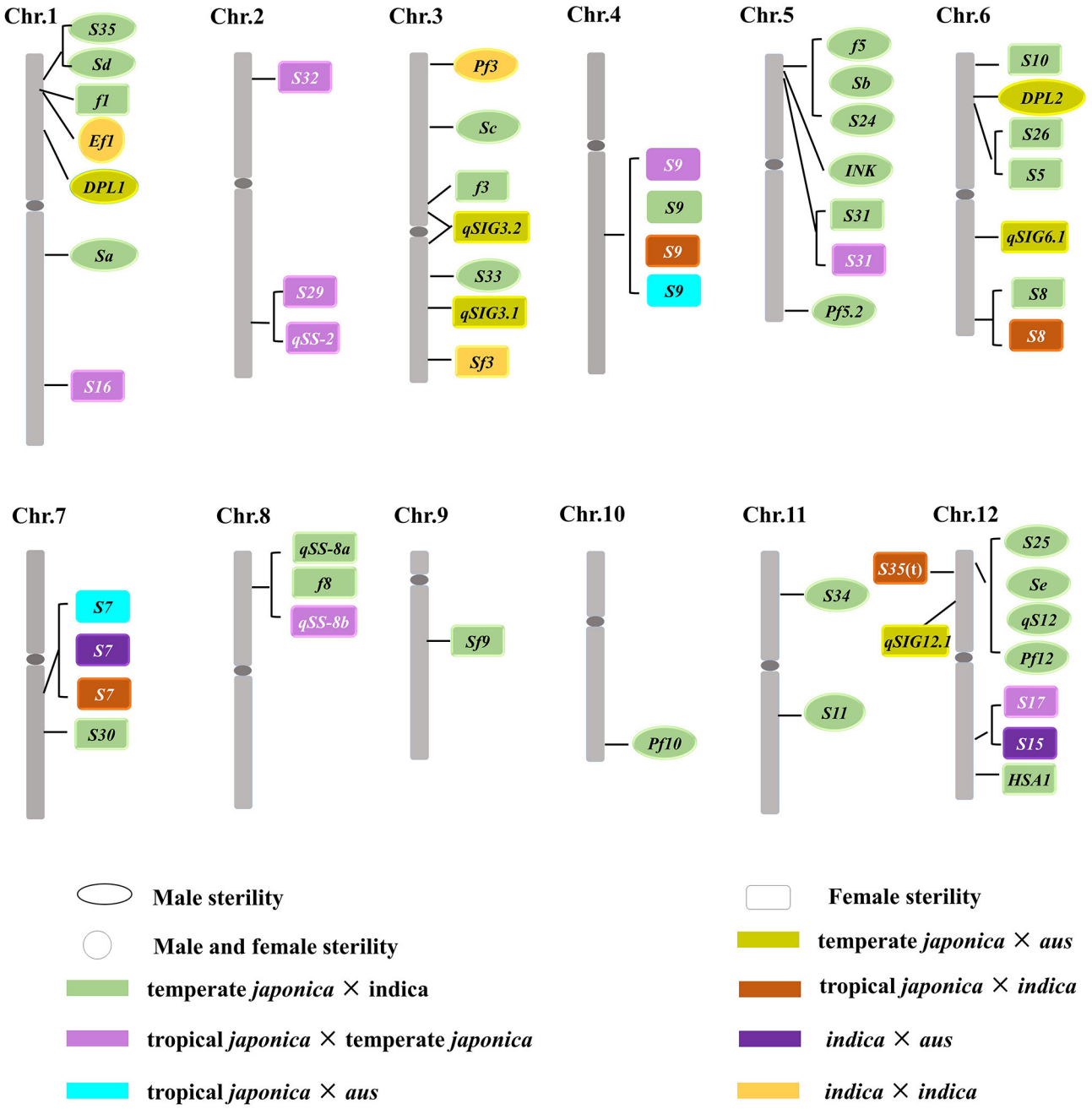

FIGURE 3 | Loci for hybrid sterility in Asian cultivated rice.

S35, *f5/Sb/S24*, *INK*, *Pf5.2*, *Pf10*, and *S25/Se/qS12/Pf12* for male gamete sterility (Sawamura and Sano, 1996; Kubo and Yoshimura, 2001; Yang et al., 2004; Li et al., 2006, 2008, 2017; Wang et al., 2006; Jing et al., 2007; Kubo et al., 2008; Long et al., 2008; Zhu et al., 2008; Zhang et al., 2011; Zhao et al., 2011; Shen et al., 2017) and *f1*, *f3*, *S5*, *S8*, *S9*, *S10*, *S30*, *S31*, *S5/S26*, *qSS-8a/f8*, *Sf9*, and *HSA1* for female gamete sterility (Sano et al., 1994; Wan et al., 1996; Wang et al., 1998, 2005; Kubo and Yoshimura, 2001; Zhu et al., 2005b; Singh et al., 2006; Zhao et al., 2006; Chen et al., 2008; Kubo et al., 2016a; Li et al., 2017). It is suggested that those loci are also involved in the divergence between temperate *japonica* and *indica* subgroups. Low-hybrid fertility was also reported among other subgroups in previous studies (Morinaga, 1968; Engle et al., 1969), and then, hybrid sterility loci were identified by genetic mapping in succession (**Figure 3**; **Supplementary Table 1**). *DPL1/DPL2* was responsible for hybrid pollen sterility (Mizuta et al., 2010), and *qSIG3.1*, *qSIG3.2*, *qSIG6.1*, as well as *qSIG12.1* gave rise

to female gamete abortion in the crosses between temperate *japonica* × *aus* (Rao et al., 2021). These results indicated that *DPL1/DPL2*, *qSIG3.1*, *qSIG3.2*, *qSIG6.1,* and *qSIG12.1* were responsible for the divergence between temperate *japonica* and *aus* subgroups. *S7* and *S15* led to female gamete sterility in the hybrid between *indica* and *aus* (Wan et al., 1996; Yu et al., 2016), and it is suggested that both loci were involved in the reproductive isolation establishment between the two subgroups. Heterozygous loci at *S9*, *S16*, *S17*, *S29*, *S31*, *S32*, *qSS-2*, and *qSS-8b* resulted in gametes abortion in the hybrid between tropical *japonica* and temperate *japonica* (Wan and Ikehashi, 1995; Wan et al., 1998; Li et al., 2005, 2007; Wang et al., 2005; Zhu et al., 2005a; Zhao et al., 2007), which inferred that those loci acted in the genetic differentiation between tropical *japonica* and temperate *japonica*. In addition, *S7*, *S8, S9*, and *S35*(t) were responsible for the female sterility in the cross between tropical *japonica* and *indica* (Wan et al., 1993, 1996; Chen et al., 2012), it also indicated that those three loci promoted the differentiation between tropical *japonica* and *indica*. *S7* and *S9* controlled hybrid sterility and contributed to the divergence between tropical *japonica* and *aus* (Yanagihara et al., 1992; Wan et al., 1996). Taken together, the different combinations of distinct hybrid sterility loci could be the major driving force for Asian cultivated rice to differentiate into five subgroups, the type, and a number of loci involved in genetic differentiation varied with the subgroup types. The results supported the classification of five subgroups in Asian cultivated rice in terms of reproductive isolation.

Allelic Differentiation at Hybrid Sterility Locus and Genetic Differentiation in Asian Cultivated Rice

In recent years, more and more evidences showed that some loci responsible for hybrid sterility were identified in several subgroups crosses simultaneously (**Supplementary Table 1**). *S7* contributed to the female sterility in the hybrid between tropical *japonica* or *indica* and *aus*, together with the tropical *japonica* and *indica* cross (Yanagihara et al., 1992; Wan et al., 1993; Yu et al., 2016), suggesting that the allelic differentiation at *S7* locus contributed to the divergence among tropical *japonica*, *indica,* and *aus* subgroups. *S8* locus conferring hybrid female sterility was reported in the different crosses between temperate *japonica* and *indica*, tropical *japonica* and *indica* (Wan et al., 1993; Singh et al., 2006), it indicated that *S8* had a potential to lead the genetic differentiation between *japonica* and *indica*. A locus *S9* for hybrid female sterility was found among the different subgroup crosses, including temperate *japonica* and tropical *japonica*, temperate *japonica* and *indica*, tropical *japonica* and *aus*, as well as tropical *japonica* and *indica* (Wan et al., 1996; Zhu et al., 2005a; Zhao et al., 2006), suggesting that allelic interaction at *S9* locus was involved in the divergence of the above subgroups. Heterozygous *S31* locus led to the hybrid female sterility in two kinds of populations, temperate *japonica* × *indica* and temperate *japonica* × tropical *japonica* (Li et al., 2005; Zhao et al., 2006), the same function had also been found in *qSS-8a/qSS-8b/f8* (Wang et al., 1998, 2005), it indicated that the genetic differentiation between temperate *japonica* and tropical *japonica* or *indica* could partially resulted from the interaction of the different haplotypes at two loci. Evidence indicated that *S15*/*S17* were mapped into the same region on chromosome 12, which might drove the divergence between *indica* and *aus*, temperate *japonica* and tropical *japonica*, respectively (Wan et al., 1996, 1998). Those six hot spots identified from different subgroup hybrids, with the same location region, as well as similar genetic behavior and phenotype, were probably orthologous genes and allelic to each other. It is suggested that these loci were conserved and contributed to hybrid sterility between different subgroups and allelic differentiation at hybrid sterility locus played important roles in triggering the divergence between distinct subgroups.

In a word, hybrid sterility is the major indication to distinguish the different subgroups in Asian cultivated rice at the subspecies level. A lot of genes or QTLs for male gamete abortion, female gamete abortion, and both male and female gamete abortions were identified in different populations. Some conserved loci governed the hybrid sterility among several subgroup crosses, whereas some loci were special for specific crosses; it suggested that multiple loci and allelic differentiation at hybrid sterility loci might be the important driving force of the divergence in Asian cultivated rice. However, the previous studies on hybrid sterility usually focused on the crosses between *japonica* and *indica*, with less progress on hybrid sterility between other subgroups. Future work should reinforce whole-genome-wide gene identification, and functional characterization of hybrid sterility loci in the hybrids between *aus* or *basmati* with other subgroups. It will help us to understand the genetic differentiation of subgroups and provide the guidelines for the heterosis strategy in rice breeding.

THE ORIGIN, THE EVOLUTION OF HYBRID STERILITY LOCI, AND THE DIVERGENCE OF ASIAN CULTIVATED RICE

Hybrid sterility resulted in the female or/and male gametes abortion in the hybrid, which hindered gene exchange and had substantial effects on the divergence between the different populations and speciation. How the hybrid sterility loci originated and evolved, fixed in the different populations, and lead to the population divergence finally was one of the hot spots in evolutionary biology.

Sa, *S5*, *Sc,* and *HSA1* Are Involved in the Differentiation of Temperate *Japonica* From *Indica*

For *Sa* locus, SaF^{+}/SaM^{+} haplotype was a major form presented in ancient Oryza species, SaF^{-}/SaM^{+} and SaF^{-}/SaM^{-} haplotypes were generated by the sequential mutations in SaF^{+} and SaM^{+} in different accessions of *O. rufipogon*. And then, the increasing frequency of SaF^{+}/SaM^{+} and SaF^{-}/SaM^{+} fixed in *indica* populations, whereas the fixation of $SaM^{-}SaF^{-}$ in *japonica* rice could be resulted from random genetic drift, together with geographical isolation or ecological adaptation (Long et al., 2008). In the light of SaF^{+}, SaM^{+}, and *SaM* consisted

of killer system, it is supposed that the hybrid fertility should be observed in the hybrid between *indica* and *O. rufipogon*, *indica* and *japonica*, or within *indica* subspecies, and then it was confirmed that *japonica* haplotype at the *Sa* locus contributed to the pollen abortion in the crosses between temperate *japonica* and *indica* or wild rice species (Long et al., 2008). Accordingly, the reproductive isolation mediated by the *Sa* locus contributed to the genetic differentiation between temperate *japonica* and *indica* subspecies.

By a reciprocal BLASTN search using three genes sequence, it was postulated that the *S5* complex could originated from Ospara3-5 by genome fragment duplication after Oryzeae tribe differentiation and produced a nascent loci *ORF3+ORF4+ORF5+* with new introns and start codons. Most of the wild relatives in rice carried the functional *ORF3+ORF4+ORF5+* allele. Mutations may occur in the partner gene by 11-bp deletion and killer gene by two SNP substitutions, leading to *ORF3+ORF4–ORF5+* (the typical *indica* genotype), and *ORF3+ORF4+ORF5–* or *ORF3+ORF4+ORF5n*, respectively. And then, the protector *ORF3+* mutated to *ORF3–* by 13bp deletion in the haplotypes of *ORF3+ORF4+ORF5–*, thus resulting in the typical *japonica* genotype of *ORF3–ORF4+ORF5–* in the pre-differentiated population. Subsequently, the increase of *indica* haplotype frequency was driven by natural selection, whereas the spread of *japonica* haplotype might be generated by the bottleneck during domestication. Once the frequencies of *ORF3+ORF4–ORF5+* and *ORF3–ORF4+ORF5–* are fixed in the corresponding populations, *indica* and *japonica* subspecies would generate deleterious interactions upon hybridization, eventually resulting in reproductive isolation mediated by *S5* locus (Du et al., 2011; Mi et al., 2020). In previous reports, 72% accession of *O. rufipogon* carried the *ORF3+ORF4+ORF5+* or neutral haplotype at the *S5* locus (Du et al., 2011; Mi et al., 2020) and this haplotype was compatible with the *japonica* or *indica* haplotypes. It suggested that it might be one of the reasons why hybrid sterility was universally observed in *indica* and *japonica* crosses, but almost never found in the crosses between *O. rufipogon* and *O. sativa.*

For the *HSA1* locus, the ancestral population of *O. rufipogon* possessed the fertile *HSA1a-j* and the sterile *HSA1b-i*^{s} alleles. Mutation in *HSA1a* of the *indica* ancestry and in *HSA1b* of the *japonica* ancestry occurred separately, which generated *indica* haplotype (*HSA1a-i*^{s}*/HSA1b-i*^{s}) and temperate *japonica* haplotype (*HSA1a-j/HSA1b-j*). The differences in adaptive and domesticated evolution at the *HSA1* locus accelerated the divergence between temperate *japonica* and *indica* subspecies (Kubo et al., 2016a). *Sc-j* and *Sc-i* haplotypes were found in the different *O. rufipogon* populations, suggesting that the origin of *Sc* occurred before the divergence between *indica* and *japonica* (Shen et al., 2017).

Therefore, hybrid sterility loci were diverged from the ancient haplotypes by a series of nucleotide variations and were further fixed in the *japonica* and *indica* subgroups, respectively, incompatible allelic interaction mediated by *Sa*, *S5*, *HSA1,* and *Sc*, activated the reproductive isolation system simultaneously, evenly restricting gene flow and resulting in the divergence of two subgroups.

DPL1/DPL2 Contributes to the Divergence Between Temperate *Japonica* and *Aus*

DPL originated from a small-scale genome duplication gene after the differentiation between *Oryza* and *Brachypodium*. Ancient *DPL* gene seemed to be functional by sequence analysis, but loss-of-function mutations of *DPL1* genes fixed in *indica* and its wild ancestor, *O. rufipogon*. Especially, all the examined accessions of *O. barthii* and *O. glaberrima* had a deletion in the *DPL1* exon, which suggested that the ancient *DPL1* allele experienced the process of many times genome deletion, nucleotide substitution, and deletion, and gradually established non-functional haplotype in the different population, and the *DPL2* gene defect was only found in the *japonica* cultivars (Mizuta et al., 2010). Based on the above data, hybrid pollen incompatibility governed by *DPL* had the potential to lead to reproductive isolation between *japonica* subspecies and *O. barthii* or *O. glaberrima*, but this phenomenon controlled by *DPL* has not been reported in interspecific hybridization population, there might be the function of *DPL* loci inhibited by an unknown factor during the stage of pollen germination. Furthermore, the origin and geographic distribution of *DPL* were investigated using 132 *O. sativa*, 296 accessions of *O. nivara,* and 388 accessions of *O. rufipogon. DPL1*$^{-}$ and *DPL2*$^{-}$ haplotypes of *O. sativa* separately emerged from an *O. nivara* population in India and *O. rufipogon* in South China, subsequently spread into other groups, especially *aus* and temperate *japonica*, respectively (Mizuta et al., 2010; Xu et al., 2021). Therefore, hybrid incompatibility established by *DPL* contributed to the genetic differentiation between *aus* and temperate *japonica* subspecies.

Taken together, the origin and evolution of *Sa* could be explained by the sequential divergence model, the mode of *S5* fitted well with the parallel-sequential divergence model, parallel divergence model could also be supported by the studies of *HSA1* and *DPL1/DPL2* (Ouyang and Zhang, 2018). Accordingly, these results supported the "multiple origins" hypothesis, in which the Asian cultivated rice has originated from the different wild relative groups. In the future, more and more cloned hybrid sterility loci would help understand the evolutional processes and history of Asian cultivated rice and find out the reasons for subspecies divergence governed by multiple hybrid sterility loci with complex mechanisms step by step. Moreover, the different haplotypes at each locus should be developed on the uniformed background as NILs to validate whether the haplotype combinations functioned in hybrid sterility or not, it will provide experimental evidence for the time and mechanism of reproductive isolation establishment.

STRATEGIES FOR OVERCOMING HYBRID STERILITY IN AN INTRASPECIFIC HYBRID BREEDING PROGRAM

Genetic diversity representing genetic differences between parental lines is fundamentally important to utilize heterosis in rice. Current hybrid rice varieties mainly utilize the heterosis from intra-subspecific crosses, mostly between *indica* parental

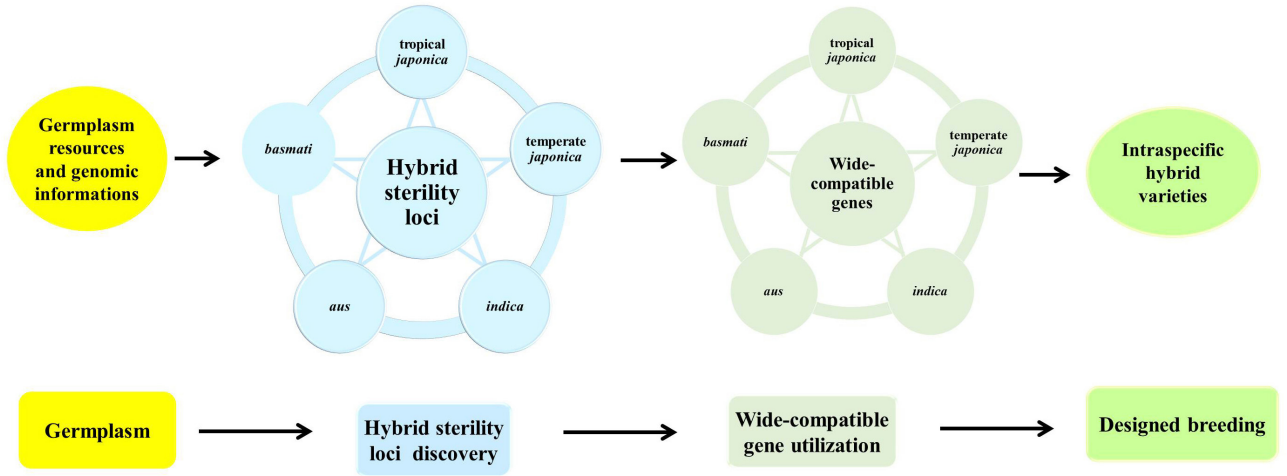

FIGURE 4 | Strategies for developing intraspecific hybrid rice. The use of diverse rice accessions in five subgroups of Asian cultivated rice representing the abundant genetic diversity, together with plenty of genome information, the loci for intraspecific hybrid sterility among five subgroup crosses will be comprehensively and systematically explored, identified, and characterized by genome-wide genetic analysis, then natural neutral genes or wide-compatible genes will be created by CRISPR-Cas9 in light of special parental lines. Applying this information, breeders could design any intraspecific crosses using known wide-compatible genes, and will be able to breed fertile hybrid varieties with strong heterosis regardless of hybrid sterility.

lines. Hybrid yields have reached a bottleneck due to the limited genetic diversity in rice. Five subgroups in Asian cultivated rice possess abundant genetic diversity, which is the basis of strong heterosis. For example, *aus* subgroup showed drought-tolerant, early-maturing traits, and 11 of 143 heterotic loci were contributed by the introgression from *aus* varieties (Lin et al., 2020); *basmati* varieties possess excellent quality traits such as slender and long grain, fluffy and soft texture after cooking, and special aroma. Therefore, *basmati* is designated as the most high priced group (Civan et al., 2019). However, high genetic diversity between subgroups might reinforce sterility barriers to maintain the independent gene pools. How to break hybrid sterility and breed intraspecific hybrids with strong heterosis has always been explored by scientists for many years. So far, three strategies have been developed to overcome hybrid sterility and enable breeders to utilize the strong hybrid vigor of *O. sativa*. The first strategy was to transfer of neutral alleles of hybrid sterility loci into the target parents to raise wide-compatibility lines for hybrid breeding (Kitamura, 1962). Neutral allele introgression, including *S5-n*, *f5*, and *Sa-n*, have been proved to enhance the pollen or spikelet fertility in the hybrid (Mi et al., 2016, 2019; Xie et al., 2017a; Ma et al., 2020). The second strategy is to develop *indica*-compatible *japonica* lines by introgression and pyramid of multiple hybrid sterility loci from *indica* into the *japonica* genetic background by using molecular marker assistant backcrossing methods (Guo et al., 2016). The third strategy is to create an artificial wide compatible line by genome editing technology, such as CRISPR/cas9 (Xie et al., 2017a,b, 2019), which is a feasible, effective, and safe method. However, hybrid sterility is a complex trait controlled by multiple loci dependent on the genetic background. If wide compatible lines eliminating the effect of the hybrid sterile loci were generated by CRISPR/cas9, the conditions are required as follows: knowing all the loci for hybrid sterility between two parental lines, clearing killer genes in the different sterility systems. Accordingly, it is necessary to utilize abundant germplasm and genome information, develop mutual intraspecific populations with five subgroups, and explore all the hybrid sterility loci by genome-wide approach, as well as characterize the molecular mechanism of the sterility loci, so as to find out or artificially create neutral allele at each locus special for target parental lines. Based on the comprehensive and systematic data, breeders could effectively and accurately utilize the stronger intraspecific heterosis by genomic breeding (**Figure 4**).

PERSPECTIVES

Five subgroups of Asian cultivated rice have provided a natural model to investigate the hybrid sterility. The genetic basis and gene identification of hybrid sterility in Asian cultivated rice have been described by many researchers, especially in *japonica* × *indica* crosses (Ouyang and Zhang, 2013, 2018; Li et al., 2020). But it is still difficult to utilize the intraspecific heterosis in rice breeding; this issue has many reasons: first, identifying only a few sterility loci is not sufficient to reveal the whole process of hybrid sterility. Second, complex variation in hybrid sterility in the hybrids differs from one cross to another, and most of the studies focus on hybrid sterility derived from one or few combinations on one genetic background, we know little about the type, number, and genetic mode of hybrid sterility loci in other subgroups, except for *japonica* and *indica*. Scattered information about sterile barriers leads to the low efficiency of intraspecific varieties improvement. There is a long way to understand the genetic basis and regulatory network of hybrid sterility.

Comprehensive and Systematic Investigation of the Genetic Basis of Hybrid Sterility Based on Five Subgroups in Asian Cultivated Rice

To understand the genetic mechanism of hybrid sterility among the different subgroups comprehensively and systematically, the genetic population representing the typical five subgroups that crossed each other should be developed to elucidate the genetic basis of intraspecific hybrid sterility. Based on the amount of data in hybrid sterility special to parental lines, a blueprint for intraspecific hybrid sterility should be drawn in the distinct combinations derived from five subgroups, clearly marking the common or special features at gene types, gene numbers, conserved genes, special genes, major effect genes, minor effect genes, and gene–gene interaction, and so on. It would provide breeders with scientific and accurate guidelines to overcome the hybrid sterility special to target parental lines. Now, 20 varieties representing five subgroups were introgressed into a temperate *japonica* variety Dianjingyou 1 and an *indica* variety IR64 by multi-round backcrosses, respectively.

Moreover, *basmati* rice is considered a unique varietal group because of its aroma and superior grain quality (Ahuja et al., 1995; Siddiq et al., 2012; Civan et al., 2019) and it displays a different pattern of genetic diversity from other subgroups (Civan et al., 2019). Over the decades, less attention has been given to the hybrid sterility between *basmati* and other subgroups, though hybrid sterility was both observed in the hybrid between a *basmati* variety (Dom Sufid) and a *japonica* variety (Dianjingyou 1) or an *indica* variety (IR64) (data not shown). By now, none of the sterile loci was identified from the *basmati* variety, so, we should pay more attention to the hybrid sterility between the *basmati* group and other subgroups to utilize the genetic diversity and unique quality traits to breed intraspecific hybrid varieties.

Molecular Mechanism and Signal Network of Hybrid Sterility Loci

To date, only six loci involved in intraspecific hybrid sterility were cloned (Chen et al., 2008; Long et al., 2008; Mizuta et al., 2010; Yang et al., 2012; Kubo et al., 2016a; Yu et al., 2016; Shen et al., 2017). We only see a few separating points, which cannot be connected to obtain a clear genetic and regulatory pathway. Thus, it is necessary to clone more genes, elucidate the allelic and non-allelic interaction, and characterize the molecular mechanism of hybrid sterility. Recent evidence indicated the DUF1618 protein not only played an important role in hybrid sterility, but also contributed to gametophyte development (Kubo et al., 2016a; Shen et al., 2017). However, its molecular function was still unclear. It will provide a rapid and easy way to overcome hybrid sterility if we can find out the "key nodes" or "switch" in the crosstalk network of hybrid sterility.

Cloning and Characterization of Segregation Distortion Factors

Segregation distortion (SD) has substantial effects in population structure and fitness of the progenies, which would result in reproductive isolation and finally speciation. SD was observed widely in the cross populations, many loci for SD were identified by genetic analysis in rice (**Figure 3**; Kinoshita, 1991, 1993; Maekawa and Inukai, 1992; Rha et al., 1995; Xu et al., 1997; Harushima et al., 2001; Matsushita et al., 2003; Reflinur et al., 2014; Li et al., 2019; Xia and Ouyang, 2020). However, only one locus *DPL1/DPL2* has been cloned and characterized (Mizuta et al., 2010), genetic and molecular mechanism of other loci remains to be unclear. Understanding the molecular events controlling SD will shed light on the reproductive isolation and speciation, as well as provide a potential approach to keep genetic diversity in populations. The above-mentioned studies will broaden our knowledge about hybrid sterility and divergence of Asian cultivated rice, and overcome the reproductive isolation so as to serve for rice intraspecific breeding.

CONCLUSION

In this article, we reviewed the comprehensive studies on the genetic differentiation and the origin of Asian cultivated rice, the comparative analysis of hybrid sterility loci among five subgroups, as well as an illustration of the relationships between hybrid sterility loci and the differentiation of subgroups in *O. sativa*. Clearly, Asian cultivated rice possesses abundant genetic diversity, which reflects the striking differences in traits, physiology, genome structure, geographic distribution, and ecological adaptation, accompanied by human civilization. Muti-dimensional discrepancies, including hybrid sterility, drive the genetic differentiation in rice. Conserved and special sterile loci contribute to the Asian cultivated rice divergence. Based on the genetic diversity of subgroups and information on the rice genome, it will facilitate gene flow by utilizing wide compatible genes for future rice breeding. Identifying loci conferring hybrid sterility between subgroups by genome-wide analysis, and elucidating the molecular mechanism and signal network will be easy to overcome the hybrid sterility by precise gene modification of key "killer genes" or "switch gene" in the sterility system.

AUTHOR CONTRIBUTIONS

DT proposed the concept. YZ conceived and wrote the paper. JW, QP, YY, YL, JZ, JL, XD, and MW edited the manuscript. All the authors reviewed and approved the final manuscript.

FUNDING

This research was supported by the National Natural Science Foundation of China (Grant Nos. 31991221, 31660380, 31201196, and 32160489), Yunnan Fundamental Research Projects (202001AS070003, 202101AS070036, 202101AT070193, and 53000210000000013809), the Yunnan Provincial Government (YNWR-QNBJ-2018-359), and the Yunnan Seed Industrialization Laboratory Program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.908342/full#supplementary-material

REFERENCES

Ahuja, S. C., Panwar, D. V. S., Uma, A., and Gupta, K. R. (1995). *Basmati Rice: The Scented Pearl*. Hisar: Directorate of Publications CCS Haryana Agricultural University).

Bourgis, F., Guyot, R., Gherbi, H., Tailliez, E., Amabile, I., Salse, J., et al. (2008). Characterization of the major fragance gene from an aromatic japonica rice and analysis of its diversity in Asian cultivated rice. *Theor. Appl. Genet.* 117, 353–368. doi: 10.1007/s00122-008-0780-9

Bradbury, L. M., Fitzgerald, T. L., Henry, R. J., Jin, Q., and Waters, D. L. (2005). The gene for fragrance in rice. *Plant Biotechnol. J.* 3, 363–370. doi: 10.1111/j.1467-7652.2005.00131.x

Chen, J., Ding, J., Ouyang, Y., Du, H., Yang, J., Cheng, K., et al. (2008). A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of indica-japonica hybrids in rice. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11436–11441. doi: 10.1073/pnas.0804761105

Chen, M., Zhao, Z., Jiang, L., and Wan, J. (2012). A new gene controlling hybrid sterility in rice (*Oryza sativa* L.). *Euphytica* 184, 15–22. doi: 10.1007/s10681-011-0467-1

Cheng, K., Wang, X., Zhou, J., Lu, Y., Lou, J., and Huang, H. (1984). Studies on indigenous rices in Yunnan and their utilization. 2. A revised classification of Asian cultivated rice (in Chinese). *Acta Agron. Sin.* 10, 271–280.

Choi, J., Lye, Z., Groen, S., Dai, X., Rughani, P., Zaaijer, S., et al. (2020). Nanopore sequencing-based genome assembly and evolutionary genomics of circum-basmati rice. *Genome Biol.* 21:21. doi: 10.1186/s13059-020-1938-2

Civan, P., Ali, S., Batista-Navarro, R., Drosou, K., Ihejieto, C., Chakraborty, D., et al. (2019). Origin of the aromatic group of cultivated rice (*Oryza sativa* L.) traced to the Indian subcontinent. *Genome Biol. Evol.* 11, 832–843. doi: 10.1093/gbe/evz039

Civan, P., Craig, H., Cox, C. J., and Brown, T. A. (2015). Three geographically separate domestications of Asian rice. *Nat. Plants* 1:15164. doi: 10.1038/nplants.2015.164

Du, H., Ouyang, Y., Zhang, C., and Zhang, Q. (2011). Complex evolution of S5, a major reproductive barrier regulator, in the cultivated rice *Oryza sativa* and its wild relatives. *New Phytol.* 191, 275–287. doi: 10.1111/j.1469-8137.2011.03691.x

Engle, L. M., Chang, T. T., and Ramirez, D. A. (1969). The cytogenetics of sterility in F1 hybrids of indica×indica and indica×javanica varieties of rice (*Oryza sativa* L.). *Philipp. Agric.* 53, 289–307.

Garris, A. J., Tai, T. H., Coburn, J., Kresovich, S., and Mccouch, S. (2005). Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169, 1631–1638. doi: 10.1534/genetics.104.035642

Glaszmann, J. C. (1987). Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet.* 74, 21–30. doi: 10.1007/BF00290078

Gross, B. L., and Zhao, Z. (2014). Archaeological and genetic insights into the origins of domesticated rice. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6190–6197. doi: 10.1073/pnas.1308942110

Guo, J., Xu, X., Li, W., Zhu, W., Zhu, H., Liu, Z., et al. (2016). Overcoming inter-subspecific hybrid sterility in rice by developing indica-compatible japonica lines. *Sci. Rep.* 6:26878. doi: 10.1038/srep26878

Gutaker, R. M., Groen, S. C., Bellis, E. S., Choi, J. Y., Pires, I. S., Bocinsky, R. K., et al. (2020). Genomic history and ecology of the geographic spread of rice. *Nat. Plants* 6, 492–502. doi: 10.1038/s41477-020-0659-6

Harushima, Y., Nakagahra, M., Yano, M., Sasaki, T., and Kurata, N. (2001). A genome-wide survey of reproductive barriers in an intraspecific hybrid. *Genetics* 159, 883–892. doi: 10.1093/genetics/159.2.883

Huang, X., Kurata, N., Wei, X., Wang, Z. X., Wang, A., Zhao, Q., et al. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490, 497–501. doi: 10.1038/nature11532

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., et al. (2011). 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

Ikehashi, H., and Araki, H. (1986). "Genetics of F1 sterility in remote crosses in rice," in *Rice Genetics Proceedings of the First Rice Genetics Symposium* (Manila: IRRI), 119–130. doi: 10.1142/9789812814265_0011

Jing, W., Zhang, W., Jiang, L., Chen, L., Zhai, H., and Wan, J. (2007). Two novel loci for pollen sterility in hybrids between the weedy strain Ludao and the japonica variety Akihikari of rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 114, 915–925. doi: 10.1007/s00122-006-0489-6

Kato, S. (1930). On the affinity of the cultivated varieties of rice plants, *Oryza sativa* L. *J. Depart. Agric. Kyushu Imper. Univ.* 2, 241–276. doi: 10.5109/22556

Kato, S., Kosaka, H., and Hara, S. (1928). On the affinity of rice varieties as shown by the fertility of hybrid plants. *Bull. Sci. Fac. Agric. Kyushu Univ.* 3, 132–147.

Khush, G. S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant Mol. Biol.* 35, 25–34. doi: 10.1023/A:1005810616885

Khush, G. S., Brar, D. S., Virk, P. S., Tang, S., Malik, S. S., Busato, G. A., et al. (2013). *Classifying Rice Germplasm by Isozyme Polymorphism and Origin of Cultivated Rice*. IRRI discussion paper series No 46. Los Banos: International Rice research Institute, 11.

Kinoshita, T. (1991). Report of the committee on gene symbolization, nomenclature and linkage group. *Rice Genet. Newslett.* 8, 2–37.

Kinoshita, T. (1993). Report of the committee on gene symbolization, nomenclature and linkage group. *Rice Genet. Newslett.* 10, 7–39.

Kishor, D. S., Seo, J., Chin, J. H., and Koh, H. J. (2020). Evaluation of whole-genome Sequence, genetic diversity, and agronomic traits of Basmati rice (*Oryza sativa* L.). *Front. Genet.* 11:86. doi: 10.3389/fgene.2020.00086

Kitamura, E. (1962). Genetics studies on sterility observed in hybrids between distantly related varieties of rice (*Oryza sativa* L.). *Bull. Chgoku Agric. Exp. Sta. Ser. A* 8, 141–205.

Kovach, M. J., Calingacion, M. N., Fitzgerald, M. A., and Mccouch, S. R. (2009). The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proc. Natl. Acad. Sci. U.S.A.* 106, 14444–14449. doi: 10.1073/pnas.0904077106

Kovach, M. J., Sweeney, M. T., and Mccouch, S. R. (2007). New insights into the history of rice domestication. *Trends Genet.* 23, 578–587. doi: 10.1016/j.tig.2007.08.012

Kubo, T., Eguchi, M., and Yoshimura, A. (2000). A new gene for F1 pollen sterility in Japonica/Indica cross of rice. *Rice Genet. Newslett.* 17, 63–69.

Kubo, T., Takashi, T., Ashikari, M., Yoshimura, A., and Kurata, N. (2016a). Two tightly linked genes at the hsa1 locus cause both F1 and F2 hybrid sterility in rice. *Mol. Plant* 9 221–232. doi: 10.1016/j.molp.2015.09.014

Kubo, T., Yamagata, Y., Eguchi, M., and Yoshimura, A. (2008). A novel epistatic interaction at two loci causing hybrid male sterility in an inter-subspecific cross of rice (*Oryza sativa* L.). *Genes Genet. Syst.* 83, 443–453. doi: 10.1266/ggs.83.443

Kubo, T., and Yoshimura, A. (2001). Linkage analysis of an F1 sterility gene in japonica/indica cross of rice. *Rice Genet. Newslett.* 18, 52–53.

Kubo, T., Yoshimura, A., and Kurata, N. (2016b). Pollen killer gene S35 function requires interaction with an activator that maps close to S24, another pollen killer gene in rice. *G3* 6, 1459–1468. doi: 10.1534/g3.116.027573

Li, D., Chen, L., Jiang, L., Zhu, S., Zhao, Z., Liu, S., et al. (2007). Fine mapping of S32(t), a new gene causing hybrid embryo sac sterility in a Chinese landrace rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 114, 515–524. doi: 10.1007/s00122-006-0450-8

Li, D., Zhu, S., Jiang, L., and Wan, J. (2005). Mapping for a new locus causing hybrid sterility in a China landrace (*Oryza sativa* L.). *Rice Genet. Newslett.* 22:20.

Li, G., Jin, J., Zhou, Y., Bai, X., Mao, D., Tan, C., et al. (2019). Genome-wide dissection of segregation distortion using multiple inter-subspecific crosses in rice. *Sci. China Life Sci.* 62, 507–516. doi: 10.1007/s11427-018-9452-8

Li, G., Li, X., Wang, Y., Mi, J., Xing, F., Zhang, D., et al. (2017). Three representative inter and intra-subspecific crosses reveal the genetic architecture of reproductive isolation in rice. *Plant J.* 92, 349–362. doi: 10.1111/tpj.13661

Li, L., Zhou, J., Zhang, Y., Yang, Y., Pu, Q., and Tao, D. (2020). New insights into the nature of interspecific hybrid sterility in rice. *Front. Plant Sci.* 11:555572. doi: 10.3389/fpls.2020.555572

Li, W., Zeng, R., Zhang, Z., Ding, X., and Zhang, G. (2006). Fine mapping of locus S-b for F1 pollen sterility in rice (*Oryza sativa* L.). *Chinese Sci. Bull.* 51, 675–680. doi: 10.1007/s11434-006-0675-6

Li, W., Zeng, R., Zhang, Z., Ding, X., and Zhang, G. (2008). Identification and fine mapping of S-d, a new locus conferring the partial pollen sterility of intersubspecific F1 hybrids in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 116, 915–922. doi: 10.1007/s00122-008-0723-5

Lin, Z., Qin, P., Zhang, X., Fu, C., Deng, H., Fu, X., et al. (2020). Divergent selection and genetic introgression shape the genome landscape of heterosis in hybrid rice. *Proc. Natl. Acad. Sci. U.S.A.* 117, 4623–4631. doi: 10.1073/pnas.1919086117

Liu, H., Xu, C., and Zhang, Q. (2004). Male and female gamete abortions, and reduced affinity between the uniting gametes as the causes for

sterility in an indica/japonica hybrid in rice. *Sex. Plant Reprod.* 17, 55–62. doi: 10.1007/s00497-004-0214-z

Long, Y., Zhao, L., Niu, B., Su, J., Wu, H., Chen, Y., et al. (2008). Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18871–18876. doi: 10.1073/pnas.0810108105

Ma, X., Wei, X., Gao, G., and Jiang, H. (2020). Development and evaluation of improved lines based on an elite rice variety 9311 for overcoming hybrid sterility in rice. *Mol. Breed.* 40:102. doi: 10.1007/s11032-020-01180-2

Maekawa, M., and Inukai, T. (1992). Genes linked with d-2 in rice. *Jpn. J. Breed. Sci.* 42(Suppl 2), 212–213. doi: 10.1270/jsbbs1951.42.213

Matsuo, T. (1952). Genecological studies on cultivated rice (in Japanese). *Bull. Nat. Inst. Agric. Sci. Jpn. D* 3, 1–111.

Matsushita, S., Iseki, T., Fukuta, Y., Araki, E., Kobayashi, S., Osaki, M., et al. (2003). Characterization of segregation distortion on chromosome 3 induced in wide hybridization between indica and japonica type rice varieties. *Euphytica* 134, 27–32. doi: 10.1023/A:1026182312730

McDermott, S. R., and Noor, M. A. (2010). The role of meiotic drive in hybrid male sterility. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 1265–1272. doi: 10.1098/rstb.2009.0264

McNally, K. L., Childs, K. L., Bohnert, R., Davidson, R. M., Zhao, K., Ulat, V. J., et al. (2009). Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12273–12278. doi: 10.1073/pnas.0900992106

Mi, J., Li, G., Huang, J., Yu, H., Zhou, F., Zhang, Q., et al. (2016). Stacking S5-n and f5-n to overcome sterility in indica-japonica hybrid rice. *Theor. Appl. Genet.* 129, 563–575. doi: 10.1007/s00122-015-2648-0

Mi, J., Li, G., Xu, C., Yang, J., Yu, H., Wang, G., et al. (2020). Artificial selection in domestication and breeding prevents speciation in rice. *Mol. Plant* 13, 650–657. doi: 10.1016/j.molp.2020.01.005

Mi, J., Y., Kim, S. R., Prahalada, G. D., Ouyang, Y., and Mou, T. (2019). An effective strategy for fertility improvement of indica-japonica hybrid rice by pyramiding S5-n, f5-n, and pf12-j. *Mol. Breed.* 39:138. doi: 10.1007/s11032-019-1044-x

Mizuta, Y., Harushima, Y., and Kurata, N. (2010). Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20417–20422. doi: 10.1073/pnas.10031 24107

Morinaga, T. (1954). "Classification of rice varieties on the basis of affinity," in *International Rice Commission, Working Party on Rice Breeding, Rep 5th Meeting,* 1–14.

Morinaga, T. (1968). Origin and geographical distribution of Japanese rice. *Jpn. Agric. Res. Q.* 1–5.

Morinaga, T., and Kuriyama, H. (1955). Japonica type rice in the subcontinent of India and Java. *Jpn. J. Breed.* 5, 149–153. doi: 10.1270/jsbbs1951.5.149

Morinaga, T., and Kuriyama, H. (1958). Intermediate type of rice in the subcontinent of India and Java. *Jpn. J. Breed.* 7, 253–269. doi: 10.1270/jsbbs1951.7.253

Myint, K. M., Courtois, B., Risterucci, A. M., Frouin, J., Soe, K., Thet, K. M., et al. (2012). Specific patterns of genetic diversity among aromatic rice varieties in Myanmar. *Rice* 5:20. doi: 10.1186/1939-8433-5-20

Oka, H. I. (1957). Genic analysis for the sterility of hybrids between distantly related varieties of cultivated rice. *J. Genet.* 55, 397–409. doi: 10.1007/BF02984059

Oka, H. I. (1974). Analysis of genes controlling F1 sterility in rice by the use of isogenic lines. *Jpn. J. Genet.* 77, 521–534. doi: 10.1093/genetics/77.3.521

Oka, H. I. (1988). *Origin of Cultivated Rice.* Tokyo: Scientific Societies Press, 181–209.

Oka, H. I., and Doida, Y. (1962). Phylogenetic differentiation of cultivated rice, XX. Analysis of the genetic basis of hybrid breakdown in rice. *Jpn. J. Genet.* 37, 24–35. doi: 10.1266/jjg.37.24

Ouyang, Y., Liu, Y. G., and Zhang, Q. (2010). Hybrid sterility in plant: stories from rice. *Curr. Opin. Plant Biol.* 13, 186–192. doi: 10.1016/j.pbi.2010.01.002

Ouyang, Y., and Zhang, Q. (2013). Understanding reproductive isolation based on the rice model. *Annu. Rev. Plant Biol.* 64, 111–135. doi: 10.1146/annurev-arplant-050312-120205

Ouyang, Y., and Zhang, Q. (2018). The molecular and evolutionary basis of reproductive isolation in plants. *J. Genet. Genomics* 45, 613–620. doi: 10.1016/j.jgg.2018.10.004

Paterson, H. E. H. (1985). "The recognition concept of species," in *Species and Speciation,* ed E. Vrba (Pretoria: Transvaal Museum), 21–29.

Rao, J., Wang, X., Cai, Z., Fan, Y., and Yang, J. (2021). Genetic analysis of s5-interacting genes regulating hybrid sterility in rice. *Rice* 14:11. doi: 10.1186/s12284-020-00452-x

Reflinur, K., B., Jang, S. M., Chu, S. H., Bordiya, Y., Akter, M. B., et al. (2014). Analysis of segregation distortion and its relationship to hybrid barriers in rice. *Rice* 7:3. doi: 10.1186/s12284-014-0003-8

Rha, E. S., Wu, S. T., and Thseng, F. S. (1995). Distorted segregation of marker genes in wide crosses of rice (*Oryza sativa* L.). *J. Agric. Ass. China* 170, 104–123.

Sano, Y., Sano, R., Eiguchi, M., and Hirano, H. Y. (1994). Gamete eliminator adjacent to the wx locus as recealed by pollen analysis in rice. *J. Hered.* 85, 310–312. doi: 10.1093/oxfordjournals.jhered.a111465

Sawamura, N., and Sano, Y. (1996). Chromosomal location of gamete eliminator, S11(t), found in an Indica-Japonica hybrid. *Rice Genet. Newslett.* 13, 70–71.

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., et al. (2014). Genomics and the origin of species. *Nat. Rev. Genet.* 15, 176–192. doi: 10.1038/nrg3644

Shen, R., Wang, L., Liu, X., Wu, J., Jin, W., Zhao, X., et al. (2017). Genomic structural variation-mediated allelic suppression causes hybrid male sterility in rice. *Nat. Commun.* 8:1310. doi: 10.1038/s41467-017-01400-y

Siddiq, E. A., Vemireddy, L. R., and Nagaraju, J. (2012). Basmati rices: genetics, breeding and trade. *Agric. Res* 1, 25–36. doi: 10.1007/s40003-011-0011-5

Singh, S. P., Sundaram, R. M., Biradar, S. K., Ahmed, M. I., Viraktamath, B. C., and Siddiq, E. A. (2006). Identification of simple sequence repeat markers for utilizing wide-compatibility genes in inter-subspecific hybrids in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 113, 509–517. doi: 10.1007/s00122-006-0316-0

Stein, J. C., Yu, Y., Copetti, D., Zwickl, D. J., Zhang, L., Zhang, C., et al. (2018). Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. *Nat. Genet.* 50, 285–296. doi: 10.1038/s41588-018-0040-0

Teng, J., Xue, Q., and Wang, Y. (1996). Ultrastructural observations of pollen and anther wall developments between subspecies in rice (*Oryza sativa* L.) (in Chinese). *J. Zhejiang Agric. Univ.* 22, 467–473.

Ting, Y. (1949a). Origination of the rice cultivation in China. *J. Coll. Agric. Sun Yat-Sen Univ.* 7, 11–24.

Ting, Y. (1949b). Chronological studies of the cultivation and the distribution of rice varieties, Keng and Sen (in Chinese). *Agr. Bull. Coll. Agric. Sun Yatsen Univ.* 6, 1–32.

Ting, Y. (1949c). A preliminary report on the cultivation and distribution of hsien and keng rices in ancient China and the classification of current cultivars (in Chinese). *Mem. Coll. Agric. Sun Yatsen Univ.* 6, 1–32.

Travis, A. J., Norton, G. J., Datta, S., Sarma, R., Dasgupta, T., Savio, F. L., et al. (2015). Assessing the genetic diversity of rice originating from Bangladesh, Assam and West Bengal. *Rice* 8:35. doi: 10.1186/s12284-015-0068-z

Vaughan, D. A. (1994). *The Wild Relatives of Rice: A Genetic Resources Handbook.* Manila: The International Rice Research Institute.

Wan, J., and Ikehashi, H. (1995). Identification of a new locus S-16 causing hybrid sterility in native rice varieties (*Oryza sativa* L.) from Tai-hu lake region and Yunnan Province, China. *Breed. Sci.* 45, 461–470. doi: 10.1270/jsbbs1951.45.461

Wan, J., Ikehashi, H., Sakai, M., Horisuei, H., and Imbe, T. (1998). Mapping of hybrid sterility gene S17 of rice (*Oryza sativa* L.) by isozyme and RFLP markers. *Rice Genet. Newslett.* 15, 151–154.

Wan, J., Yamaguchi, Y., Kato, H., and Ikehashi, H. (1996). Two new loci for hybrid sterility in cultivated rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 92, 183–190. doi: 10.1007/BF00223375

Wan, J., Yanagihara, S., Kato, H., and Ikehashi, H. (1993). Multiple alleles at a new locus causing hybrid sterility between a Korean indica variety and a javanica variety in rice (*Oryza sativa* L.). *Jpn. J. Breed.* 43, 507–516. doi: 10.1270/jsbbs1951.43.507

Wang, C., Zhu, C., Zhai, H., and Wan, J. (2005). Mapping segregation distortion loci and quantitative trait loci for spikelet sterility in rice (*Oryza sativa* L.). *Genet. Res.* 86, 97–106. doi: 10.1017/S0016672305007779

Wang, G., He, Y., Xu, C., and Zhang, Q. (2006). Fine mapping of f5-Du, a gene conferring wide-compatibility for pollen fertility in inter-subspecific hybrids of rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 112, 382–387. doi: 10.1007/s00122-005-0141-x

Wang, J., Liu, K., Xu, C., Li, X., and Zhang, Q. (1998). The high level of wide-compatibility of variety 'Dular' has a complex genetic basis. *Theor. Appl. Genet.* 97, 407–412. doi: 10.1007/s001220050910

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, 43–49. doi: 10.1038/s41586-018-0063-9

Xia, F., and Ouyang, Y. (2020). Recurrent breakdown and rebalance of segregation distortion in the genomes: battle for the transmission advantage. *aBIOTECH*. 1, 246–254. doi: 10.1007/s42994-020-00023-0

Xie, Y., Niu, B., Long, Y., Li, G., Tang, J., Zhang, Y., et al. (2017a). Suppression or knockout of SaF/SaM overcomes the Sa-mediated hybrid male sterility in rice. *J. Integr. Plant Biol.* 59, 669–679. doi: 10.1111/jipb.12564

Xie, Y., Tang, J., Xie, X., Li, X., Huang, J., Fei, Y., et al. (2019). An asymmetric allelic interaction drives allele transmission bias in interspecific rice hybrids. *Nat. Commun.* 10:2501. doi: 10.1038/s41467-019-10488-3

Xie, Y., Xu, P., Huang, J., Ma, S., Xie, X., Tao, D., et al. (2017b). Interspecific hybrid sterility in rice is mediated by OgTPR1 at the S1 locus encoding a peptidase-like protein. *Mol. Plant.* 10, 1137–1140. doi: 10.1016/j.molp.2017.05.005

Xu, X., Ge, S., and Zhang, F. (2021). Genetic and geographic patterns of duplicate DPL genes causing genetic incompatibility within rice: implications for multiple domestication events in rice. *Rice Sci.* 28, 58–68. doi: 10.1016/j.rsci.2020.11.007

Xu, Y., Zhu, L., Xiao, J., Huang, N., and Mccouch, S. R. (1997). Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* 253, 535–545. doi: 10.1007/s004380050355

Yan, J., Zou, D., Li, C., Zhang, Z., Song, S., and Wang, X. (2020). SR4R: an integrative SNP resource for genomic breeding and population research in rice. *Genomics Proteomics Bioinformatics* 18, 173–185. doi: 10.1016/j.gpb.2020.03.002

Yanagihara, S., Kato, H., and Ikehashi, H. (1992). A new locus for multiple alleles causing hybrid sterility between an Aus variety and javanica varieties in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 90, 182–188. doi: 10.1007/BF00222200

Yang, C., Chen, Z., Zhuang, C., Mei, M., and Liu, Y. (2004). Genetic and physical fine-mapping of the Sc locus conferring indica-japonica hybrid sterility in rice (*Oryza sativa* L.). *Chinese Sci. Bull.* 49, 1718–1721. doi: 10.1360/04wc0197

Yang, J., Zhao, X., Cheng, K., Du, H., Ouyang, Y., Chen, J., et al. (2012). A killer-protector system regulates both hybrid sterility and segregation distortion in rice. *Science* 337, 1336–1340. doi: 10.1126/science.1223702

Yu, Y., Zhao, Z., Shi, Y., Tian, H., Liu, L., Bian, X., et al. (2016). Hybrid sterility in rice (*Oryza sativa* L.) involves the tetratricopeptide repeat domain containing protein. *Genetics* 203, 1439–1451. doi: 10.1534/genetics.115.183848

Zhang, G., and Lu, Y. (1993). Genetic studies of the hybrid sterility in cultivated rice (*Oryza sativa*). II. Gene mode of F1 pollen sterility (in Chinese). *Acta Genet. Sin.* 20, 222–228.

Zhang, G., Lu, Y., Liu, G., Yang, J., and Zhang, H. (1989). Genetic studies of the hybrid sterility in culti vated rice (*Oryza sativa*). I. Diallel analysis of the hybrid sterility among isogenic F1 sterile lines (in Chinese). *Chinese J. Rice Sci.* 3, 97–101.

Zhang, G., Lu, Y., Liu, G., Yang, J., and Zhang, H. (1993). Genetic studies of the hybrid sterility in cultivated rice (*Oryza sativa*). III. Allele differentiation of F1 pollen sterility in different types of varieties (in Chinese). *Acta Genet. Sin.* 20, 541–551.

Zhang, G., Lu, Y., Zhang, H., Yang, J., and Liu, G. (1994). Genetic studies of hybrid sterility in cultivated rice (*Oryza sativa*), IV. Genotypes of pollen sterility genes (In Chinese). *Acta Genet. Sin.* 21, 31–41.

Zhang, H., Zhang, C., Sun, Z., Yu, W., Gu, M., Liu, Q., et al. (2011). A major locus qS12, located in a duplicated segment of chromosome 12, causes spikelet sterility in an indica-japonica rice hybrid. *Theor. Appl. Genet.* 123, 1247–1256. doi: 10.1007/s00122-011-1663-z

Zhao, Z. (2011). New archaeobotanic data for the study of the origins of Agriculture in China. *Curr Anthropol S*295–S306. doi: 10.1086/659308

Zhao, Z., Jiang, L., Zhang, W., Yu, C., Zhu, S., Xie, K., et al. (2007). Fine mapping of S31, a gene responsible for hybrid embryo-sac abortion in rice (*Oryza sativa* L.). *Planta* 226, 1087–1096. doi: 10.1007/s00425-007-0553-8

Zhao, Z., Wang, C., Jiang, L., Zhu, S., Ikehashi, H., and Wan, J. (2006). Identification of a new hybrid sterility gene in rice (*Oryza sativa* L.). *Euphytica* 151, 331–337. doi: 10.1007/s10681-006-9154-z

Zhao, Z., Zhu, S., Zhang, Y., Bian, X., Wang, Y., Jiang, L., et al. (2011). Molecular analysis of an additional case of hybrid sterility in rice (*Oryza sativa* L.). *Planta* 233, 485–494. doi: 10.1007/s00425-010-1313-8

Zhu, S., Jiang, L., and Wang, C. M. (2005b). The origin of weedy rice Ludao in China deduced by a genome wide analysis of its hybrid sterility genes. *Breed. Sci.* 55, 409–414. doi: 10.1270/jsbbs.55.409

Zhu, S., Wang, C., Zheng, T., Zhao, Z., Ikehashi, H., and Wan, J. (2005a). A new gene located on chromosome 2 causing hybrid sterility in a remote cross of rice. *Plant Breed.* 124, 440–445. doi: 10.1111/j.1439-0523.2005.01143.x

Zhu, W., Li, W., Ding, X., Zhang, Z., Zeng, R., Zhu, H., et al. (2008). Preliminary identification of F1 pollen sterility gene S-e in *Oryza sativa* (in Chinese). *J. South China Agric. Univ.* 29, 1–5.

Zhu, X., Cao, X., and Zhu, Q. (1996). Cytological studies on spikelet sterility of indica-japonica hybrids in rice (in Chinese). *Chin. J. Rice Sci.* 10, 71–78.

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

ORIGINAL RESEARCH
published: 28 June 2022
doi: 10.3389/fpls.2022.908000

A Novel Combination of Genes Causing Temperature-Sensitive Hybrid Weakness in Rice

*Than Kutay Soe[1,2], Mai Kunieda[1], Hidehiko Sunohara[1,3], Yoshiaki Inukai[4], Vincent Pamugas Reyes[1], Shunsaku Nishiuchi[1] and Kazuyuki Doi[1]**

[1] Laboratory of Information Sciences in Agricultural Lands, Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya, Japan, [2] Department of Botany, University of Yangon, Yangon, Myanmar, [3] Environmental Control Center Co., Ltd., Hachioji, Japan, [4] International Center for Research and Education in Agriculture, Nagoya University, Nagoya, Japan

Reproductive isolation is an obstacle for plant breeding when a distant cross is demanded. It can be divided into two main types based on different growth stages: prezygotic isolation and postzygotic isolation. The hybrid weakness, which is a type of postzygotic isolation, can become a problem in crop breeding. In order to overcome reproductive isolation, it is necessary to elucidate its mechanism. In this study, genetic analysis for low temperature-dependent hybrid weakness was conducted in a rice F_2 population derived from Taichung 65 (T65, Japonica) and Lijiangxintuanheigu (LTH, Japonica). The weak and severe weak plants in F_2 showed shorter culm length, late heading, reduced panicle number, decreased grain numbers per panicle, and impaired root development in the field. Our result also showed that hybrid weakness was affected by temperature. It was observed that 24°C enhanced hybrid weakness, whereas 34°C showed recovery from hybrid weakness. In terms of the morphology of embryos, no difference was observed. Therefore, hybrid weakness affects postembryonic development and is independent of embryogenesis. The genotypes of 126 F_2 plants were determined through genotyping-by-sequencing and a linkage map consisting of 862 single nucleotide polymorphism markers was obtained. Two major quantitative trait loci (QTLs) were detected on chromosomes 1 [*hybrid weakness j 1* (*hwj1*)] and 11 [*hybrid weakness j 2* (*hwj2*)]. Further genotyping indicated that the hybrid weakness was due to an incompatible interaction between the T65 allele of *hwj1* and the LTH allele of *hwj2*. A large F_2 populations consisting of 5,722 plants were used for fine mapping of *hwj1* and *hwj2*. The two loci, *hwj1* and *hwj2*, were mapped in regions of 65-kb on chromosome 1 and 145-kb on chromosome 11, respectively. For *hwj1,* the 65-kb region contained 11 predicted genes, while in the *hwj2* region, 22 predicted genes were identified, two of which are disease resistance-related genes. The identified genes along these regions serve as preliminary information on the molecular networks associated with hybrid weakness in rice.

Keywords: reproductive isolation, hybrid weakness, genotyping-by-sequencing (GBS), rice, QTL

OPEN ACCESS

Edited by:
Kazuki Matsubara, Institute of Crop Science (NARO), Japan

Reviewed by:
Takahiro Tezuka, Osaka Metropolitan University, Japan
Jeonghwan Seo, National Institute of Crop Science, Rural Development Administration, South Korea

***Correspondence:**
Kazuyuki Doi
kdoi@agr.nagoya-u.ac.jp

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

Received: *30 March 2022*
Accepted: *26 May 2022*
Published: *28 June 2022*

Citation:
Than Kutay Soe, Kunieda M, Sunohara H, Inukai Y, Reyes VP, Nishiuchi S and Doi K (2022) A Novel Combination of Genes Causing Temperature-Sensitive Hybrid Weakness in Rice. Front. Plant Sci. 13:908000. doi: 10.3389/fpls.2022.908000

INTRODUCTION

In plant breeding, the development of new varieties requires expansion of genetic diversity, subsequent selection, and uniformization. The expansion of genetic diversity is the initial and one of the important steps to exploiting the diversity of genes available for breeding. For this purpose, several methods such as introduction, crossing, mutagenesis, and transgenic have been used. Among these methods, the crossing is the most frequently used. The genetic diversity of materials used for crossing, especially of distant relatives, often cause abnormality of the hybrids and progeny, called the reproductive barrier. To facilitate the use of various genetic resources and to transfer useful genes to new varieties, an intermediate set of materials that breeders can use as the starting materials, which is known as pre-breeding (Kumar and Shukla, 2014), is proposed.

In rice, crosses between different genomic species of the genus *Oryza* are often very difficult due to the crossing barriers and abnormal chromosome pairing in meiosis (Vaughan et al., 2003). There are several reports on the successful construction of pre-breeding materials (introgression lines) obtained from the crosses between *Oryza sativa* and other AA-genome species (e.g., Yoshimura et al., 2010). In addition, *O. sativa* has been classified into Japonica and Indica, and crosses between this variety group are also performed (e.g., Kubo et al., 2002). The reproductive isolation was observed in these cross combinations (e.g., Kubo and Yoshimura, 2002). Therefore, reproductive isolation generally becomes a major obstacle in plant breeding.

Reproductive barriers are highly related to the genetic differentiation of populations within and between species (Coyne and Orr, 2004; Rieseberg and Willis, 2007). Based on the stage of growth, these reproductive barriers can be classified as prezygotic isolation and postzygotic isolation. Prezygotic isolation is a common phenomenon that prevents both an inter and an intra-specific crossing. This type of reproductive isolation is due to various factors such as geographical isolation, the difference in flowering time, pollinator specificity, incompatibility in pollen tube growth (Stebbins, 1950). In contrast, postzygotic reproductive isolation happens after the zygotes or hybrids are developed. This type of reproductive isolation often leads to embryonic lethality, seed inviability, weakness, and sterility. One of the fundamental theories on the mechanism of postzygotic isolation was explained by the Bateson–Dobzhansky–Muller (BDM) model (Dobzhansky, 1937; Muller, 1942; Coyne and Orr, 2004). This model postulates that the deleterious interaction of two or more genes, derived from different species or populations, causes postzygotic isolation.

Hybrid weakness refers to the phenomenon in which the hybrid, derived from two normal parents, shows defective development such as necrotic leaves, small stature, or poor growth (Bomblies and Weigel, 2007). This phenomenon has been reported in many other plant species, including *Arabidopsis thaliana* (Bomblies et al., 2007), *Phaseolus vulgaris* (Shii et al., 1980), interspecific crosses among *Gossypium* (Lee, 1981), interspecific crosses among *Nicotiana* (Tezuka et al., 2007), and interspecific crosses among *Capsicum* (Shiragaki et al., 2021).

Hybrid weakness in rice is frequently observed (Koide et al., 2008). To date, several studies have identified genes that are associated with hybrid weakness in rice. The very first study of hybrid weakness genes, L_1 and L_2 were reported by Oka (1957). In most cases, hybrid weakness was controlled by the complementary interaction of unlinked loci, e.g., *Hwc1* and *Hwc2* (Ichitani et al., 2001, 2007; Kuboyama et al., 2009), *hwd1* and *hwd2* (Fukuoka et al., 1998), *hwe1* and *hwe2* (Kubo and Yoshimura, 2002), *hwg1* and *hwg2* (Fukuoka et al., 2005), *hwh1* and *hwh2* (Jiang et al., 2008), *hwi1* and *hwi2* (Chen et al., 2013, 2014), and *hbd2* and *hbd3* (Matsubara et al., 2007; Yamamoto et al., 2007).

Previous studies suggested that the over-activated immune responses might be an essential mechanism for triggering hybrid weakness of rice, *Arabidopsis,* and other plant species (Bomblies et al., 2007; Alcázar et al., 2009, 2010; Jeuken et al., 2009; Chen et al., 2013, 2014; Shiragaki et al., 2021). For example, in *Arabidopsis*, an autoimmune response was activated to cause hybrid weakness by gene interaction between two loci and at least one of them coded nucleotide-binding site leucine-rich repeat (NB-LRR) protein (Bomblies et al., 2007). NB-LRR genes are the most common class of disease resistance (R) genes in plants (Jones and Dangl, 2006). In rice, one of the interacting loci for hybrid weakness coded NB-LRR genes (Yamamoto et al., 2010). Chen et al. (2014) reported that the interaction of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes from wild rice and the subtilisin-like protease gene from Indica rice activated the autoimmune response resulted in hybrid weakness. Temperature is known to influence disease resistance to pathogens. A high temperature mostly inhibits disease resistance (e.g., Dropkin, 1969). Only a few genes for hybrid weakness in rice have been cloned (Yamamoto et al., 2010; Chen et al., 2014; Nadir et al., 2019). Therefore, the molecular mechanism of hybrid weakness in rice still needs to be elucidated.

Clarifying the mechanisms of reproductive barriers such as hybrid weakness is important not only to understand speciation and evolution but also to overcome these barriers for crop breeding. Here, we report a new combination of loci causing temperature-sensitive hybrid weakness in rice. The genetic analysis of hybrid weakness was carried out using F_2 populations derived from a cross between Taichung 65 (*Oryza sativa* Japonica) and Lijiangxintuanheigu (*O. sativa* Japonica). The two loci, *hwj1* and *hwj2*, associated with hybrid weakness, were fine mapped at the 65-kb region and 145-kb of chromosomes 1 and 11, respectively.

MATERIALS AND METHODS

Plant Materials

In this study, F_2 populations derived from a cross between Taichung 65 (T65) and Lijiangxintuanheigu (LTH) were used (**Figure 1**). In summary, the seeds were pre-germinated in water at room temperature for 3 days and were sown. The 30-day-old seedlings were transplanted ten plants per row with 20 cm between the hill and 30 cm spacing between rows in Togo field, Nagoya University. The field experiments were done three times,

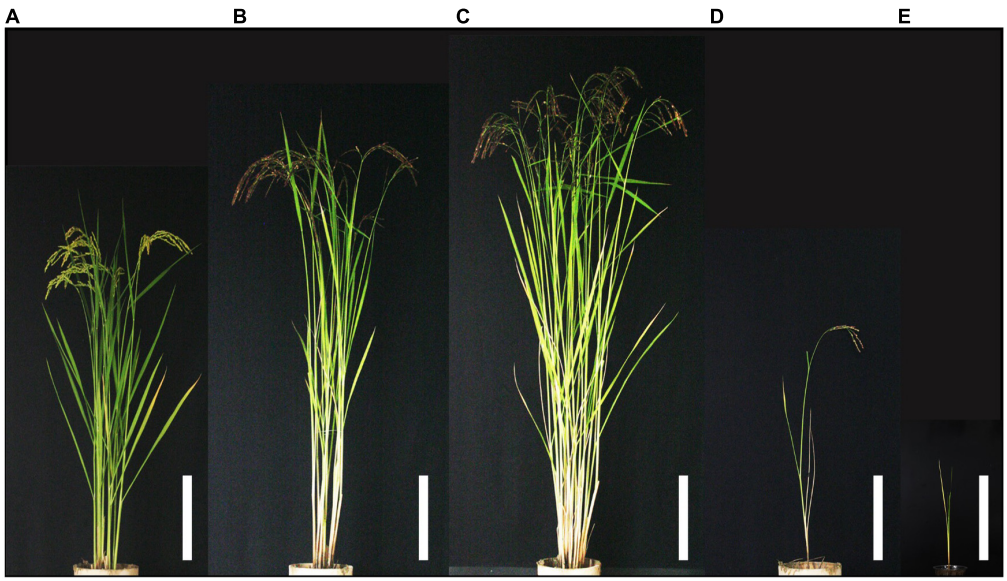

FIGURE 1 | Morphology of F_2 plants and parents at the reproductive stage. **(A)** T65, **(B)** Lijiangxintuanheigu (LTH), **(C)** normal plant, **(D)** weak plant, and **(E)** severe weak plant. Scale bars = 20 cm.

in 2016, 2018, and 2019. In 2016, an F_2 population consisting of 344 plants was grown and among them, 126 plants were used for genotyping by sequencing (GBS) and quantitative trait loci (QTL) analysis. In 2018 and 2019, large F_2 populations consisting of 5,722 plants were used for fine mapping of *hwj1* and *hwj2*.

Phenotypic Characterization of Hybrid Weakness

In 2016, culm length, days to heading, and panicle number per plant were recorded using 306 plants. In 2019, the number of grains per panicle, and filled and unfilled grain numbers in the main panicle were counted using ten mature plants for each genotype. Days to heading were recorded thrice in a week. Harvesting of the panicles of each plant was done 40–45 days after heading and samples were dried in a greenhouse. After the QTL analysis, all 344 F_2 plants (in 2016) were genotyped using the DNA markers linked to the two causal loci, and trait values were compared among the genotypes using R software[1].

Genotyping-by-Sequencing and Quantitative Trait Loci Analysis

For DNA extraction, the leaf of each plant was taken three weeks after transplanting and was oven-dried at 52°C overnight. The DNA was extracted following a modified Dellaporta method (Dellaporta et al., 1983). The quality of extracted DNA was checked by electrophoresis on a 0.8% agarose gel in 0.5 × TBE buffer. For GBS materials, the QuantiFluor dsDNA System and Quantus fluorometer instrument (Promega, Madison, WI, United States) were used for the quantification of extracted DNA.

[1]http://r-project.org/

The genotypes of 126 F_2 plants, from the F_2 population consisting of 344 individual plants derived from the cross between T65 and LTH (in 2016), were determined following the protocol of Poland et al. (2012). Two combinations of the restriction enzymes, *Kpn*I-*Msp*I and *Pst*I-*Msp*I were used for the construction of the next generation sequencing (NGS) libraries. The NGS libraries were sequenced using Illumina MiSeq (Illumina, San Diego, CA, United States). The informatics was done following the pipeline used by Furuta et al. (2017).

The number of panicles of the 126 F_2 plants was used as an indicator for QTL mapping for the population in 2016. The R/qtl package (Broman et al., 2003) was used for QTL analysis using "scanone" and "scantwo" functions for detecting 1-way and 2-way QTL. The LOD score significance threshold was calculated using 1,000 permutations.

Fine Mapping of Causal Loci

In 2018 and 2019, a total of 5,722 plants from F_2 populations were genotyped using InDel markers surrounding the two loci (**Tables 1**, **2**). The informative plants were used for further fine mapping. To narrow down the candidate regions, cleaved amplified polymorphic sequence (CAPS) markers were designed and used (**Supplementary Tables 1**, **2**).

Temperature Sensitivity Test

To analyze the sensitivity to temperature, T65, LTH, and their F_2 seeds were pre-germinated in water at room temperature for 3 days. The germinated seeds were sown on a plastic mesh in the water at 24 and 34°C in a growth chamber with continuous light. Ten days after sowing, the shoot length and root length of each plant were measured. Alongside these, the number of roots of each plant was also counted. To identify significant differences among genotypes, one-way analysis of variance (ANOVA)

TABLE 1 | List of InDel markers for fine mapping of *hwj1* locus.

No.	Marker	Primer sequence	Position
1	IDM0144_f	ccccactgttcccaaaccgt	17,405,250
	IDM0144_r	tttaacccccctcaacttttcactct	
2	IDM0145_f	agcattggtatttacgtcgctaca	18,526,363
	IDM0145_r	actttgagcctcatttgttaccca	
3	IDM0147_f	gtcaccgttggtagccccac	18,863,466
	IDM0147_r	gcagcagtggagtgacacca	
4	IDSO0101_f	tgtgacacctgtttttatcttccgta	17,724,364
	IDSO0101_r	gcatgactggagatcttcgaatactta	
5	IDMSO0102_f	tccctcgatctcacggaatta	17,797,920
	IDMSO0102_r	cagggttagacgggaacgtg	
6	IDMSO0104_f	cgggtgaaataaccgggagt	18,272,024
	IDMSO0104_r	tgcatggtttaaccgggagt	
7	IDMSO0105_f	tctgcccttacctccctgaa	18,460,793
	IDMSO0105_r	tgtgctgtgtgtcgcgtatg	

TABLE 2 | List of InDel markers for fine mapping of *hwj2* locus.

No.	Marker	Primer sequence	Position
1	IDM1104_f	tgttccagccaacgaacaca	22,425,872
	IDM1104_r	tctctacgacgcggcaaact	
2	IDM1106_f	cgtgggcaggaggtggaga	23,414,083
	IDM1106_r	tctgggccgtcagatctgct	
3	IDM1109_f	tgcaagattgtggcaagcgc	24,322,691
	IDM1109_r	agcaggcttgttggtgagtga	
4	IDMSO1101_f	actttccgatcgattagttgaca	23,488,904
	IDMSO1101_r	tccgtacgtaatctctagaaagaagaa	
5	IDMSO1102_f	tgttggtttgtgagggcagtc	23,693,281
	IDMSO1102_r	gcatctcgctgttggctgt	
6	IDMSO1103_f	cacgggggtcactggataag	23,748,995
	IDMSO1103_r	cacgccaaggtcatgcttc	
7	IDMSO1104_f	tgcctttcgagttcttcacca	24,085,843
	IDMSO1104_r	gcagctcaaagaaaaagaagagca	
8	IDMSO1105_f	gggacagatggtttttcttgtg	24,206,507
	IDMSO1105_r	tcagacaaccgttcacgaaa	
9	IDMSO1106_f	tcccattcctcatcactaactcc	24,210,608
	IDMSO1106_r	tgggttatgcatacgacagca	

followed by Tukey's HSD (honestly significant difference) test ($p < 0.05$) was performed through R software.

Morphological and Histological Characterizations of Embryo

Prior to morphological and histological analyses of seeds, genotypes were determined using half seeds of the F_2 populations. To observe the morphology and histology of the embryos, half seeds of all genotypes including parents were soaked at 24°C. The F_2 seeds which possessed severe weak genotype (*aabb*), weak genotype (*aaBb*, *Aabb)*, and normal genotype (*AABB*), together with T65 and LTH, were soaked in water at 24°C. Embryos were observed 1 day after soaking under a digital microscope (Keyence VHX-600). The embryos were then fixed with 4% paraformaldehyde overnight at 4°C and embedded in SCEM medium (SECTION-LAB, Hiroshima, Japan). The cut surface was covered with an adhesive film (Cryofilm type IIC9, SECTION-LAB, Hiroshima, Japan) and frozen sections (20 μm thickness) were performed with a cryostat (CM 1850 Leica Microsystems, Germany) according to the Kawamoto method (Kawamoto, 2003) and stained with hematoxylin and eosin. The mounting semi-permanent slides were examined under a microscope (OLYMPUS BX52 equipped with DP72 camera).

Plant Growth and Quantitative Reverse Transcription PCR

The pathogen-related (PR) genes: *PR1a*, *PR1b*, *PR2*, *PR4*, *PBZ1,* or *PR10*, were identified as genes whose mRNA expression was induced as a response to the pathogen (van Loon, 1985). In rice, Yamamoto et al. (2010) and Chen et al. (2013, 2014) reported that mRNA expression of the PR genes was increased in plants showing symptoms of hybrid weakness and those genes demonstrated an autoimmune response. To confirm the autoimmune response of hybrid weakness in our study, mRNA expression of PR genes (*PBZ1*, *PR1a*, *PR4*, and *JIOsPR10*) were determined using 7-day-old seedlings grown at 24°C. Primers used for qRT-PCR are listed in **Table 3**. In summary, the F_2 seeds which possessed severe weak genotype (*aabb*), weak genotype (*aaBb*, *Aabb)*, and normal genotype (*AaBB, AABb*), together with T65 and LTH were pre-germinated for 3 days. The germinated seeds were then sown on a mesh on the water in a growth chamber at 24°C. In this study, the electric heaters were used to control the water temperature. However, the air temperature was not controlled, but stable at approximately 22°C. Therefore, the temperature around seedlings was regarded to be at 24°C.

The seedlings, leaf blades, and sheaths, were sampled 7 days after sowing, and the tip of the leaf blades was used for DNA extraction, and the remaining parts of the leaf blades and sheaths were used for total RNA extraction using the RNeasy plant mini kit (Qiagen Sciences, MD, United States) according to the manufacturer's instructions. The RNA concentration was measured through QuantiFluor® RNA System and Quantus fluorometer instrument (Promega, WI, United States). The ReverTra Ace® qPCR RT Master Mix with gDNA Remover (TOYOBO, Osaka, Japan) was utilized to synthesize cDNA according to the manufacturer's instructions using 500 ng total RNA. qRT-PCR was performed using StepOnePlus Real-Time PCR (Life Technologies, MD, United States).

TABLE 3 | List of primer sequences for qRT-PCR.

Gene	Primer sequence	References
Actin_f	TGTATGCCAGTGGTCGTACCA	Chen et al., 2013
Actin_r	CCAGCAAGGTCGAGACGAA	
PBZ1_f	TACACCATGAAGCTTAACCCTGCC	Chen et al., 2013
PBZ1_r	TCGAGCACATCCGACTTTAGGACA	
PR1a_f	GGTTATCCTGCTGCTTGCTGGT	Chen et al., 2013
PR1a_r	GTTGTGCGGGTCCACGAAGT	
PR4_f	AGTATGGATGGACCGCCTTCTGT	Chen et al., 2013
PR4_r	CTCGCAATTATTGTCGCACCTGTTC	
JIOsPR10_f	CCGGACGCTTACAACTAAATCG	Chen et al., 2013
JIOsPR10_r	CACTTCTCAATCACTGCTTGGAA	

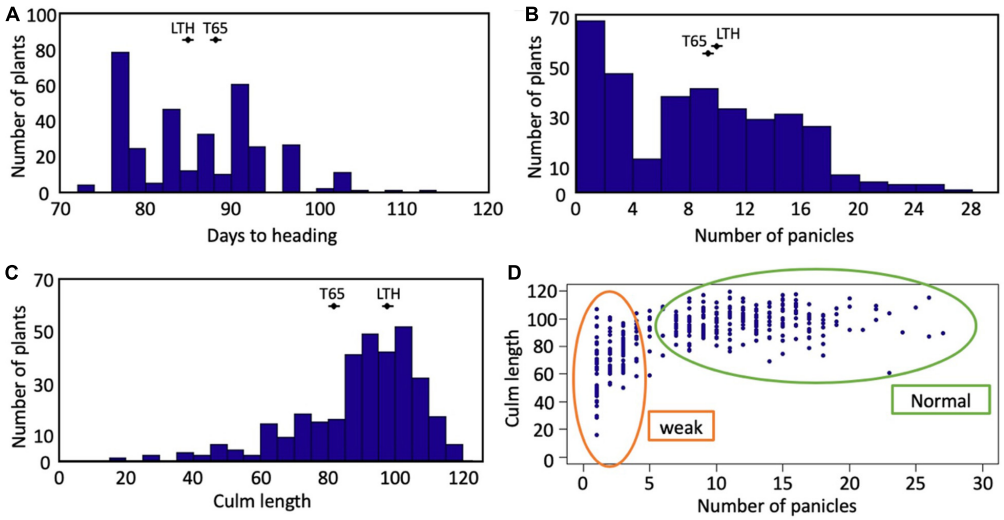

FIGURE 2 | Segregation of traits in F_2 population. Frequency distributions of **(A)** days to heading, **(B)** number of panicles, and **(C)** culm length. The averages of T65 and LTH were shown in each panel. **(D)** Scatter plot between number of panicles and culm length.

RESULTS

Segregation of Weakness

Segregation of weak plants was observed in the field condition (**Supplementary Figure 1**). The frequency distributions of days to heading, the number of panicles, and the culm length of the F_2 population are shown in **Figure 2**. Segregation of these traits was continuous for days to heading and culm length (**Figures 2A,C**). A valley curve was observed in the number of panicles and can therefore be used to distinguish normal plants from weak plants (**Figure 2B**). In the correlation between the number of panicles and culm length (**Figure 2D**), the plants which had four or fewer panicles corresponded well to the weak plants which were visually discriminated.

Quantitative Trait Loci Mapping

The genotypes of the 126 F_2 plants from the population in 2016 were determined by GBS and a linkage map consisting of 862 single nucleotide polymorphism (SNP) markers were obtained. QTL analysis was performed using panicle number per plant as the indicator of hybrid weakness. We identified two significant QTL at a 5% level on chromosomes 1 and 11 (**Table 4** and **Supplementary Figure 2**). Since the segregation of the weakness was controlled by these two major QTLs, each QTL was named *hybrid weakness j 1* (*hwj1*) (chromosome 1) and *hybrid weakness j 2* (*hwj2*) (chromosome 11). The interaction between *hwj1* and *hwj2* was detected by two-factor QTL analysis (**Figure 3**).

Interaction of *hwj1* and *hwj2* Caused the Weakness

Since the hybrid weakness trait was associated with *hwj1* and *hwj2,* and the interaction between these two loci was significant

TABLE 4 | Quantitative trait loci detected in F_2 population.

QTL	Chr[a]	Trait	Position (Mb)	LOD	PVE (%)[b]	AE[c]	DE[d]
hwj1	1	Panicle number	18.53	4.76	15.97	1.53	–4.56
hwj2	11	Panicle number	22.37	12.78	37.32	–5.14	–0.59

[a] *Chromosome.*
[b] *Percentage of variance explained by the QTL.*
[c] *Additive effects of the marker calculated as (average of LTH-average of T65)/2.*
[d] *Dominance effect calculated as average of heterozygotes-(average of LTH + average of T65)/2.*

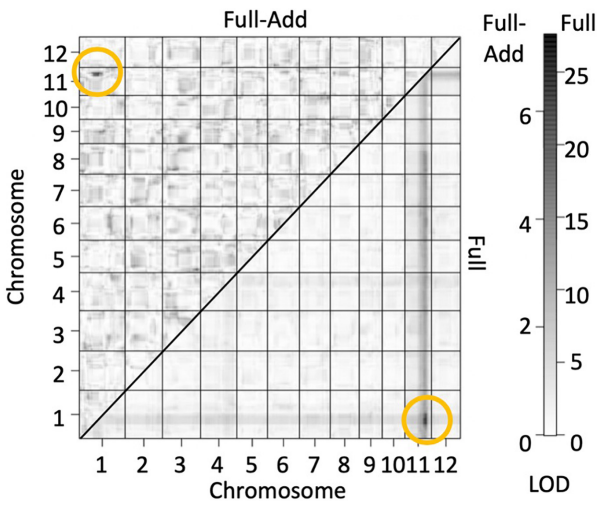

FIGURE 3 | Heat map showing quantitative trait loci (QTL) for number of panicles by assuming interaction of two loci. Bottom and upper diagonals show the LOD score of full models (LOD_T) and full-additive model ($LOD_{Int} = LOD_T - LOD_x - LOD_y$), respectively. The yellow circles indicate sites with high LOD values and interaction between the two factors. The scale on the right shows LOD values of upper (left) and bottom (right) diagonals.

(**Figure 3** and **Supplementary Figure 2**), the genotypes of the 344 F_2 plants were confirmed using the InDel markers linked to *hwj1* and *hwj2*. In this study, *hwj1* and *hwj2* loci were designated as *A* and *B*, wherein T65 and LTH genotypes were designated as *aaBB* and *AAbb*, respectively. The F_2 plants carrying the genotypes of *Aabb* and *aaBb* showed weakness in the field. On the other hand, the double recessive genotype, *aabb*, showed a severe weak phenotype (**Supplementary Figure 3**).

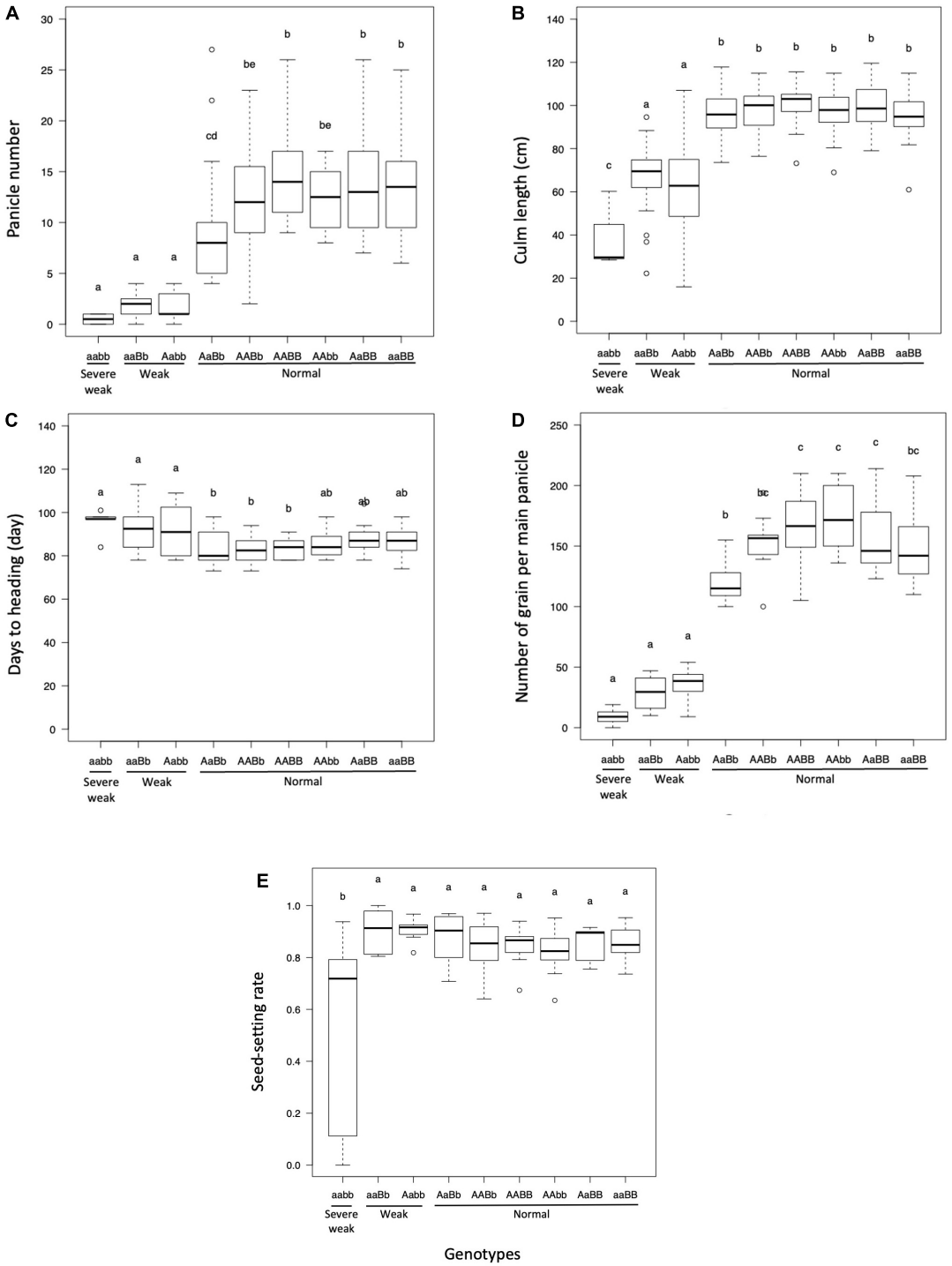

FIGURE 4 | Boxplot of evaluated traits of F_2 population in field condition: **(A)** panicle number, **(B)** culm length (cm), **(C)** days to heading, **(D)** number of grains per main panicle, and **(E)** seed-setting rate. Different letters in each phenotype represents significant differences among the genotypes based on one-way ANOVA followed by Tukey's HSD test ($p < 0.05$).

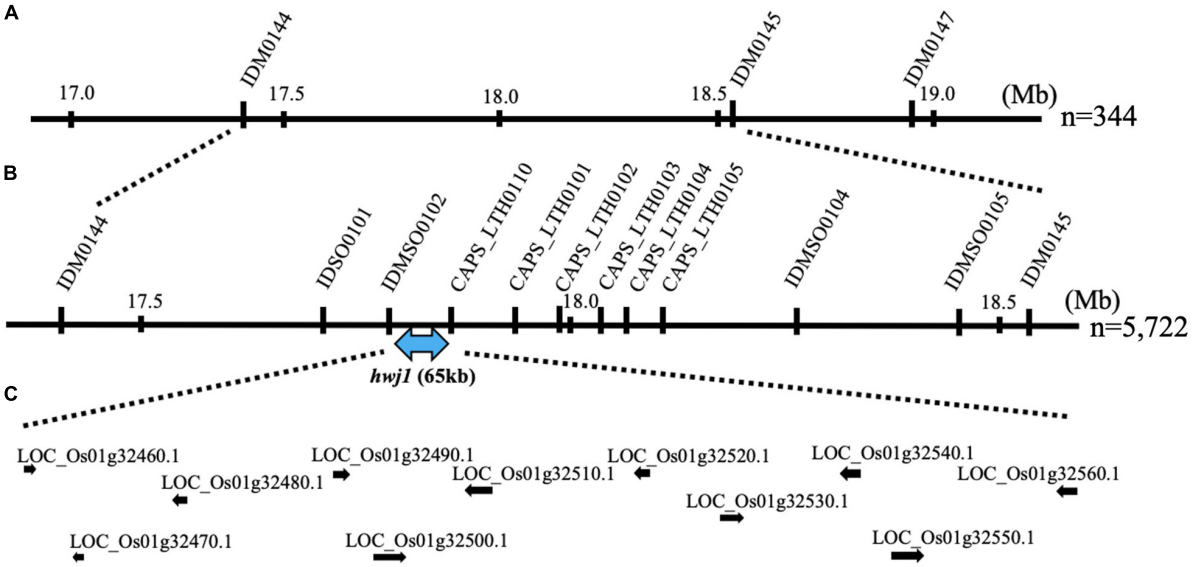

FIGURE 5 | Fine mapping of the *hwj1* locus. **(A)** Initial mapping of *hwj1* on chromosome 1. Numbers show physical positions in megabase (Mb), **(B)** fine mapping of *hwj1*. The *hwj1* locus was delimited between flanking markers IDMSO0102 and CAPS_LTH0110. Blue-double head arrow shows candidate region, and **(C)** annotated genes in the candidate region. The black arrows represent the annotated genes.

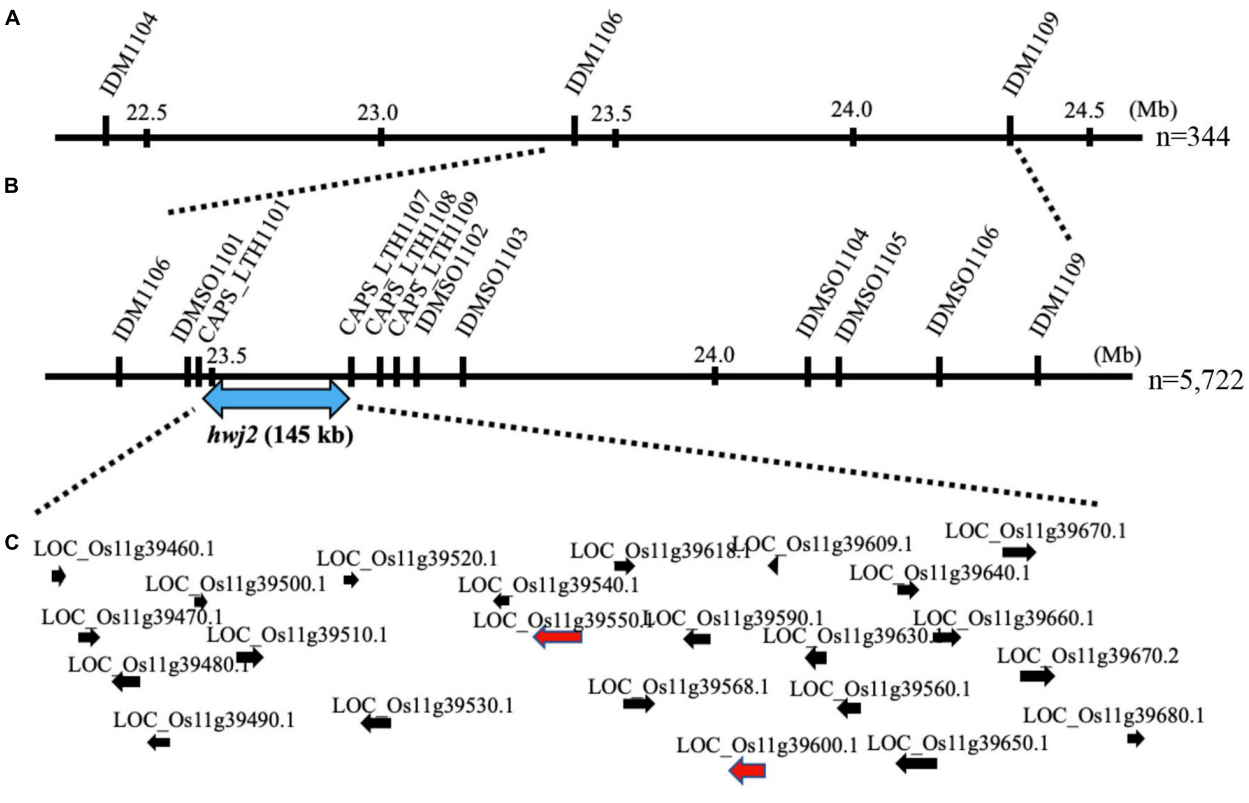

FIGURE 6 | Fine mapping of the *hwj2* locus. **(A)** Initial mapping of *hwj2* on chromosome 11. Numbers show physical positions in megabase (Mb), **(B)** fine mapping of *hwj2*. The *hwj2* locus was delimited between flanking markers CAPS_LTH1101 and CAPS_LTH1107. Blue-double head arrow shows the candidate region, and **(C)** annotated genes in the candidate region. The red arrows represent NB-LRR disease-resistance gene and disease resistance-related gene. The black arrows represent other annotated genes.

Phenotypic Characterization of Hybrid Weakness

The panicle number and culm length of plants with weak and severe weak genotypes were significantly reduced as compared to plants with normal genotypes (**Figures 4A,B**). The days to heading of the plants with weak and severe weak genotypes were delayed, but not significantly different from other genotypes in the field condition (**Figure 4C**). The number of grains per main panicle was significantly reduced in plants with weak and severe weak genotypes (**Figure 4D**). The seed setting rate of severe weak plants was significantly decreased compared to other genotypes (**Figure 4E**).

Fine Mapping of the Causal Loci

To map the genes responsible for *hwj1* on chromosome 1 and *hwj2* on chromosome 11, larger F_2 populations obtained from the T65 and LTH cross were grown in 2018 and 2019. A total of 5,722 plants were used for fine mapping, and the weakness phenotype was discriminated based on visual observation and the number of panicles. Using InDel and CAPS markers, the location of the identified causal loci were fine mapped. The *hwj1* locus was delimited to 65-kb by the flanking markers IDMSO_0102 and CAPS_LTH0110 on chromosome 1 (**Figure 5**). On the other hand, the location of the *hwj2* locus was delimited to 145-kb by the flanking markers CAPS_LTH1101 and CAPS_LTH1107 on chromosome 11 (**Figure 6**).

According to the annotation databases[2,3], a total of eleven annotated genes were located in the region of *hwj1* (**Figure 5** and **Supplementary Table 3**), and 22 annotated genes are located on the region of *hwj2* and includes one NB-LRR disease-resistance gene and one disease resistance-related gene containing NB-ARC domain (**Figure 6** and **Supplementary Table 4**).

Temperature Affects the Hybrid Weakness Caused by *hwj1* and *hwj2*

To investigate the effect of temperature on the weak phenotype, the growth of the seedlings was examined at 24 and 34°C using 10-day-old seedlings. It was observed that 24°C enhances the weak phenotype and 34°C rescues the weakness (**Figures 7**, **8**). At 24°C, the shoot length and the number of roots of severe weak genotypes (*aabb*) and weak genotypes (*aaBb*, *Aabb*) were significantly different from the normal genotypes (*AaBb, AABb, AABB, AAbb, AaBB,* and *aaBB*) (**Figures 8A,E**). In the same way, at 24°C, the root length of severe weak and weak genotypes was shorter compared to normal genotypes (**Figure 8C**). In addition, the root length of the F_1 genotype (*AaBb*) was significantly different from other normal genotypes (**Figure 8C**). Similarly, the F_1 genotype (*AaBb*) had fewer roots compared to other normal genotypes (**Figure 8E**). At 34°C, the shoot length, root length, and the number of roots showed no significant difference (**Figures 8B,D,F**), except for the root length of severe weak genotype (*aabb*) (**Figure 8D**).

[2]http://rice.plantbiology.msu.edu/

[3]http://rapdb.dna.affrc.go.jp/

Morphological and Histological Observations of Embryo in Hybrid Weakness

To confirm whether an abnormality occurred in the embryo, morphological and histological analyses on 1-day-soaked embryos at 24°C were conducted. The seeds were genotyped before soaking using the half seeds. All the F_2 genotypes and parents showed no visible morphological differences (**Figure 9**). In addition, structures of the shoot meristems and radicles observed by vertical sections were normal in weak and severe weak genotypes (**Figures 10D–F**), compared to those of normal genotypes and the parents (**Figures 10A–C**).

Quantitative Reverse Transcription PCR of Pathogenesis-Related Genes

To confirm the autoimmune response that was often detected in previous studies on hybrid weakness, the mRNA expressions of PR genes were compared among genotypes using the seedlings grown at 24°C. The result of qRT-PCR did not show any significant difference in the mRNA expression when compared across genotypes (**Figure 11**). However, the *PBZ1* showed a tendency of increased mRNA levels in severe weak genotypes (**Figure 11A**). The same tendency was observed in *JIOsPR10* of weak and severe weak genotypes (**Figure 11D**).

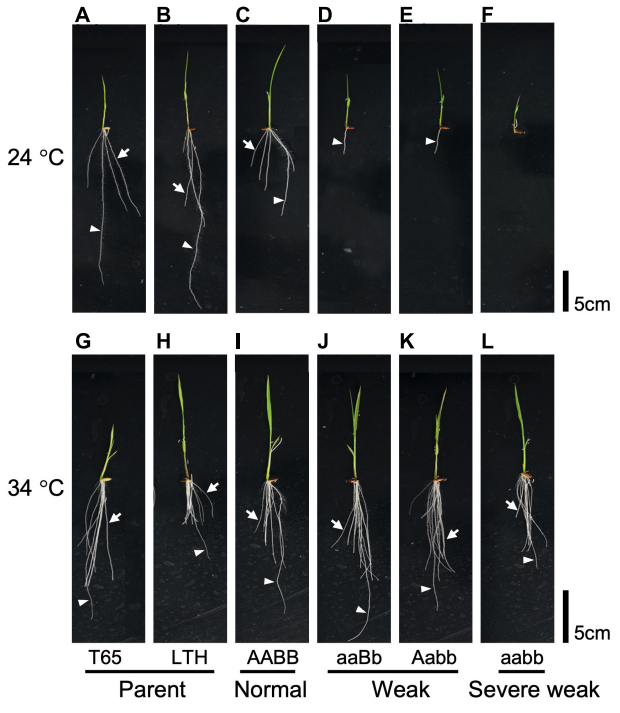

FIGURE 7 | Morphology of 10-day-old seedlings grown under different temperatures. Seedlings grown at 24°C **(A)** T65, **(B)** LTH, **(C)** normal genotype, **(D,E)** weak genotypes, and **(F)** severe weak genotype; seedlings grown at 34°C **(G)** T65, **(H)** LTH, **(I)** normal genotype, **(J,K)** weak genotypes, and **(L)** severe weak genotype. Arrowheads point primary root. Arrows point crown root. Scale bars = 5 cm.

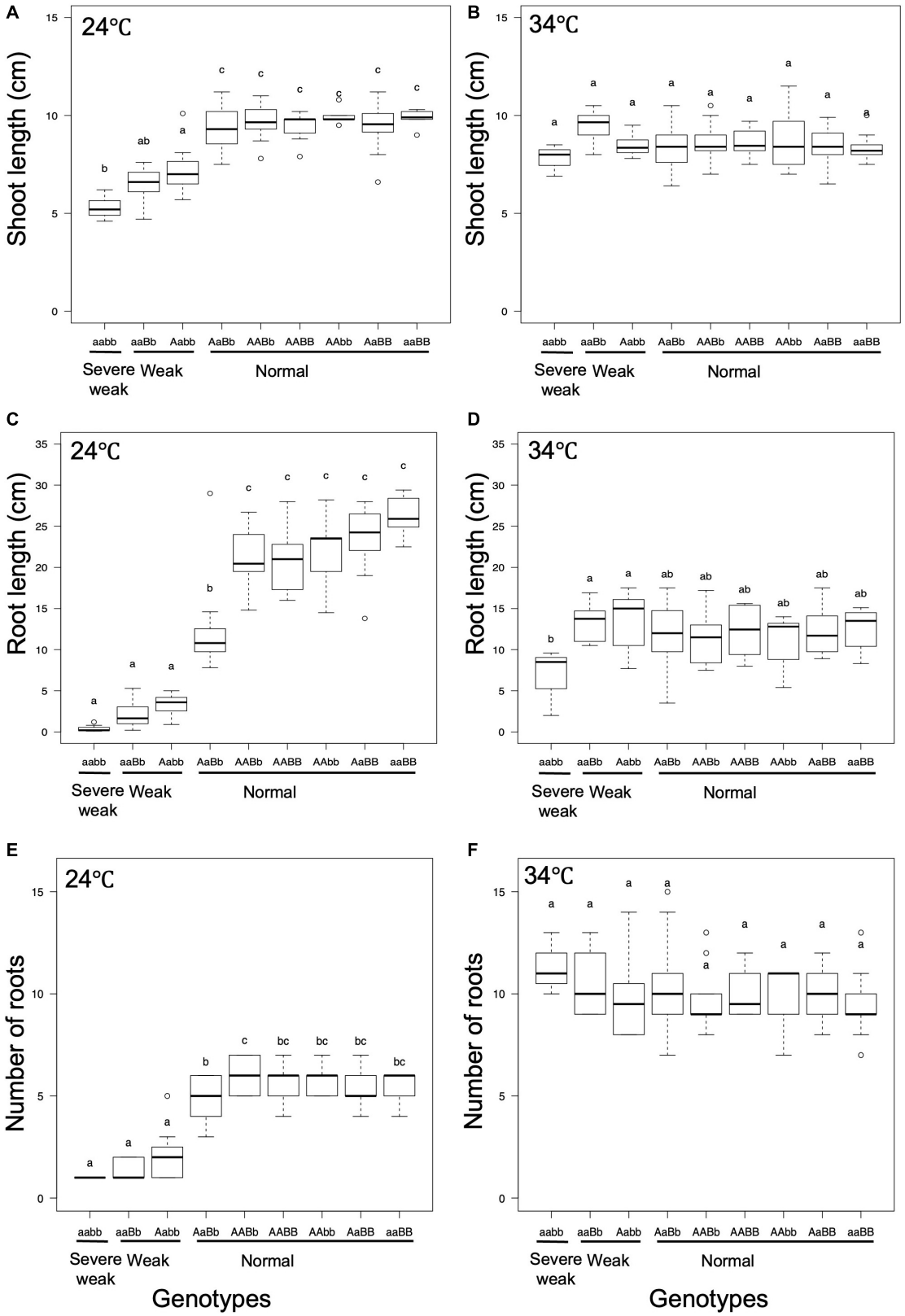

FIGURE 8 | Boxplots of phenotypic evaluations of **(A,B)** shoot length, **(C,D)** root Length, and **(E,F)** number of roots of 10-day-old seedlings grown at 24 and 34°C. Different letters in each phenotype represent significant differences among the genotypes based on one way ANOVA followed by Tukey's HSD test ($p < 0.05$).

DISCUSSION

In the present study, the segregation of weak plants was observed in the F_2 population derived from the cross between T65 and LTH. Generally, the symptoms of hybrid weakness include short stature, less tiller number, impaired root development, late heading date, and significantly decreased grain yield. All these phenomena were observed in weak and severe weak plants (**Figure 4**) under field conditions. The primary cause of weakness was the suppression of crown root formation at a relatively low temperature and the phenotypes of plants with weak and severe weak genotypes were rescued at high temperatures (**Figure 7**). In contrast to the clear difference observed in the seedling test at 24°C (**Figure 7**), continuous segregation was observed in the field (**Figure 2**). This may be due to the temperature in the field which is between 24 and 34°C, and the expression of the weak phenotype was not stable. However, root systems of the weak and severe weak plants in field condition were smaller than the normal plants (data not shown).

Through GBS and QTL analysis, it was indicated that the hybrid weakness was caused by the interaction of two major QTLs. The hybrid weakness explained by the interaction of two

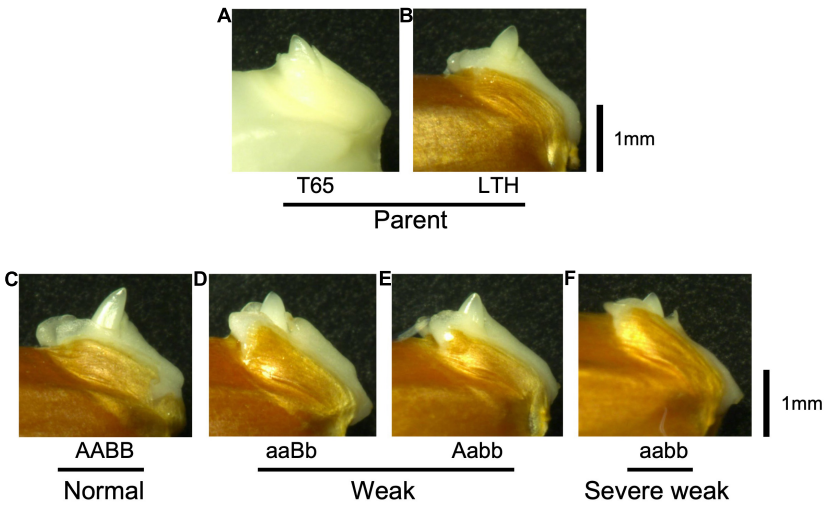

FIGURE 9 | Morphology of embryos of F_2 and the parents 1 day after soaking in water at 24°C. **(A)** T65, **(B)** LTH, **(C)** normal genotype, **(D,E)** weak genotypes, and **(F)** severe weak genotype. Scale bars = 1 mm.

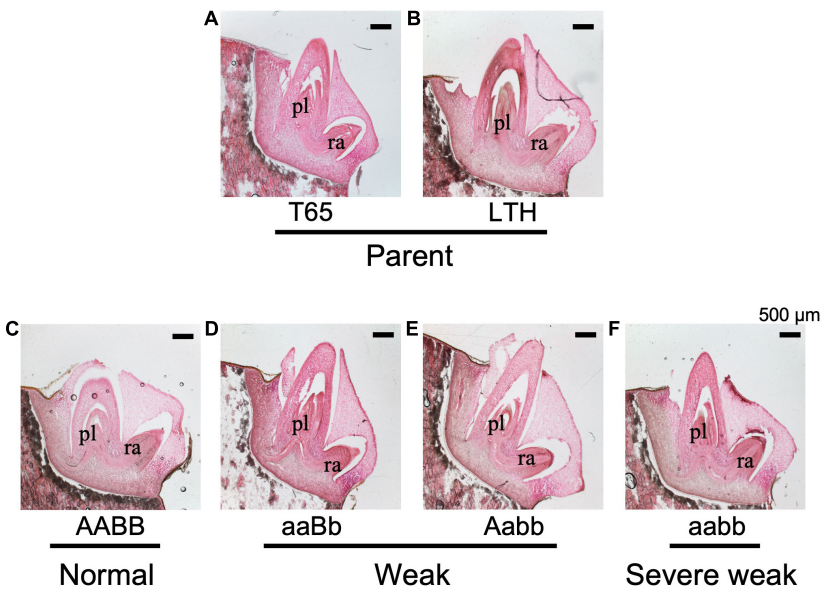

FIGURE 10 | Histological characteristics of embryos of F_2 and the parents in germinating seeds 1 day after soaking in water at 24°C. **(A)** T65, **(B)** LTH, **(C)** normal genotype, **(D,E)** weak genotypes, and **(F)** severe weak genotype. The sections (20 μm thick) were stained with hematoxylin and eosin. pl, plumule (shoot meristem); ra, radicle. Scale bars = 500 μm.

genes has been reported by many scientists (e.g., Oka, 1957; Kubo and Yoshimura, 2002; Yamamoto et al., 2007, 2010; Chen et al., 2014). The two loci in the present study were designated as *hwj1* (chromosome 1) and *hwj2* (chromosome 11). The genotype of T65 and LTH was assumed to be *hwj1hwj1Hwj2Hwj2* (*aaBB*) and *Hwj1Hwj1hwj2hwj2* (*AAbb*), respectively. The plants possessing 3 recessive alleles (*aaBb* and *Aabb*) showed a weak phenotype, and the *aabb* genotype resulted in severe weakness, thus the expected segregation ratio of normal: weak: severe weak plants is 11: 4: 1 (**Supplementary Figure 3**). This model of inheritance is similar to that of Fukuoka et al. (2005), although most of the previous studies reported a segregation ratio of normal: weak = 15:1.

Because of the interaction of the two loci and the unstable phenotype in the field, fine mapping of the two loci was done using plants with clear phenotypes. A total of 5,722 F_2 plants were used for fine mapping, and the region of *hwj1* and *hwj2* were narrowed down to 65-kb on chromosome 1 (**Figure 5**) and 145-kb on chromosome 11 (**Figure 6**), respectively. In the 65-kb candidate region of *hwj1*, 11 genes were predicted, and some of these are retrotransposon-like genes (Nadir et al., 2019). On the other hand, the candidate region of *hwj2* contains two disease resistance-related genes, thus these two genes were primary candidates for *hwj2*. The sequencing of the candidate regions and transgenic experiments are needed to confirm the causal genes.

In general, several hybrid weakness cases in some plants were reported to be low temperature-dependent and phenotypes of weak plants were rescued at high temperature (Saito et al., 2007; Alcázar et al., 2009; Shiragaki et al., 2021; Yoneya et al., 2021). To date, only one high temperature-dependent hybrid weakness in rice was reported (Chen et al., 2014). In our study, it was observed that the hybrid weakness was induced by low temperature (24°C), and weakness syndrome did not appear at high temperature (34°C). At 24°C, the shoot length, root length, and the number of roots of 10-day-old weak and severe weak genotypes were significantly shorter compared to normal genotypes (**Figures 8A,C,E**). Moreover, the root length of severe weak genotypes (*aabb*) was shorter than 1 cm (**Figure 8C**) and they did not grow crown roots (**Figure 7F**). This morphological feature of severe weak phenotype is the same as the previous description by Ichitani et al. (2007) wherein plants that were

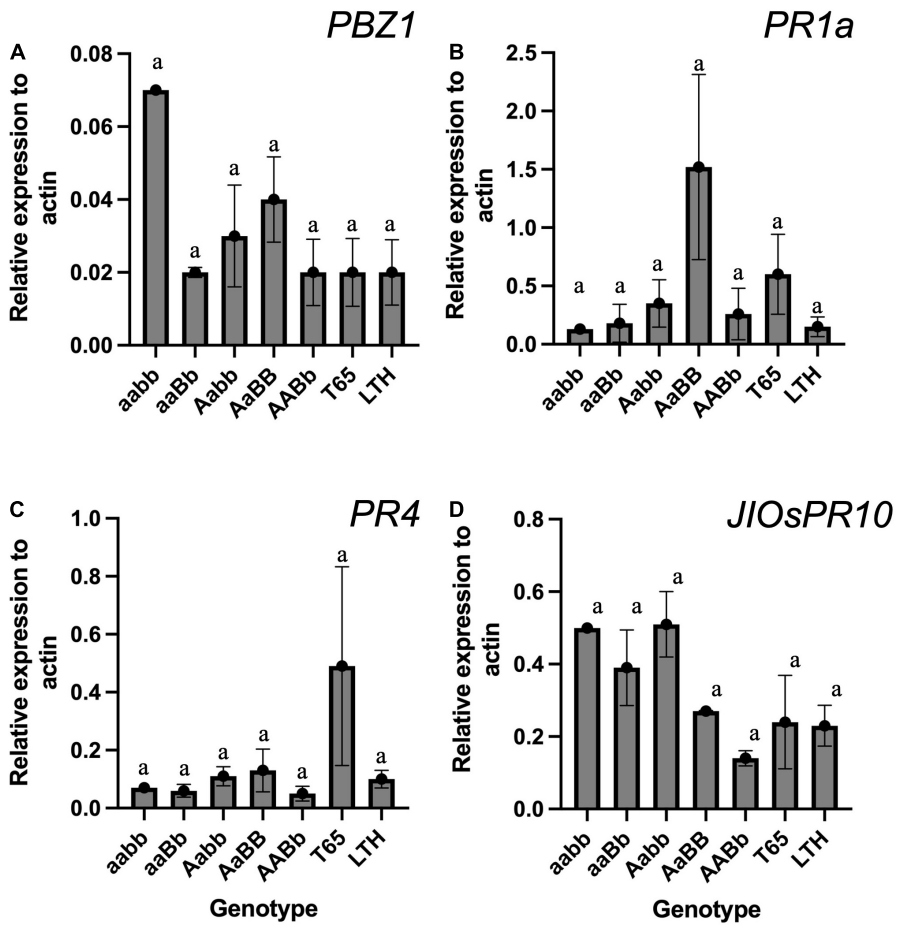

FIGURE 11 | The expression level of pathogen-related genes (PRs) in the leaf blades and sheaths of severe weak genotype (*aabb*), weak genotypes (*aaBb* and *Aabb*), normal genotype (*AaBB* and *AABb*), and the parents (T65 and LTH) at 7-day-old seedling stage. **(A)** *PBZ1*, **(B)** *PR1a*, **(C)** *PR4*, and **(D)** *JIOsPR1*. Each qRT-PCR analysis was performed with three biological replications, but severe weak plants had no replication. Bars represent the means ± SE. Actin was used as the endogenous control for normalization.

homozygotes for *Hwc1* and *Hwc2* showed a very short shoot and did not grow roots. In our study, the root length of weak genotypes (*aaBb*, *Aabb*) was shorter than 5 cm with few crown roots (**Figures 8C,E**). Furthermore, the root length and number of roots of the F_1 genotype (*AaBb*) were different from other normal genotypes (*AABb, AABB, AAbb, AaBB,* and *aaBB*) (**Figures 8C,E**). Conversely, at 34°C, the shoot length, root length, and number of roots of the severe weak and weak genotypes were restored. To date, several studies were also reported wherein plants with weak genotypes were rescued in high temperatures such as in *Nicotiana* species (Yamada et al., 1999) and *Gossypium* species (Philips, 1977). Thus, high temperature treatment might be generally effective in rescuing hybrid plants from weakness. Taken together, hybrid weakness in rice affects more the root than the shoot.

According to morphological and histological observations, the embryos of weak and severe weak genotypes were normal (**Figure 9**). Taken together, hybrid weakness simply affects postembryonic development and is independent of embryogenesis. In addition, hybrid weakness is triggered by low temperature after germination. In rice, the temperature sensitivity was reported by the previous studies such as *Hwc* genes (Saito et al., 2007), *hw3/hw4* genes (Fu et al., 2013), *thb1* (Yoneya et al., 2021) where low temperature enhanced the weakness, while *Hwi* genes (Chen et al., 2013) enhanced the weakness at high temperature.

In our study, the candidate region of *hwj2* contained disease resistance-related genes. So far, at least one of the causal gene combinations for rice hybrid weakness was disease-related genes (e.g., Yamamoto et al., 2010; Chen et al., 2014). Therefore, it was hypothesized that of *hwj1* and *hwj2* were disease-related genes. According to the annotation databases (see text footnote 2 and 3), one LRR (leucine-rich repeat) family protein (LOC_Os11g39550) and one disease resistance-related gene containing NB-ARC domain (LOC_Os11g39600) is located in the candidate region of *hwj2,* implying that these genes are the causal genes. However, many other genes are also located in that region (**Supplementary Table 4**). The *Hwa1* and *Hwa2* (Ichitani et al., 2011) and *hwg2* (Fukuoka et al., 2005) shared the same candidate region as *hwj2* on chromosome 11. Thus, there is a possibility that a common allele of an NB-LRR gene is the cause of hybrid weakness observed in different crosses and different counterpart genes.

In previous studies, the mRNA expression levels of the PR genes were upregulated in plants showing symptoms of hybrid weakness (Yamamoto et al., 2010; Chen et al., 2014). Contrary to the expectation, the results of this study did not show a significant increase in the mRNA levels of PR genes (**Figure 11**). However, a tendency of autoimmune response was observed in *PBZ1* of severe weak genotype (**Figure 11A**), and the same tendency was found in *JIOsPR10* of both weak and severe weak genotypes (**Figure 11D**). Thus, the authors presume that the hybrid weakness is accompanied by temperature sensitivity and is caused by an autoimmune response. This result should be carefully translated because RNA was extracted from the aerial part of young seedlings where no symptom was observed. To date, several studies have shown that the activation of PR genes is associated with the occurrence of symptoms under various types of conditions (Yamamoto et al., 2010; Chen et al., 2014; Shiragaki et al., 2019). Therefore, sampling factors such as growth stage and organ source must be taken into consideration.

CONCLUSION

In the present study, the phenotypic property of the F_2 weakness observed in the T65/LTH cross was characterized and two interacting loci, *hwj1* and *hwj2*, were identified. These loci were identified to be associated with temperature-sensitive hybrid weakness. To clarify the molecular mechanism of the hybrid weakness, cloning of the *hwj1* and *hwj2* by further sequencing and transgenic experiments is necessary. In addition, it is also necessary to clarify the mRNA expression of PR genes to confirm if the hybrid weakness was caused through the mechanism of an autoimmune response. The distribution of *hwj1* and *hwj2* alleles is also of future interest.

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: https://www.ddbj.nig.ac.jp/, DRA014130.

AUTHOR CONTRIBUTIONS

HS and KD: conceptualization and data curation. TKS, MK, HS, YI, and KD: investigation. TKS and HS: writing—original draft preparation. TKS, VPR, SN, and KD: writing—review and editing. KD: supervision. YI and KD: funding acquisition. All authors read and approved the submitted version.

FUNDING

This work was supported in part by National Bioresource Project Rice (NBRP Rice), RIKEN Cluster for Science, Technology and Innovation Hub (RCSTI), SATREPS project "Rice genome breeding system for developing rice plant for non-irrigated area," and JSPS KAKENHI Grant Numbers 20KK0138/21K05522/22H02310.

ACKNOWLEDGMENTS

TKS was grateful to Japan International Cooperative Agency (JICA) Innovative Asia program for the scholarship. We would like to thank Kimiyo Inukai for excellent technical assistance on qRT-PCR.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.908000/full#supplementary-material

REFERENCES

Alcázar, R., García, A. V., Kronholm, I., de Meaux, J., Koornneef, M., Parker, J. E., et al. (2010). Natural variation at Strubbelig receptor kinase 3 drives immune-triggered incompatibilities between *Arabidopsis thaliana* accessions. *Nat. Genet.* 42, 1135–1139. doi: 10.1038/ng.704

Alcázar, R., García, A. V., Parker, J. E., and Reymond, M. (2009). Incremental steps toward incompatibility revealed by *Arabidopsis* epistatic interactions modulating salicylic acid pathway activation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 334–339. doi: 10.1073/pnas.0811734106

Bomblies, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J. L., et al. (2007). Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol.* 5:e236. doi: 10.1371/journal.pbio.0050236

Bomblies, K., and Weigel, D. (2007). Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nat. Rev. Genet.* 8, 382–393. doi: 10.1038/nrg2082

Broman, K. W., Wu, H., Sen, S., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19, 889–890. doi: 10.1093/bioinformatics/btg112

Chen, C., Chen, H., Lin, Y.-S., Shen, J.-B., Shan, J.-X., Qi, P., et al. (2014). A two-locus interaction causes interspecific hybrid weakness in rice. *Nat. Commun.* 5:3357. doi: 10.1038/ncomms4357

Chen, C., Chen, H., Shan, J.-X., Zhu, M.-Z., Shi, M., Gao, J.-P., et al. (2013). Genetic and physiological analysis of a novel type of interspecific hybrid weakness in rice. *Mol. Plant* 6, 716–728. doi: 10.1093/mp/sss146

Coyne, J. A., and Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.

Dellaporta, S. L., Wood, J., and Hicks, J. B. (1983). A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.* 1, 19–21. doi: 10.1007/BF02712670

Dobzhansky, T. (1937). *Genetics and the Origin of Species*. New York, NY: Columbia University Press.

Dropkin, V. (1969). The necrotic reaction of tomatoes and other hosts resistant to *Meloidogyne*: reversal by temperature. *Phytopathology* 59, 1632–1637.

Fu, C.-Y., Wang, F., Sun, B.-R., Liu, W.-G., Li, J.-H., Deng, R.-F., et al. (2013). Genetic and cytological analysis of a novel type of low temperature-dependent intrasubspecific hybrid weakness in rice. *PLoS One* 8:e73886. doi: 10.1371/journal.pone.0073886

Fukuoka, S., Namai, H., and Okuno, K. (1998). RFLP mapping of the genes controlling hybrid breakdown in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 97, 446–449. doi: 10.1007/s001220050915

Fukuoka, S., Newingham, M. C. V., Ishtiaq, M., Nagamine, T., Kawase, M., and Okuno, K. (2005). Identification and mapping of two new loci for hybrid breakdown in cultivated rice. *Rice Genet. Newslett.* 22:29.

Furuta, T., Ashikari, M., Jena, K. K., Doi, K., and Reuscher, S. (2017). Adapting genotyping-by-sequencing for rice F2 populations. *G3* 7, 881–893. doi: 10.1534/g3.116.038190

Ichitani, K., Fukuta, Y., Taura, S., and Sato, M. (2001). Chromosomal location of Hwc2, one of the complementary hybrid weakness genes, in rice. *Plant Breed.* 120, 523–525. doi: 10.1046/j.1439-0523.2001.00644.x

Ichitani, K., Namigoshi, K., Sato, M., Taura, S., Aoki, M., Matsumoto, Y., et al. (2007). Fine mapping and allelic dosage effect of Hwc1, a complementary hybrid weakness gene in rice. *Theor. Appl. Genet.* 114, 1407–1415. doi: 10.1007/s00122-007-0526-0

Ichitani, K., Taura, S., Tezuka, T., Okiyama, Y., and Kuboyama, T. (2011). Chromosomal location of HWA1 and HWA2, complementary hybrid weakness genes in rice. *Rice* 4, 29–38. doi: 10.1007/s12284-011-9062-2

Jeuken, M. J. W., Zhang, N. W., McHale, L. K., Pelgrom, K., den Boer, E., Lindhout, P., et al. (2009). Rin4 causes hybrid necrosis and race-specific resistance in an interspecific lettuce hybrid. *Plant Cell* 21, 3368–3378. doi: 10.1105/tpc.109.070334

Jiang, W., Chu, S.-H., Piao, R., Chin, J.-H., Jin, Y.-M., Lee, J., et al. (2008). Fine mapping and candidate gene analysis of hwh1 and hwh2, a set of complementary genes controlling hybrid breakdown in rice. *Theor. Appl. Genet.* 116, 1117–1127. doi: 10.1007/s00122-008-0740-4

Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286

Kawamoto, T. (2003). Use of a new adhesive film for the preparation of multi-purpose fresh-frozen sections from hard tissues, whole-animals, insects and plants. *Arch. Histol. Cytol.* 66, 123–143. doi: 10.1679/aohc.66.123

Koide, Y., Onishi, K., Kanazawa, A., and Sano, Y. (2008). "Genetics of speciation in rice," in *Rice Biology in the Genomics Era Biotechnology in Agriculture and Forestry*, eds H.-Y. Hirano, Y. Sano, A. Hirai, and T. Sasaki (Berlin: Springer), 247–259. doi: 10.1007/978-3-540-74250-0_19

Kubo, T., Aida, Y., Nakamura, K., Tsunematsu, H., Doi, K., and Yoshimura, A. (2002). Reciprocal chromosome segment substitution series derived from Japonica and Indica cross of rice (*Oryza sativa* L.). *Breed. Sci.* 52, 319–325. doi: 10.1270/jsbbs.52.319

Kubo, T., and Yoshimura, A. (2002). Genetic basis of hybrid breakdown in a Japonica/Indica cross of rice, *Oryza sativa* L. *Theor. Appl. Genet.* 105, 906–911. doi: 10.1007/s00122-002-1059-1

Kuboyama, T., Saito, T., Matsumoto, T., Wu, J., Kanamori, H., Taura, S., et al. (2009). Fine mapping of HWC2, a complementary hybrid weakness gene, and haplotype analysis around the locus in rice. *Rice* 2, 93–103.

Kumar, V., and Shukla, Y. (2014). Pre-breeding: its applications in crop improvement. *Double Helix Res.* 16, 199–202.

Lee, J. A. (1981). Genetics of D3 complementary lethality in *Gossypium hirsutum* and *G. barbadense*. *J. Hered.* 72, 299–300. doi: 10.1093/oxfordjournals.jhered.a109510

Matsubara, K., Ando, T., Mizubayashi, T., Ito, S., and Yano, M. (2007). Identification and linkage mapping of complementary recessive genes causing hybrid breakdown in an intraspecific rice cross. *Theor. Appl. Genet.* 115, 179–186. doi: 10.1007/s00122-007-0553-x

Muller, H. (1942). Isolating mechanisms, evolution, and temperature. *Biol. Symp.* 6, 71–125. doi: 10.1093/genetics/161.3.939

Nadir, S., Li, W., Zhu, Q., Khan, S., Zhang, X.-L., Zhang, H., et al. (2019). A novel discovery of a long terminal repeat retrotransposon-induced hybrid weakness in rice. *J. Exp. Bot.* 70, 1197–1207. doi: 10.1093/jxb/ery442

Oka, H.-I. (1957). Phylogenetic differentiation of cultivated rice. XV. complementary lethal genes in rice. *Jpn. J. Genet.* 32, 83–87. doi: 10.1266/jjg.32.83

Philips, L. L. (1977). Interspecific incompatibility in *Gossypium*. IV. Temperature-conditional lethality in hybrids of *G. klozschianum*. *Am. J. Bot.* 64, 914–915.

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One* 7:e32253. doi: 10.1371/journal.pone.0032253

Rieseberg, L. H., and Willis, J. H. (2007). Plant speciation. *Science* 317, 910–914. doi: 10.1126/science.1137729

Saito, T., Ichitani, K., Suzuki, T., Marubashi, W., and Kuboyama, T. (2007). Developmental observation and high temperature rescue from hybrid weakness in a cross between Japanese rice cultivars and Peruvian rice cultivar "Jamaica.". *Breed. Sci.* 57, 281–288. doi: 10.1270/jsbbs.57.281

Shii, C. T., Mok, M. C., Temple, S. R., and Mok, D. W. S. (1980). Expression of developmental abnormalities in hybrids of *Phaseolus vulgaris* L. *J. Hered.* 71, 218–222. doi: 10.1093/oxfordjournals.jhered.a109353

Shiragaki, K., Furukawa, H., Yokoi, S., and Tezuka, T. (2021). Temperature-dependent sugar accumulation in interspecific capsicum F1 plants showing hybrid weakness. *J. Plant Res.* 134, 1199–1211. doi: 10.1007/s10265-021-01340-1

Shiragaki, K., Iizuka, T., Ichitani, K., Kuboyama, T., Morikawa, T., Oda, M., et al. (2019). HWA1- and HWA2-mediated hybrid weakness in rice involves cell death, reactive oxygen species accumulation, and disease resistance-related gene upregulation. *Plants* 8:450. doi: 10.3390/plants8110450

Stebbins, G. L. (1950). "Isolation and the origin of species" in *Variation and Evolution in Plants*, ed. G. L. Stebbins (New York, NY: Columbia University Press), 189–250. doi: 10.7312/steb94536-007

Tezuka, T., Kuboyama, T., Matsuda, T., and Marubashi, W. (2007). Possible involvement of genes on the Q chromosome of *Nicotiana tabacum* in expression of hybrid lethality and programmed cell death during interspecific hybridization to *Nicotiana debneyi*. *Planta* 226, 753–764. doi: 10.1007/s00425-007-0522-2

van Loon, L. C. (1985). Pathogenesis-related proteins. *Plant Mol. Biol.* 4, 111–116. doi: 10.1007/BF02418757

Vaughan, D. A., Morishima, H., and Kadowaki, K. (2003). Diversity in the *Oryza* genus. *Curr. Opin. Plant Biol.* 6, 139–146. doi: 10.1016/S1369-5266(03)00009-8

Yamada, T., Marubashi, W., and Niwa, M. (1999). Detection of four lethality types in interspecific crosses among *Nicotiana* species through the use of three rescue methods for lethality. *Breed. Sci.* 49, 203–210. doi: 10.1270/jsbbs.49.203

Yamamoto, E., Takashi, T., Morinaka, Y., Lin, S., Kitano, H., Matsuoka, M., et al. (2007). Interaction of two recessive genes, hbd2 and hbd3, induces hybrid breakdown in rice. *Theor. Appl. Genet.* 115, 187–194.

Yamamoto, E., Takashi, T., Morinaka, Y., Lin, S., Wu, J., Matsumoto, T., et al. (2010). Gain of deleterious function causes an autoimmune response and Bateson–Dobzhansky–Muller incompatibility in rice. *Mol. Genet. Genomics* 283, 305–315. doi: 10.1007/s00438-010-0514-y

Yoneya, Y., Wakabayashi, T., and Kato, K. (2021). The temperature sensitive hybrid breakdown 1 induces low temperature-dependent intrasubspecific hybrid breakdown in rice. *Breed. Sci.* 71, 268–276. doi: 10.1270/jsbbs.20129

Yoshimura, A., Nagayama, H., Sobrizal, Kurakazu, T., Sanchez, P. L., Doi, K., et al. (2010). Introgression lines of rice (*Oryza sativa* L.) carrying a donor genome from the wild species, *O. glumaepatula* Steud. and *O. meridionalis* Ng. *Breed. Sci.* 60, 597–603. doi: 10.1270/jsbbs.60.597

Conflict of Interest: HS was employed by Environmental Control Center Co., Ltd.

The remaining 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

MINI REVIEW
published: 12 July 2022
doi: 10.3389/fpls.2022.898206

Influence of Gender Bias on Distribution of Hybrid Sterility in Rice

*Zin Mar Myint and Yohei Koide**

Laboratory of Plant Breeding, Research Faculty of Agriculture, Hokkaido University, Sapporo, Japan

Hybrid sterility genes define species identities, setting reproductive barriers between distantly related *Oryza* relatives. They induce allelic-specific selective gametic abnormalities by killing pollens, embryo sacs, or both, and thus resulting in the male specific transmission ratio distortion (*m*TRD), female specific transmission ratio distortion (*f*TRD), and/or sex-independent transmission ratio distortion (*si*TRD) in hybrids. Although more than 50 hybrid sterility genes have been reported, comprehensive analysis on the distributional pattern of TRD systems in *Oryza* species is limited. In this review, we surveyed the TRD systems and the underlying possible mechanisms in these species. In rice, pollen killers which cause *m*TRD are often observed in higher frequency than egg killers and gamete eliminators, which are factors affecting *f*TRD and *si*TRD, respectively. Due to the rather massive population of pollen grains, their reduction in the number caused by hybrid sterility possesses a smaller selective disadvantage to the hybrid individuals, in contrast to female gamete abortion. The pattern of TRD distribution displays less abundancy in *si*TRD. It suggests that fixation of *si*TRD might require a certain time rather than single sex-specific factors. The presence of linked sterility factors worked for *m*TRD and *f*TRD, and strength of their linkage in chromosomal regions might determine the type of sterility and TRD. The study of TRD systems has a potential to reveal the relationships between selfish genes and their functions for reproductive isolation.

Keywords: rice, hybrid sterility, selfish genes, reproductive isolating barrier, transmission ratio distortion

OPEN ACCESS

Edited by:
Mallikarjuna Swamy, International Rice Research Institute (IRRI), Philippines

Reviewed by:
Jing Li, Yunnan Academy of Agricultural Sciences, China
Aris Hairmansis, Indonesian Center for Rice Research, Indonesia

***Correspondence:**
Yohei Koide
ykoide@abs.agr.hokudai.ac.jp

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

Received: *17 March 2022*
Accepted: *11 May 2022*
Published: *12 July 2022*

Citation:
Zin Mar Myint and Koide Y (2022) Influence of Gender Bias on Distribution of Hybrid Sterility in Rice. Front. Plant Sci. 13:898206. doi: 10.3389/fpls.2022.898206

INTRODUCTION

Based on interbreeding, the biological species concept stated, "Species are groups of interbreeding natural populations that are reproductively isolated from other such groups." (Mayr, 1996; Coyne and Orr, 2004). It literally means that the boundaries between the population of one species and that of another are defined by reproductive barriers, and the mechanism itself is referred to as "reproductive isolation", in which genetically based intrinsic barriers prevent gene flow between populations of different species (Nosil, 2013). Reproductive isolation maintains species identity, and evolutionary biologists are highly interested in this area of study. Reproductive isolation is considered to be incidentally acquired as a by-product of other divergences between species' populations (Nosil, 2013). The genetic variations gradually increase over time, and the accumulation of those changes and divergences is probably either neutral or wholesome to its own genetic background while it works as a deleterious factor in other alternative genetic backgrounds, exclusively in a heterozygous state (Turelli, 1998). It is possible for hybrid dysfunction to appear due to a divergence in multiple genomic regions across the whole genome in a wide variety of species populations. The Bateson-Dobzhansky-Muller (BDM) model is a model in which negatively epistatic interaction between divergent alleles contributes deleterious effects within

hybrid populations (Bateson, 1909; Dobzhansky, 1937; Muller, 1942; Kubo et al., 2008; Nosil, 2013; Xie et al., 2019a) and is a widely accepted model of the emergence of reproductive barriers. However, are genes for reproductive isolation always neutral or wholesome when they spread in a population? By summarizing genes for hybrid sterility, a form of reproductive barrier, we will discuss how selfish genes act to build species barriers in rice.

REPRODUCTIVE BARRIERS IN RICE

In plants, reproductive barriers can be basically divided into two categories, namely prezygotic reproductive isolation and postzygotic reproductive isolation, based on the developmental stage in which they give rise (Koide et al., 2008b; Ouyang and Zhang, 2013; Zin Mar et al., 2021). While prezygotic isolation occurs during the formation of the zygote, the latter restricts the introgressive gene flow in crossed populations, inducing hybrid arrest after fertilization at different developmental stages and/or other advanced generations (Chen et al., 2016). Hybrid sterility, the gametic disorder at their reproductive stage with the failure to produce fertile male and/or female gametes in normally grown hybrid plants, is one form of postzygotic reproductive isolation. The genus *Oryza*, which consists of two cultivated rice species and 22 wild rice relatives (Ammiraju et al., 2010), is a valuable pool for improvement of agricultural traits. In rice genetics, interspecific and intraspecific hybrid sterility, which is one of the most pronounced forms of postzygotic reproductive isolation, has been the most extensively investigated subject across a wide variety of genomic regions in a substantial number of different rice populations (Ouyang and Zhang, 2013; Li et al., 2020).

THE GENETIC MODELS OF HYBRID STERILITY

Based upon either sporophytic or gametophytic function and the number of loci involved, altogether four separate genic models underly rice hybrid sterility: namely (1) One-locus sporo-gametophytic interaction, (2) Duplicate gametic lethals, (3) One-locus sporophytic interaction, and (4) Complementary sporophytic interaction (Oka, 1957; Sano et al., 1979; Koide et al., 2008b). Among them, one-locus sporo-gametophytic interaction fits with genetic mechanisms for the majority of sterility loci in rice (Oka, 1957; Xie et al., 2019a). Therefore, we will focus on single or tightly linked sterility loci in this review (see **Table 1**). Hybrid sterility loci, such as *S1*, *S2,* and *S5*, can perfectly explain how one-locus sporo-gametophytic model works in rice sterility (Sano et al., 1979; Chen et al., 2009; Yang et al., 2012, 2016; Xie et al., 2017; Koide et al., 2018; Zin Mar et al., 2021). As an example, for *S2* locus, although $S2^g$ (allele derived from *oryza glaberrima*) and $S2^s$ (allele derived from *O. sativa*) are neutral in their respective backgrounds, the incompatible interaction between these two alleles occurred in heterozygous hybrids. As a result, neither male nor female gametes which carry $S2^g$ allele survive, causing preferential transmission of selfish allele $S2^s$ in later generations (Sano et al., 1979; Zin Mar et al., 2021). Such a preferential transmission of one of the two alleles is referred to as transmission ratio distortion (TRD). Selfish genes causing TRD take advantage over their alternative alleles, and preferentially promote their own distribution in populations, with the help of incompatibilities that occur in heterozygous hybrids.

SEX-SPECIFICITY OF TRD IN HYBRID STERILITY IN RICE

Although several underlying mechanisms behind the TRD phenomenon, including non-random segregation of chromosomes during meiosis (Pardo-Manuel de Villena and Sapienza, 2001; Birchler et al., 2003; Fishman and Willis, 2005; Koide et al., 2008a), unequal gametic success in fertilization (Price, 1997; Diaz and Macnair, 1999; Seymour et al., 2019), and embryo lethality (Lyttle, 1991; Silver, 1993; Price, 1997; Diaz and Macnair, 1999; Ubeda and Haig, 2005; Moyle, 2006) has been reported in plants, selective gametic abnormality (Lyttle, 1991; Silver, 1993; Ubeda and Haig, 2005; Moyle, 2006; Koide et al., 2008a) caused by hybrid sterility locus is the most frequently observed in *Oryza* species. The selective abnormality occurs in either male/female-gametes or sex-independently. In this report, the former and latter are termed *m*TRD (male specific transmission ratio distortion)/*f* TRD (female specific transmission ratio distortion) and *si*TRD (sex-independent transmission ratio distortion), respectively, (Maguire, 1963; Sano, 1983; Koide et al., 2008c; Ouyang and Zhang, 2013). Therefore, in general, hybrid sterility genes in rice can be divided into three subcategories: pollen killer (PK: which induces hybrid male sterility and *m*TRD), egg killer (EK: which results in hybrid female sterility and *f* TRD), and gamete eliminator (GE: which eliminates both pollen and embryo sac causing *si*TRD) (**Figure 1**).

Several loci causing TRD have been cloned and their underlying mechanisms have been studied. The *S5* locus is the most renowned hybrid sterility locus in the study of rice reproductive barriers. This locus is an example of "Killer-Protector system", and its effect is shown in the failure of embryo sac fertility (We note that "Killer" in the Killer-Protector system and "Killer" in PK or EK have different actions. The former "Killer" induces the abortion of gametes with both alleles. The "Protector" selectively protects gametes with one of two alleles from the "Killer" function. As a result, the preferential dysfunction of gametes with one of two alleles occurs. The factor causing preferential function is referred to as PK, EK, or GE. Thus, PK and EK is a concept of a factor with the combined function of "Killer" and "Protector".). Three different tightly linked genes (*ORF3*, *ORF4*, and *ORF5*) regulate the tripartite complex in indica-japonica crossed species. The indica-derived *S5* allele (*ORF3*+, *ORF4*–, and *ORF5*+) contained both functionally active Protector and Killer genes, *ORF3*+ and *ORF5*+, while its allelic japonica-derived *S5* region (*ORF3*–, *ORF4*+, and *ORF5*–) carries inactive Protector and Killer genes, *ORF3*– and *ORF5*–. However, the indica-derived Killer factor *ORF5*+ enacts its killing function with the assistance of its japonica-derived partner *ORF4*+ in the heterozygous combination. Meanwhile, *ORF3*+, which

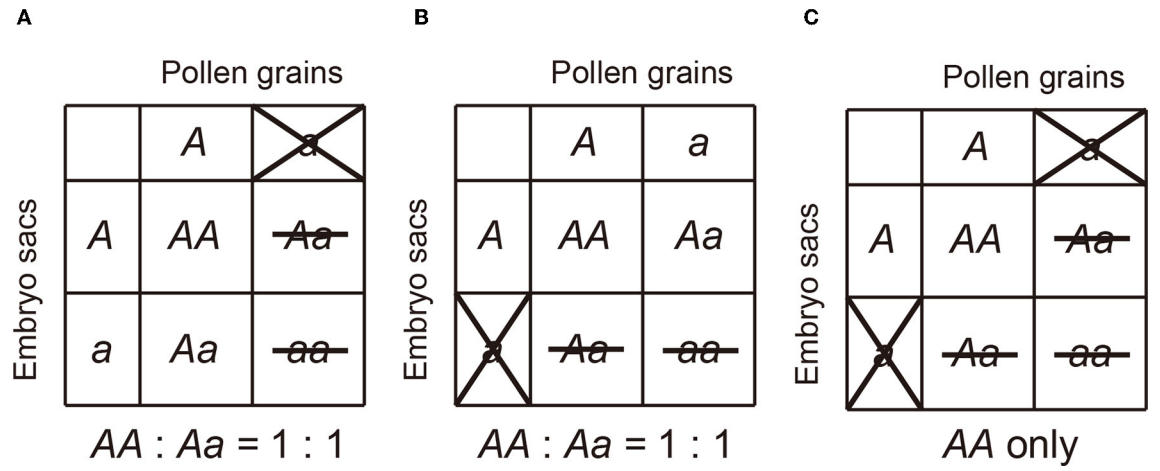

FIGURE 1 | The genetic actions of three categories of hybrid sterility in rice. **(A)** Pollen killer, **(B)** Egg killer, **(C)** Gamete eliminator.

derives from the indica, selectively protects the embryo sac in which it resides. In the absence of stress responsive gene *ORF3+*, the unsolved endoplasmic reticulum (ER) stress induced by *ORF5+* in partner with *ORF4+*, triggers programmed cell death (PCD), which results in embryo sac abortion (Yang et al., 2012).

The *S1* locus is one of the major reproductive barriers between Asian and African rice species and a remarkable *si*TRD locus. The *O. glaberrima* allele ($S1^g$) takes advantage over *O. sativa* allele ($S1^s$) within populations, thus preferentially transmitting $S1^g$ and eliminating both male and female gametes when two alleles meet in a heterozygous state ($S1^s/S1^g$) (Koide et al., 2008c, 2018). According to Xie et al. (2017), CRISPR/Cas9-generated knockout mutants of *OgTPR1* on $S1^g$ region (which encodes a protein with two trypsin-like peptidase domains and one ribosome biogenesis regulatory protein domain) produce normally fertile pollen and embryo sacs in crossing to *O. sativa* parent. The tripartite gamete Killer-Protector complex involving *S1A4*, *S1TPR* (*OgTPR1*), and *S1A6* (*SSP*) of the $S1^g$ region, generates a sterility signal in sporophytic cells and protect itself with *S1TPR* in gametophytic cells, whereas its allelic $S1^s$ region lacks both functional killer and protector (Xie et al., 2019b).

The evolution of Killer-Protector system should be explained by models in which the deleterious effect of the Killer did not occur in a lineage (e.g., BDM model or parallel-sequential divergence model, Ouyang and Zhang, 2013), because without Protector, Killer induces abortion of gametes with both alleles highly reducing the chance of its fixation in a population. In the study of *S5* sterility locus, the combination of non-functional Protector, and functional Killer and its partner (*ORF3–*, *ORF4+*, and *ORF5+*) was not found in the surveyed population (Yang et al., 2012). It may suggest that the combination that has Killer alone, without having functional Protector, cannot survive long in a population and also suggest that Killer can exist with functional Protector. However, after the emergence of Killer-Protector system, a single hybrid sterility locus with two tightly linked genes produces incompatibility when it meets the divergent allele in a different population. Then, the selfish nature of the system (i.e., TRD) caused by Killer-Protector system may facilitate its spread in a population.

DISCUSSION

Evolution of Sex-Specificity in TRD in Rice

Approximately 50 hybrid sterility loci were identified in diverged species from the genus *Oryza*, and so far, cloning and characterization at a molecular level were already implemented for 10 loci out of those sterility loci or pairs (Ouyang and Zhang, 2013; Li et al., 2020). Among 49 hybrid sterility loci surveyed in this review, the number of loci for *m*TRD is the largest, in contrast to that of *fTRD* and *siTRD* (**Table 1**). In addition, the sterility based on gamete specificity disproportionately distributes between interspecific and intraspecific hybrids (**Table 1**). The distribution of *m*TRD loci is much wider in interspecific hybrids. On the other hand, all the *f* TRD loci are detected exclusively in intraspecific crosses. With regards to *siTRD*, it takes <15% of the total number of hybrid sterility loci, and it is very few in proportions compared to the other two sterility factors, *m*TRD and *f* TRD. Such a bias in frequency of sex-specificity in TRD system observed in *Oryza* might reflect different evolutionary pressures acting on the system, suggesting that TRD systems are not driven only by mutation and genetic drift. Therefore, uncovering the mechanisms underlying the observed pattern of TRD systems may be a key to understanding the evolution of reproductive isolation in *Oryza*. Although there is no direct evidence, several conjectures are possible to explain these patterns.

The Abundance of *m*TRD

The simple difference between male and female gametes is the quantity of them produced by a single plant. In a single flower of rice, thousands of pollen grains are produced, while only one egg is produced. Such a difference in the quantity of gametes can cause differences in selection pressure acting on factors which

TABLE 1 | Hybrid sterility loci in rice and their sex-specificity-related transmission ratio distortion.

Sex-specific transmission ratio distortion (TRD)		Loci	References
*m*TRD	Interspecific cross	*S3*	Sano, 1983
		S12(*t*)	Sano, 1994
		S13	Koide et al., 2007
		S18	Doi et al., 1998
		S19	Taguchi et al., 1999
		S20	Doi et al., 1999
		S21	Miyazaki et al., 2007
		S22A and *S22B*	Sakata et al., 2021
		S23(*t*)	Fang et al., 2019
		S27	Yamagata et al., 2010
		S28	Yamagata et al., 2010
		S29(*t*)	Hu et al., 2006
		S34(*t*)	Zhang et al., 2005
		S36	Win et al., 2009
		S38	Xu et al., 2014
		S39	Xu et al., 2014
		S44	Zhao et al., 2012
		S51	Li et al., 2018
		S52	Li et al., 2018
		S53	Li et al., 2018
		S54	Li et al., 2018
		S55/qHMS7	Li et al., 2018
			Yu et al., 2018
		S56	Zhang et al., 2018
	Intrapecific cross	*S14(t)*	Sano, 1994
		S24(*t*)	Kubo et al., 2000
		S25(*t*)	Kubo et al., 2001
		S35	Kubo et al., 2016
		Sa	Long et al., 2008
		Sb	Li et al., 2006
		Sc	Shen et al., 2017
		Sd	Guiquan et al., 1994
		Se	Guiquan et al., 1994
		Sf	Guiquan et al., 1994
*f*TRD	Intraspecific cross	*S5*	Yang et al., 2012
		S-7	Yu et al., 2016
		S-8	Wan et al., 1993
		S-9	Wan et al., 1996
		S15	Wan et al., 1996
		S16	Wan and Ikehashi, 1995
		S32(*t*)	Li et al., 2005
		S29(t)	Zhu et al., 2005b
		S30(t)	Zhu et al., 2005a
*si*TRD	Interspecific cross	*S1*	Xie et al., 2019b
		S2	Zin Mar et al., 2021
		S6	Koide et al., 2008a
		S33(*t*)	Ren et al., 2005
		S37	Xu et al., 2014
	Intraspecific cross	*S10*	Sano et al., 1994
		S11(*t*)	Sawamura and Sano, 1996

induce the abortion of them. One can easily imagine that the reduction of pollen numbers causes less effect on fecundity than that of egg numbers. The PK, which is one factor causing *m*TRD, induces abortion of only pollen grains of a specific genotype. Therefore, *m*TRD caused by a PK can spread in a population with a small selective disadvantage. In contrast, female-specific TRD (*f*TRD) and sex-independent TRD (*si*TRD) induced by EK and GE, respectively, cause abortion of eggs. Thus, *f*TRD and *si*TRD offers a selective disadvantage on the individual. The abundance of male sterility in hybrids is also found in *Drosophila* (Presgraves and Meiklejohn, 2021). A much smaller selective disadvantage of male sterility genes might help them exist within diverse populations at a higher pace of frequency. However, we also note that other experimental factors might be able to explain the abundance of *m*TRD reported: pollen sterility is easier to detect/analyze than egg sterility because a large number of pollen grains are available for the assay.

Uncommonly Observed Sex-Independent TRD

Among the three types of TRD systems, *si*TRD is less frequently observed (**Table 1**). In addition, most of them are observed in interspecific cross combinations. These observations suggest that hybrid sterility genes causing both male and female sterility occur less frequently and more time is necessary for fixation than ones causing single sex-specific sterility.

Although the *S1* locus for *si*TRD has been cloned, molecular mechanisms causing sex-independent sterility in pollen grains and eggs are unknown. Because pollen grains and eggs are developed in physically separated tissues (i.e., anthers and ovaries), it is difficult to imagine that abortion of gametes in one sex causes gametic abortion of another sex. The sex-independent abortion of gametes might be caused by disturbance of the biological/developmental process common in two sexes. Therefore, the rareness of *si*TRD might reflect less abundancy of the biological/developmental process common in two sexes than in one sex.

Another possible mechanism for the emergence of *si*TRD is a combination of factors for *m*TRD and *f*TRD. If the two genes, each of which causes *mTRD* and *fTRD*, respectively, are located in tightly linked regions on a chromosome, the region is expected to behave like a factor for *si*TRD. Because of the limited number of cloned loci for *si*TRD, it is still unknown how often such a "pseudo-*si*TRD" occurrs. In the case of the *S1* locus, which causes *si*TRD in inter-specific hybrids, Koide et al. (2008c) reported the change of sex-specificity of TRD depending on the length of introgressed chromosomal segments. The line with a long introgressed segment on chromosome 6 from *O. glaberrima* in the genetic background of *O. sativa* causes both pollen and embryo sac abortion when crossed with *O. sativa*. As a result, *si*TRD is observed in the next generation of hybrids. In contrast, when the line with a short introgressed segment from *O. glaberrima* was crossed with *O. sativa*, only pollen abortion and *m*TRD were observed. These results suggested the presence of two linked factors responsible for *m*TRD and *f*TRD in the region (We note that no other research groups have reported the change of sex-specificity of

TRD induced by the *S1* locus.). Another locus, the *S6* for *si*TRD, has been suggested to be a compound locus of *f*TRD and pollen competition (Koide et al., 2012). In the *si*TRD induced by the *S6* locus, preferential abortion was observed in ovules, but not in pollen grains, suggesting that *m*TRD was caused by competition of pollens with different genotypes. The degree of TRD was altered only for male gametes when genetic background was changed. These results suggested the presence of two different genes for *m*TRD and *f*TRD in a closely linked region, though no direct evidence of these two factors were reported. As we described above, the gene for *m*TRD may be easier to evolve than that for *f*TRD, because of its small selective disadvantage on fecundity. Therefore, if the *si*TRD system originated *via* tight linkage between *m*TRD and *f*TRD, the rate of its emergence is dependent on how often *f*TRD evolved in the chromosomal region closely linked to the *m*TRD. It should also depend on the strength of recombination between these two factors in the chromosomal region.

Although the common genetic basis of *si*TRD (i.e., one factor or two factors) is still unknown, the evolution of *si*TRD is dependent on the balance between transmission advantage through pollen and disadvantage of female gamete abortion. Therefore, population size and outcrossing rate also affect the evolutionary process of *si*TRD. Uncovering the molecular basis and evolutionary trajectories of *si*TRD system will provide clearer insight into how selfish elements relate to the development of species' barriers as theorized by Frank (1991).

CONCLUSIONS

In *Oryza* genus, reproductive isolation is excessively influenced by PK rather than EK and GE, which results in preferential occurrence of *m*TRD in contrast to the other two types, *f*TRD and *si*TRD. Compared to single sex-specificity, factors controlling *si*TRD are less frequently observed, mostly in interspecific hybridizations. Unveiling the underlying cause(s) behind this disproportionate pattern of TRD systems will shed light on the evolutionary process of reproductive barriers between rice relatives. Since our understanding on TRD systems remains very limited with confined experimental factors, further efforts are required to extend our investigation on many other selfish genes that exist and their distribution in *Oryza* genus.

AUTHOR CONTRIBUTIONS

Zin Mar Myint and YK contributed to conception, analyzed, and wrote the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

YK was funded by JSPS KAKENHI Grant Number 21H02160.

ACKNOWLEDGMENTS

We thank Prof. Y. Kishima for his support for the project.

REFERENCES

Ammiraju, J. S. S., Song, X., Luo, M., Sisneros, N., Angelova, A., Kudrna, D., et al. (2010). The *Oryza* BAC resource: a genus-wide and genome scale tool for exploring rice genome evolution and leveraging useful genetic diversity from wild relatives. *Breed. Sci.* 60, 536–543. doi: 10.1270/jsbbs.60.536

Bateson, W. (1909). "Heredity and variation in modern lights," in *Darwin and Modern Science*, ed A. C. Seward (Cambridge: Cambridge University Press), 85–101.

Birchler, J. A., Dawe, R. K., and Doebley, J. F. (2003). Marcus Rhoades, preferential segregation and meiotic drive. *Genetics* 164, 835–841. doi: 10.1093/genetics/164.3.835

Chen, C., E, Z., and Lin, H.-X. (2016). Evolution and molecular control of hybrid incompatibility in plants. *Front. Plant Sci.* 7, 01208. doi: 10.3389/fpls.2016.01208

Chen, Z., Hu, F., Xu, P., Li, J., Deng, X., Zhou, J., et al. (2009). QTL analysis for hybrid sterility and plant height in interspecific populations derived from a wild rice relative, *Oryza longistaminata*. *Breed. Sci.* 59, 441–445. doi: 10.1270/jsbbs.59.441

Coyne, J. A., and Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.

Diaz, A., and Macnair, M. R. (1999). Pollen tube competition as a mechanism of prezygotic reproductive isolation between *Mimulus nasutus* and its presumed progenitor *M. guttatus*. *New Phytol.* 144, 471–478. doi: 10.1046/j.1469-8137.1999.00543.x

Dobzhansky, T. (1937). Genetic nature of species differences. *Am. Nat.* 71, 404–420. doi: 10.1086/280726

Doi, K., Taguchi, K., and Yoshimura, A. (1999). RFLP mapping of *S20* and *S21* for F1 pollen semi-sterility found in backcross progeny of *Oryza sativa* and *O. glaberrima*. *Rice Genet. Newsl.* 16, 65–68.

Doi, K., Tagushi, K., and Yoshimura, A. (1998). A new locus affecting high F1 pollen sterility found in backcross progenies of *japonica* rice and African rice. *Rice Genet. Newsl.* 15, 146–148.

Fang, C., Li, L., He, R., Wang, D., Wang, M., Hu, Q., et al. (2019). Identification of *S23* causing both interspecific hybrid male sterility and environment-conditioned male sterility in rice. *Rice* 12, 10. doi: 10.1186/s12284-019-0271-4

Fishman, L., and Willis, J. H. (2005). A novel meiotic drive locus almost completely distorts segregation in *Mimulus* (Monkeyflower) hybrids. *Genetics* 169, 347–353. doi: 10.1534/genetics.104.032789

Frank, S. A. (1991). Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. *Evolution.* 45, 262–267. doi: 10.1111/j.1558-5646.1991.tb04401.x

Guiquan, Z., Yonggen, L., Hua, Z., Jinchang, Y., and Guifu, L. (1994). Genetic studies on the hybrid sterility in culitivated rice (*Oryza sativa*) IV: genotypes for F1 pollen sterility. *Acta Genet. Sin.* 21, 34–41.

Hu, F., Xu, P., Deng, X., Zhou, J., Li, J., and Tao, D. (2006). Molecular mapping of a pollen killer gene *S29*(t) in *Oryza glaberrima* and colinear analysis with *S22* in *O. glumaepatula*. *Euphytica* 151, 273–278. doi: 10.1007/s10681-006-9146-z

Koide, Y., Ikegaya, T., Nishimoto, D., Kanazawa, A., and Sano, Y. (2007). Hybrid male sterility induced by *S13* found in a distantly related rice species. *Genes Genet. Syst.* 82, 522–522.

Koide, Y., Ikenaga, M., Sawamura, N., Nishimoto, D., Matsubara, K., Onishi, K., et al. (2008a). The evolution of sex-independent transmission ratio distortion involving multiple allelic interactions at a single locus in rice. *Genetics* 180, 409–420. doi: 10.1534/genetics.108.090126

Koide, Y., Ogino, A., Yoshikawa, T., Kitashima, Y., Saito, N., Kanaoka, Y., et al. (2018). Lineage-specific gene acquisition or loss is involved in interspecific hybrid sterility in rice. *Proc. Nat. Acad. Sci. U. S. A.* 115, E1955–E1962. doi: 10.1073/pnas.1711656115

Koide, Y., Onishi, K., Kanazawa, A., and Sano, Y. (2008b). "Genetics of speciation in rice," in *Rice Biology in the Genomics Era. Biotechnology in Agriculture and*

Forestry, eds H. Y. Hirano, Y. Sano, A. Hirai, and T. Sasaki (Berlin, Heidelberg: Springer), 247–259.

Koide, Y., Onishi, K., Nishimoto, D., Baruah, A. R., Kanazawa, A., and Sano, Y. (2008c). Sex-independent transmission ratio distortion system responsible for reproductive barriers between Asian and African rice species. *New Phytol.* 179, 888–900. doi: 10.1111/j.1469-8137.2008.02490.x

Koide, Y., Shinya, Y., Ikenaga, M., Sawamura, N., Matsubara, K., Onishi, K., et al. (2012). Complex genetic nature of sex-independent transmission ratio distortion in Asian rice species: the involvement of unlinked modifiers and sex-specific mechanisms. *Heredity.* 108, 242–247. doi: 10.1038/hdy.2011.64

Kubo, T., Eguchi, M., and Yoshimura, A. (2000). A new gene for F1 pollen sterility in *japonica/indica* cross of rice. *Rice Genet. Newsl.* 17, 63–64.

Kubo, T., Eguchi, M., and Yoshimura, A. (2001). A new gene for F1 pollen sterility located on chromosome 12 in *japonica/indica* cross of rice. *Rice Genet. Newsl.* 18, 54.

Kubo, T., Yamagata, Y., Eguchi, M., and Yoshimura, A. (2008). A novel epistatic interaction at two loci causing hybrid male sterility in an inter-subspecific cross of rice (*Oryza sativa* L.). *Genes Genet. Syst.* 83, 443–453. doi: 10.1266/ggs.83.443

Kubo, T., Yoshimura, A., and Kurata, N. (2016). Pollen killer gene *S35* function requires interaction with an activator that maps close to S24, another pollen killer gene in rice. *G3.* 6, 1459–1468. doi: 10.1534/g3.116.027573

Li, D., Zhu, S., Jiang, L., and Wan, J. (2005). Mapping for a new locus causing hybrid sterility in a China landrace (*Oryza sativa* L.). *Rice Genet. Newsl.* 22, 17–19.

Li, J., Zhou, J., Xu, P., Deng, X., Deng, W., Zhang, Y., et al. (2018). Mapping five novel interspecific hybrid sterile loci between *Oryza sativa* and *Oryza meridionalis*. *Breed. Sci.* 68, 516–523. doi: 10.1270/jsbbs.18001

Li, J., Zhou, J., Zhang, Y., Yang, Y., Pu, Q., and Tao, D. (2020). New insights into the nature of interspecific hybrid sterility in rice. *Front. Plant Sci.* 11, 555572. doi: 10.3389/fpls.2020.555572

Li, W., Zeng, R., Zhang, Z., Xiaohua, D., and Guiquan, Z. (2006). Fine mapping of locus *S-b* for F1 pollen sterility in rice (*Oryza sativa* L.). *Chin. Sci. Bull.* 51, 675–680. doi: 10.1007/s11434-006-0675-6

Long, Y., Zhao, L., Niu, B., Su,J., Wu, H., Chen, Y., et al. (2008). Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes. *Proc. Natl. Acad. Sci. U. S. A.* 105, 18871–18876. doi: 10.1073/pnas.0810108105

Lyttle, T. W. (1991). Segregation distorters. *Annu. Rev. Genet.* 25, 511–581. doi: 10.1146/annurev.ge.25.120191.002455

Maguire, M. P. (1963). High transmission frequency of a tripsacum chromosome in corn. *Genetics* 48, 1185–1194. doi: 10.1093/genetics/48.9.1185

Mayr, E. (1996). What is a species, and what is not? *Philos. Sci.* 63, 262–277. doi: 10.1086/289912

Miyazaki, Y., Doi, K., Yasui, H., and Yoshimura, A. (2007). Identification of a new allele of F1 pollen sterility gene, *S21*, detected from the hybrid between *Oryza sativa* and *O. rufipogon*. *Rice Genet. Newsl.* 23, 36–38.

Moyle, L. C. (2006). Genome-wide associations between hybrid sterility QTL and marker transmission ratio distortion. *Mol. Biol. Evol.* 23, 973–980. doi: 10.1093/molbev/msj112

Muller, H. J. (1942). Isolating mechanisms, evolution and temperature. *Biol. Symp.* 6, 71–125.

Nosil, P. (2013). "Reproductive isolation," in *Brenner's Encyclopedia of Genetics,* eds S. Molley and K. Hughes (San Diego, CA: Academic Press), 173–176.

Oka, H.-I. (1957). Genic analysis for the sterility of hybrids between distantly related varieties of cultivated rice. *J. Genet.* 55, 397–409. doi: 10.1007/BF02984059

Ouyang, Y., and Zhang, Q. (2013). Understanding reproductive isolation based on the rice model. *Annu. Rev. Plant Biol.* 64, 111–135. doi: 10.1146/annurev-arplant-050312-120205

Pardo-Manuel de Villena, F., and Sapienza, C. (2001). Nonrandom segregation during meiosis: the unfairness of females. *Mamm. Genome* 12, 331–339. doi: 10.1007/s003350040003

Presgraves, D. C., and Meiklejohn, C. D. (2021). Hybrid sterility, genetic conflict and complex speciation: lessons from the *Drosophila simulans* clade species. *Front. Genet.* 12, 669045. doi: 10.3389/fgene.2021.669045

Price, C. S. C. (1997). Conspecific sperm precedence in *Drosophila*. *Nature* 388, 663–666. doi: 10.1038/41753

Ren, G., Xu, P., Deng, X., Zhou, J., Hu, F., Li, J., et al. (2005). A new gamete eliminator from *Oryza glaberrima*. *Rice Genet. Newsl.* 22, 45–47.

Sakata, M., Takano-Kai, N., Miyazaki, Y., Kanamori, H., Wu, J., Matsumoto, T., et al. (2021). Domain unknown function DUF1668-containing genes in multiple lineages are responsible for F1 pollen sterility in rice. *Front. Plant Sci.* 11, 632420. doi: 10.3389/fpls.2020.632420

Sano, Y. (1983). A new gene controlling sterility in F1 hybrids of two cultivated rice species. *J. Hered.* 74, 435–439. doi: 10.1093/oxfordjournals.jhered.a109832

Sano, Y. (1994). Pollen-killers in rice. *Jpn. J. Breed.* 44, 298.

Sano, Y., Chu, Y.-E., and Oka, H.-I. (1979). Genetic studies of speciation in cultivated rice. 1. Genic analysis for the F1 sterility between *O. sativa* L. and *O. glaberrima* Steud. *Jpn. J. Genet.* 54, 121–132. doi: 10.1266/jjg.54.121

Sano, Y., Sano, R., Eiguchi, M., and Hirano, H. (1994). Gamete eliminator adjacent to the *wx* locus as recealed by pollen analysis in rice. *J. Hered.* 85, 310–312. doi: 10.1093/oxfordjournals.jhered.a111465

Sawamura, N., and Sano, Y. (1996). Chromosomal location of gamete eliminator, *S11*(t), found in an *indica-japonica* hybrid. *Rice Genet Newslett.* 13, 70–71.

Seymour, D. K., Chae, E., Arioz, B. I., Koenig, D., and Weigel, D. (2019). Transmission ratio distortion is frequent in *Arabidopsis thaliana* controlled crosses. *Heredity* 122, 294–304. doi: 10.1038/s41437-018-0107-9

Shen, R., Wang, L., Liu, X., Wu, J., Jin, W., Zhao, X., et al. (2017). Genomic structural variation-mediated allelic suppression causes hybrid male sterility in rice. *Nat. Commun.* 8, 1–10. doi: 10.1038/s41467-017-01400-y

Silver, L. M. (1993). The peculiar journey of a selfish chromosome: mouse t haplotypes and meiotic drive. *Trends Genet.* 9, 250–254. doi: 10.1016/0168-9525(93)90090-5

Taguchi, K., Doi, K., and Yoshimura, A. (1999). RFLP mapping of *S19*, a gene for F1 pollen semi-sterility found in backcross progeny of *Oryza sativa* and *O. glaberrima*. *Rice Genet. Newsl.* 16, 70–71.

Turelli, M. (1998). The causes of Haldane's rule. *Science.* 282, 889–891. doi: 10.1126/science.282.5390.889

Ubeda, F., and Haig, D. (2005). On the evolutionary stability of Mendelian segregation. *Genetics* 170, 1345–1357. doi: 10.1534/genetics.104.036889

Wan, J., and Ikehashi, H. (1995). Identification of a new locus S-16 causing hybrid sterility in native rice varieties (*Oryza sativa* L.) from Tai hu Lake region and Yunnan province, China. *Jpn. J. Breed.* 45, 461–470. doi: 10.1270/jsbbs1951.45.461

Wan, J., Yamaguchi, Y., Kato, H., and Ikehashi, H. (1996). Two new loci for hybrid sterility in cultivated rice (*Oryza sativa* L.). *Theoret. Appl. Genet.* 92, 183–190. doi: 10.1007/BF00223375

Wan, J., Yanagihara, S., Kato, H., and Ikehashi, H. (1993). Multiple alleles at a new locus causing hybrid sterility between a Korean indica variety and a javanica variety in rice (*Oryza sativa* L.). *Jpn. J. Breed.* 43, 507–516. doi: 10.1270/jsbbs1951.43.507

Win, K., Kubo, T., Miyazaki, Y., Doi, K., Yamagata, Y., and Yoshimura, A. (2009). Identififfication of two loci causing F1 pollen sterility in inter- and intraspecifific crosses of rice. *Breed. Sci.* 59, 411–418 doi: 10.1270/jsbbs.59.411

Xie, Y., Shen, R., Chen, L., and Liu, Y.-G. (2019a). Molecular mechanisms of hybrid sterility in rice. *Sci. China Life Sci.* 62, 737–743. doi: 10.1007/s11427-019-9531-7

Xie, Y., Tang, J., Xie, X., Li, X., Huang, J., Fei, Y., et al. (2019b). An asymmetric allelic interaction drives allele transmission bias in interspecific rice hybrids. *Nat. Commun.* 10, 2501. doi: 10.1038/s41467-019-10488-3

Xie, Y., Xu, P., Huang, J., Ma, S., Xie, X., Tao, D., et al. (2017). Interspecific hybrid sterility in rice is mediated by OgTPR1 at the S1 locus encoding a peptidase-like protein. *Mol. Plant.* 10, 1137–1140. doi: 10.1016/j.molp.2017.05.005

Xu, P., Zhou, J., Li, J., Hu, F., Deng, X., Feng, S., et al. (2014). Mapping three new interspecific hybrid sterile loci between *Oryza sativa* and *O. glaberrima*. *Breed. Sci.* 63, 476–482. doi: 10.1270/jsbbs.63.476

Yamagata, Y., Yamamoto, E., Aya, K., Win, K., Doi, K., Sobrizal Ito, T., et al. (2010). Mitochondrial gene in the nuclear genome induces reproductive barrier in rice. *Proc. Natl. Acad. Sci. U. S. A.* 107, 1494–1499. doi: 10.1073/pnas.0908283107

Yang, J., Zhao, X., Cheng, K., Du, H., Ouyang, Y., Chen, J., et al. (2012). A Killer-Protector system regulates both hybrid sterility and segregation distortion in rice. *Science* 337, 1336–1340. doi: 10.1126/science.1223702

Yang, Y., Zhou, J., Li, J., Xu, P., Zhang, Y., and Tao, D. (2016). Mapping QTLs for hybrid sterility in three AA genome wild species of *Oryza*. *Breed. Sci.* 66, 367–371. doi: 10.1270/jsbbs.15048

Yu, X., Zhao, Z., Zheng, X., Zhou, J., Kong, W., Wang, P., et al. (2018). A selfish genetic element confers non-Mendelian inheritance in rice. *Science* 360, 1130–1132. doi: 10.1126/science.aar4279

Yu, Y., Zhao, Z., Shi, Y., Tian, H., Liu, L., Bian, X., et al. (2016). Hybrid sterility in rice (*Oryza sativa* L.) involves the tetratricopeptide repeat domain containing protein. *Genetics* 203, 1439–1451. doi: 10.1534/genetics.115.183848

Zhang, Y., Zhou, J., Li, J., Yang, Y., Xu, P., and Tao, D. (2018). Mapping of *S56*(t) responsible for interspecific hybrid sterility between *Oryza sativa* and *Oryza glumaepatula*. *Breed. Sci.* 68, 242–247. doi: 10.1270/jsbbs.17116

Zhang, Z., Xu, P., Hu, F., Zhou, J., Li, J., Deng, X., et al. (2005). A new sterile gene from *Oryza glaberrima* on chromosome 3. *Rice Genet. Newsl.* 22, 26–28.

Zhao, J., Li, J., Xu, P., Zhou, J., Hu, F., Deng, X., et al. (2012). A new gene controlling hybrid sterility between *Oryza sativa* and *Oryza longistaminata*. *Euphytica* 187, 339–344. doi: 10.1007/s10681-012-0691-3

Zhu, S., Jiang, L., Wang, C., Zhai, H., Li, D., and Wan, J. (2005a). The origin of weedy rice Ludao in China deduced by genome wide analysis of its hybrid sterility genes. *Breed. Sci.* 55, 409–414. doi: 10.1270/jsbbs.55.409

Zhu, S., Wang, C., Zheng, T., Zhao, Z., Ikehashi, H., and Wan, J. (2005b). A new gene located on chromosome 2 causing hybrid sterility in a remote cross of rice. *Plant Breed.* 124, 440–445. doi: 10.1111/j.1439-0523.2005.01143.x

Zin Mar, M., Koide, Y., Ogata, M., Kuniyoshi, D., Tokuyama, Y., Hikichi, K., et al. (2021). Genetic mapping of the gamete eliminator locus, *S2*, causing hybrid sterility and transmission ratio distortion found between *Oryza sativa* and *Oryza glaberrima* cross combination. *Agriculture* 11, 268. doi: 10.3390/agriculture11030268

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership