https://doi.org/10.17221/49/2022-CJGPB

Further fine mapping and candidate gene prediction for a new restoring fertility gene *Rf(fa)* in rice

Yu Li[1], Xiaomi Chen[1], Tao Lan[1], Jing Zhang[2], Ziheng Chen[2], Wenting Yang[3], Xinmei Lin[3]*

[1]*Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops, Fujian Agriculture and Forestry University, Fuzhou, P.R. China*
[2]*College of Agronomy, Fujian Agriculture and Forestry University, Fuzhou, P.R. China*
[3]*Jinshan College, Fujian Agriculture and Forestry University, Fuzhou, P.R. China*

**Corresponding author: yl0934@163.com*

Citation: Li Y., Chen X.M., Lan T., Zhang J., Chen Z.H., Yang W.T., Lin X.M. (2023): Further fine mapping and candidate gene prediction for a new restoring fertility gene *Rf(fa)* in rice. Czech J. Genet. Plant Breed., 59: 43–53.

Abstract: *Rf(fa)*, a new restoring fertility gene in rice, was previously located to a large region on Chromosome 10. The large number of genes within the region made cloning of *Rf(fa)* difficult. To perform the cloning and further elucidate the molecular mechanism, we reconstructed a mapping segregation population (BC_1F_1) of 12 000 plants. Using the population and polymorphism of simple sequence repeat (SSR) molecular markers, we finally mapped *Rf(fa)* between the two SSR molecular markers MM2000 and RM25658, within a 78.87 kb region. By *de novo* sequencing of a restoring line of CMS-FA hybrid rice, we obtained the genomic sequence of the mapping region, which provided the basis for the prediction of the candidate gene(s) of the target gene and for the comparison of genomic sequence differences between wild and cultivated rice. Within the mapping region, the genomic sequence of the wild rice was significantly different from that of cultivated rice. There were ten genes in the final mapping region. A pentatricopeptide repeat (PPR) protein gene was predicted as the candidate gene of *Rf(fa)*. Our results laid a solid foundation for the final cloning and molecular mechanism analysis of the gene. The identified molecular markers tightly linked to *Rf(fa)* will facilitate the marker assisted selection in breeding of CMS-FA hybrid rice.

Keywords: fertility restorer; gene fine mapping; prediction of candidate gene(s); rice

CMS-FA hybrid rice, a new type of three-line hybrid rice, was developed recently by using the sterile cytoplasm and the corresponding nuclear restorer gene of common wild rice (*O. rufipogon* L.) originating in Zhangzhou, Fujian Province, China. Both the cytoplasmic male sterility gene and nuclear fertility restoring gene are different from those of the other hybrid rice systems reported (Wang et al. 2008b). Therefore, *Rf(fa)* is a new restoring gene for three-line hybrid rice.

In 1966, Longping Yuan, "the father of hybrid rice" in China, first put forward the concept of heterosis utilization for self-pollination crop (Yuan 1966), thus opened the prelude of three-line hybrid rice research. In 1970, the research group led by Longping Yuan found a rice male sterile mutant plant among wild rices growing in Hainan, China. In 1973, the research team led by Longan Yan successfully developed Zhenshan 97A, the first male sterile line for three-line hybrid rice based on the mutant. Wild-abortion hybrid

Supported by the Nature Science Foundation of Fujian Province (2020J01527), the Innovation Foundation of Fujian Agriculture and Forestry University (KHF220003, KFA17419A, KFA18113A).

https://doi.org/10.17221/49/2022-CJGPB

rice (CMS-WA) system was established in 1973, and it has been applied in production since 1976. So far, besides WA hybrid rice, the other developed three-line hybrid rices include CMS-BT, CMS-HL, CMS-DA, CMS-LD, CMS-D, CMS-DT, CMS-CW, CMS-D1 and CMS-FA etc. (Virmaniss 1988; Li et al. 2016; Xie et al. 2018), the restoring genes involved include *Rf1*, *Rf2*, *Rf3*, *Rf4*, *Rf5*, *Rf6*, *Rf17*, *Rf98* etc., the molecular mechanisms for some of the genes have been elucidated preliminarily (Wang et al. 2006; Sota & Kinya 2009; Eitabashi et al.2011; Hua et al. 2012; Tomohiko & Kinya 2014; Huang et al. 2015; Li et al. 2016). *Rf1* is the restorer for CMS-BT, including *Rf1a* and *Rf1b*. *Rf1b* is near the upstream of *Rf1a*. *Rf1a* and *Rf5* are actually the same gene but are named *Rf5* in CMS-HL, located between two molecular markers RM1108 and RM5373 on Chromosome 10 of rice (Liu et al. 2004). However, the molecular mechanisms of *Rf1a* and *Rf5* differ: in BT system, *Rf1a* restored male fertility by being directly involved in the processing of the cytoplasmic male sterility gene *ORF79* transcript, cutting the transcript into two fragments of 1.5 kb and 0.45 kb (Wang et al. 2006); but in HL system, *Rf5* restores male fertility via a complex with the glycine-rich protein GRP162 (Hua et al. 2012). *Rf1a* and *Rf1b* restore the male fertility of BT hybrid rice by blocking the production of the sterile protein ORF79 via endonucleolytic cleavage (Rf1a) or degradation (Rf1b) of the double cistrons of *atp6-orfH79* mRNA; when the two protein genes exist simultaneously, *Rf1a* is epistatic over *Rf1b* in the processing (Wang et al. 2006). *Rf2*, located on Chromosome 2 of rice, is the restorer for LD hybrid rice, encoding a protein (with 152 amino acid residues and a glycine-rich domain) targeted to mitochondria. Rf2 protein contains only 80 amino acid residues excluding mitochondrial signal peptide sequence (Eitabashi et al. 2011), and needs to interact with the protein with a ubiquitin-binding domain to restore the male fertility of CMS-LD line because of the lack of a RNA-binding domain (Shinya et al. 2014). *Rf3* and *Rf4*, restorers for CMS-WA, located on Chromosome 1 and Chromosome 10 respectively, restore male fertility by regulating the expression of the sterile gene *WA352*: Rf3 restores male fertility by preventing the translation of *WA352* mRNA into proteins, whereas Rf4 by blocking the production of *WA352* mRNAs (Luo et al. 2013). However, Rf4 can only partially restore male fertility, complete fertility restoration requires the presence of both *Rf3* and *Rf4* genes (Ponnuswany et al. 2020). Recently, a new gene *Rf18(t)*, encoding hexokinase, is reported to be another fertility restorer for CMS-WA (Zhang et al. 2022), suggesting the complexity of the fertility restoration of CMS-WA hybrid rice. *Rf5* and *Rf6* are restorers for CMS-HL, but their molecular mechanisms are different: Rf5 restores fertility via a complex with the glycine-rich protein GRP162 as mentioned above, *Rf6* functions with hexokinase 6 to rescue cytoplasmic male sterility (Hua et al. 2012; Huang et al. 2015). *Rf17* from Chinese wild rice, encoding a mitochondria protein of unknown function, was mapped between two molecular markers SNP7-16 and SNP7-4; the down-regulation expression can restore the male fertility of CMS-CW male sterile lines while the normal expression will lead to the abortion of pollens (Fujii & Toriyama 2009). *Rf98*, located between two DNA markers SSRH10027 and IndelT003 on Chromosome 10, can partially restore the male fertility of RT98 type (Keisuke et al. 2016). Most of these *Rf* genes in rice belong to PPR protein genes, except for *Rf2*, *Rf17* and *Rf18(t)*. *Rf2* in maize, encoding aldehyde dehydrogenase, is also not a PPR protein gene (Cui et al. 1996). Thus, it can be seen that no two cloned *Rf* genes have shared completely the same fertility mechanism, neither have *Rf1a* and *Rf5*. Therefore, the mapping and cloning of *Rf(fa)* will certainly help to reveal a new fertility mechanism in hybrid rice.

CMS-FA hybrid rice has been utilized in production for ten years and has excellent comprehensive performance. Systematic studies related to the inheritance of *Rf(fa)* (Wang et al. 2008a), parent resource utilization rate (Wang et al. 2008b), yield-related traits and heterosis (Wang et al. 2010), genetic effect of nucleo-cytoplasmic interaction (Chiu et al. 2013) etc. have been carried out. However, the previous mapped range was still too large, based on which it is difficult to realize the cloning. To date, we still lack a comprehensive and deep understanding of the fertility mechanism of CMS hybrid rice due to too few restorers cloned across the world. The further fine mapping of *Rf(fa)* will lay a solid foundation for the ultimate cloning and molecular mechanism analysis of the gene.

MATERIAL AND METHODS

Plant material. Rice varieties Jinnong 1A, Jinnong 1B, 9311-9 were a male sterile line, a male sterile maintainer line and a restorer line for CMS-FA respectively, of which Jinnong 1A was developed

https://doi.org/10.17221/49/2022-CJGPB

by crossing the common wild rice (♀) with an *indica* rice cultivar Zaohui 89 (♂) to get F_1 seeds first, and then by continuous backcrossing with Zaohui 89 as the recurrent parent, hence its cytoplasm was the same as that of the common wild rice from Zhangzhou, Fujian. 9311-9 is an iso-cytoplasmic restorer line of Jinnong 1A, with the nuclear background of 9311, but carrying the target gene *Rf(fa)*.

Construction of mapping segregation population. First, hybrid F_1 was obtained from the cross Jinnong 1A (♀) × 9311-9 (♂), and then a backcross of Jinnong 1A (♀) × the F_1 (♂) was made to produce the BC_1F_1 mapping segregation population with the size of 12 000 plants. Theoretically, the fertile plants and sterile plants in the population would segregate at 1 : 1.

Field sampling and DNA extraction. The BC_1F_1 population was planted in the breeding base in Sanya, Hainan Province, China, in the winter of 2016. During the flowering and grain filling period in April 2017, typical male sterile plants were selected, and the fresh leaves were harvested for extraction of DNA. Sterile plants and fertile plants were determined mainly by their plant phenotypes or characteristics in the field. The genomic DNA was extracted by cetyltrimethyammonium bromide (CTAB) method (Murray & Thompson 1980).

Linkage genetic analysis. The linkage between the SSR marker and *Rf(fa)* was analyzed to further reduce the range of the previous mapping region. Primers of SSR molecular markers (Table 1) were designed according to the rice SSR molecular markers map data published by Zhejiang University, China (Zhang et al. 2007). First, we identified those SSR molecular markers that exhibit polymorphism within the previous mapping region between RM6100 and MM2023. Then, using 5 088 male sterile plants as genetic analysis materials, we identified crossing-over plants at RM25670 and RM25633, which were good in polymorphism, flanking the target gene and having clear bands. Subsequently, we used these crossing-over plants and the polymorphism markers to perform the linkage genetic analysis between

Table 1. Primers used in the study

Primer names	Forward primer (5'-3')	Reverse primer (5'-3')	Purpose
RM25633	ATGCTTTCTTGAGCCTCCAAACC	GTAACCCAAACTCCCAAAGTGTAGC	gene mapping
RM5629	CAACAACATCTGTGAGGGTTTCG	CACCACCATCTCCTCTTTCACC	gene mapping
MM1981	TTATACCCCGTCCGTCAAAA	TTACGTTTCATTGATCGCCA	gene mapping
MM1992	CCTCCGACGAACTGACTTTC	CACGCAGTAGGTCCACTCCT	gene mapping
RM6100	TTCCCTGCAAGATTCTAGCTACACC	TGTTCGTCGACCAAGAACTCAGG	gene mapping
MM2000	TCCGTTCCCAGATACTGGTC	CCCCTCAACGGGATATTTCT	gene mapping
MM2008	AGTGGGCAGCTAGGGTTTTT	GATGGAAGCAAAACCAGGAA	gene mapping
RM25658	ACGATCAGACGCCTAACTACAGC	CCCTCCTGCTAACTTGAGACG	gene mapping
MM2014	CGTAGCGGCGAAGGATATAG	TGATCTTTTTGTCAGCTGCG	gene mapping
MM2023	CCTTTTAGTGGAAATCCGCA	GGACGAAAACAGAGCTTGGA	gene mapping
MM2027	TGCTTCCTCCTGCAATTCTT	TTGCATCTGCATAAGCATCC	gene mapping
MM2043	CATTTGATCCGAAATCTCAGC	CAAGAAGGGAAGCAGGTCAG	gene mapping
RM25663	GCTTGAGCTCGTCGTGGAAGG	CTCCTGCGTTCTCCATCGAAGC	gene mapping
RM25670	CCATGAGCTCCCACTGGTTCC	CTCTGTCGTGTAGTTCTGGGTGTGG	gene mapping
Gene1RT	CAAGGAGGGAGCGGTGAAGGA	GGAGGGGTAGTCGGAGGAGGT	RT-PCR of Gene 1
Gene2RT	GAATGCAGGGAACACGAACAC	GGAACAGCGGGACGTGCTCCCGG	RT-PCR of Gene 2
Gene3RT	ATCTTCTGGTAGTGATGCTGGTC	CTCTCGTGTTATAGTTGTTGTCC	RT-PCR of Gene 3
Gene4RT	CAAATTGCTCATTCCTGTCCACC	ACGCACGCCTCGCCCCATCCCGT	RT-PCR of Gene 4
Gene5RT	AGCTCGAGAAGACAATTAAGTCA	GAAGAAGGAAAAGAAGAAAAAAC	RT-PCR of Gene 5
Gene6RT	GCCCTCAACATCCTTTCCCCTGC	ATTGCCCATTTCACCACCCGTCC	RT-PCR of Gene 6
Gene7RT	GCGGAGGGAGACTGTGGTGAG	AAGGAGGAGATGGCAGGGATG	RT-PCR of Gene 7
Gene8RT	CACGTCGCTGCCCCGCTGCCA	CCACTCCCTCCCCCTTTCTCC	RT-PCR of Gene 8
Gene9RT	GTGTCGTACAGCACCGTCATC	ACTCTTCAGCCTTCTCCCATT	RT-PCR of Gene 9
Gene10RT	CAGGTTAAAGAAGTTGGCAAA	ACCAGGACTAATAGATGAGGG	RT-PCR of Gene 10

https://doi.org/10.17221/49/2022-CJGPB

the target gene and each polymorphism SSR marker. Once a crossing-over plant was identified, the range where the target gene was would be reduced by some genetic distance; we continued the analysis until no crossing-over plant or no polymorphism marker were available.

PCR reaction system (20 µL) consisted of 10 µL Premix *Taq* DNA polymerase (TaKaRa, Japan), 1 µL template DNA (100 ng), 1 µL of forward and reverse primer (10 mM) each, ddH_2O 7 µL. PCR reaction procedure: 94 °C for 30 s, 35 cycles of 98 °C for 10 s, 51–58 °C for 30 s and 72 °C for 20 s, with a final extension at 72 °C for 2 min. The PCR products were separated by electrophoresis on 6% denaturing polyacrylamide gel and visualized by silver staining.

***De novo* sequencing and gene prediction.** Using *de novo* sequencing method, we sequenced the genome of a CMS-FA restoring line variety 9311-9 to obtain the genomic sequence of the previous mapped range between RM6100 and MM2023. Genes of the mapping region were predicted by using FGENESH online software (http://www.softberry.com).

RNA expression analysis. Total RNA was extracted from 9311-9 young panicle tissues using TaKaRa MiniBEST Plant RNA Extraction Kit (Dalian, China), and reverse transcription was performed using AMV first chain cDNA synthesis kit (Sangon Biotech, Shanghai, China). PCR primers (Table 1) were designed according to mRNA sequences of the predicted genes within the final mapping region. RT-PCR conditions used were as follows: 94 °C for 30 s, followed by 35 cycles of 98 °C for 10 s, 51–55 °C for 30 s, 72 °C for 30 s, and a final extension at 72 °C for 10 min. RT-PCR products were purified and recycled by agarose gel electrophoresis and using Agarose Gel DNA Extraction Kit (TaKaRa), and finally sequenced by TA-cloning method.

Informatics analysis. DNA sequence informatics analysis (blast) was performed by using NCBI on-line software (https://www.ncbi.nlm.nih.gov/ncbisearch/) and The EMBL-EBI on-line software (https://www.ebi.ac.uk/Tools/msa/clustalo/).

RESULTS AND DISCUSSION

Plant characteristics of BC_1F_1 population. The difference between male sterile plants and male fertile plants in the BC_1F_1 population was obvious in the field. When seeds matured, panicles of male fertile plants showed normal yellow, while those of male sterile plants remained green (Figure 1). During the grain filling period, male fertile plants displayed obvious characteristics of bearing seeds, while male sterile plants were characteristic of enclosed panicles, and were shorter without bearing seeds. All these features made it easy to distinguish male sterile plants from male fertile plants. In WA, BT and HL-type hybrid rice, there are two restoring fertility genes at least for each, one of restoring genes may not completely restore male fertility. Huang et al. (2012) reported that hybrid F_1 plants containing either of *Rf5* and *Rf6* displayed 50% fertile pollen grains, while those with both *Rf5* and *Rf6* showed 75% normal pollens. Rice variety Swarna with *Rf4* was identified to be a partial restorer, but it exhibited complete fertility restoration when another restorer *Rf3* was introduced into Swarna (Ponnuswamy et al. 2020). In comparison, CMS-FA rice has only one restorer *Rf(fa)*, which can completely restore male fertility (Figure 1).

***Rf(fa)* was mapped between MM2000 and RM25658.** To further fine map the target gene, *Rf(fa)*, we first conducted the polymorphism analysis of SSR molecular markers between RM6100 and MM2023, which located at both ends of the previous mapping region (Li et al. 2016), to obtain new polymorphism SSR molecular markers. As a result, three SSR markers were found to exhibit polymorphism between two parents Jinnong 1A and 9311-9, they were MM2014, RM25658 and MM2000 (Figure 2; Figure S1 in Electronic Supplementary Material (ESM)). Using the three markers we identified three crossing-over plants

Figure 1. Panicle feature of male sterile (left) and fertile (right) plants from the BC_1F_1 mapping segregation population

bar = 2 cm

https://doi.org/10.17221/49/2022-CJGPB

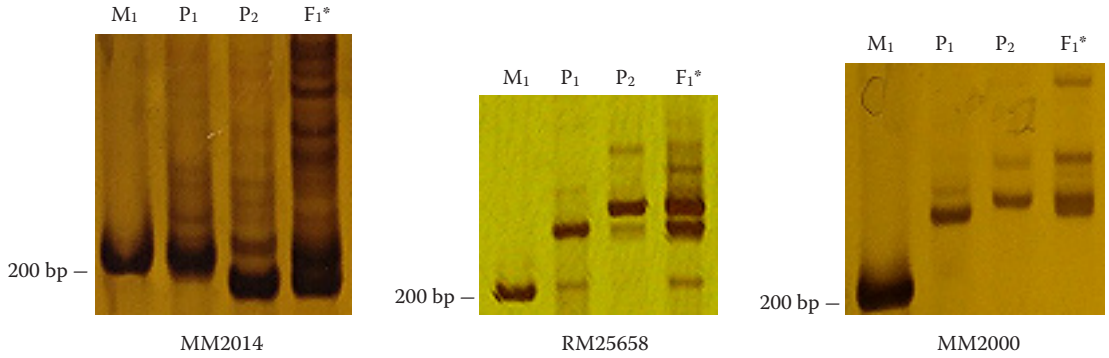

Figure 2. Polymorphism SSR markers identified within the previous mapping region between RM6100 and MM2023
M_1 – DL1000 DNA marker; P_1 – Jinnong 1A (♀); P_2 – 9311-9 (♂); F_1 – F_1 of Jinnong 1A × 9311-9; *F_1 and any crossing-over plant from the BC_1F_1 population shared the same band pattern, thus used as the control

(No. 538, 2579 and 4400) at MM2000 (Figure 3A; Figure S2 in ESM), two crossing-over plants (No. 2322 and 3276) at MM2014 (Figure 3B; Figure S2 in ESM) and one crossing-over plant (No. 2322) at RM25658 (Figure 3C; Figure S2 in ESM). Thus the target gene *Rf(fa)* was finally mapped between two SSR molecular markers MM2000 and RM25658 (Figure 4), a range of 78.87 kb based on the de novo sequencing result.

The polymorphism of MM2000 marker between the parents was extremely poor (Figure S1 in ESM; Figure 2; Figure 3A). To find out what the maximum resolution which could be achieved by the polyacrylamide gel electrophoresis was, we sequenced the marker band DNAs of Jinnong 1A and 9311-9 by TA cloning, the result (data unpublished) showed that the length for MM2000 marker band DNA of 9311-9 was 216 bp, while that of Jinnong 1A was 213 bp, thus the minimum difference between two genotypes which could be determined by polyacrylamide gel electrophoresis might be about 3 bp. Although SSR marker sequences are highly susceptible to mutation (Tautz 1989), and there are more than 50 000 SSR

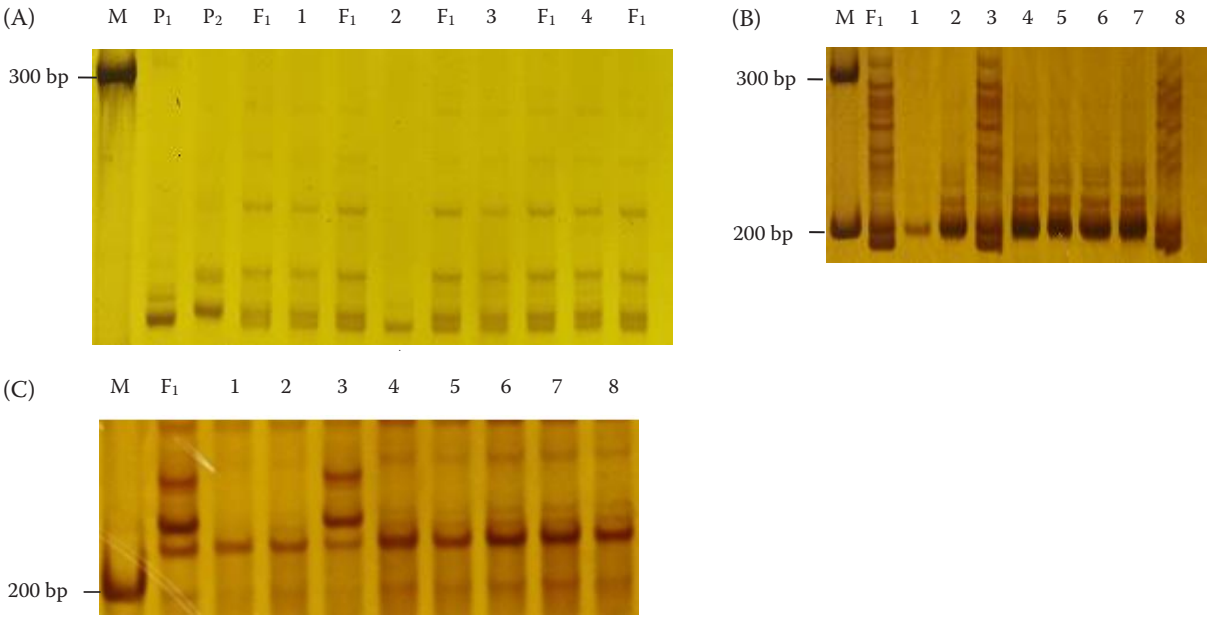

Figure 3. Plants displaying crossing-over bands at MM2000 (A), MM2014 (B) and RM25658 (C)
A: M – DL1000 DNA marker; P_1 – Jinnong 1A (♀); P_2 – 9311-9 (♂); 1, 3, 4 – recombinants (plant numbers: 538, 2 579, 4 400); 2 – non-crossing over plant; F_1 – F_1 of Jinnong 1A × 9311-9
B: M – DL1000 DNA marker; F_1 – F_1 of Jinnong 1A × 9311-9; 1, 2, 4~7 – non-crossing-over plants; 3, 8 – crossing-over plants
C: M – DL1000 DNA marker; F_1 – F_1 of Jinnong 1A × 9311-9; 1~2, 4~8 – non-crossing over plants, 3 – crossing-over plant

https://doi.org/10.17221/49/2022-CJGPB

molecular markers developed so far in rice (Zhang et al. 2007), still it seems insufficient for gene fine mapping sometimes. In the absence of polymorphism molecular markers available or the polymorphism is too poor to discriminate the difference, it might solve the problem to find out the sequence difference between two genotypes by the sequencing of the DNA from molecular marker bands.

Mapping segregation population use. Development and application of rice SSR molecular markers greatly facilitate the mapping and cloning of rice functional genes. The mapping segregation population size required for preliminary mapping may be only a few hundred individuals, while that needed for fine mapping may be much too much larger. In the previous study, Li et al. (2016) used a BC_1F_1 population of about 800 plants (380 male sterile plants used for genetic analysis) and an F_2 population of about 2 000 individuals (570 male sterile plants used for analysis) to perform the preliminary mapping. In this study, we used a re-constructed BC_1F_1 population of 12 000 plants (5 088 male sterile plants used for analysis) to locate the target gene to a 78.87 kb region, identified a crossing-over plant at RM25658 and three crossing-over plants at MM2000, which might have reached the limit the population could achieve. Studies have shown that the mapping population size required for an actual mapping varies, depending on the position of the gene on the chromosome and the mapping segregation population type used. Rice photosensitive nuclear sterility gene *pm3* was mapped to a 28.4 kb area using an F_2 population of 7 000 plants (Ding et al. 2012); Keisuke et al. (2016) used an F_2 population of 1 248 plants and a BC_1F_1 population of 6 432 plants to realize the fine mapping of *Rf98* gene; Zhang et al.(2002) used 117 completely sterile plants from an F_2 population to preliminarily map *Rf4* gene; the rice Zebra leaf mutant gene *zebra1349* was mapped to a region of 89 kb with a genetic distance of 0.21 cM using 1 192 recessive plants from a segregation population (Guo et al. 2016). Tan et al. (2015) located *Green-Revertible Chlorina* Gene *grc2* in rice between two molecular markers (31 kb) using 960 mutant individuals from an F_2 population. Thus the population size required for the mapping of different genes may vary greatly, while that for preliminary mapping may be only more than 100 plants.

Genomic sequence divergence between wild rice and cultivated rice. The length of the genomic sequence between RM6100 and MM2023 was 144.6 kb based on the de novo sequencing result, which spanned the whole mapping region determined by this study. Comparing this sequence with those of a few cultivated rice varieties, we found the following differences: (1) The sequence from the wild rice species was significantly longer than those of cultivated rice varieties. The corresponding sequences for Nipponbare, Minghui 63 and Zhenshan 97B were 120.8, 114.7 and 140.2 kb, namely 23.8, 29.9 and 4.4 kb shorter than that of the wild rice respectively. Obviously, the difference varied depending on different genetic backgrounds or genotypes. (2) Between RM6100 and MM2023, the gene number in Nip-

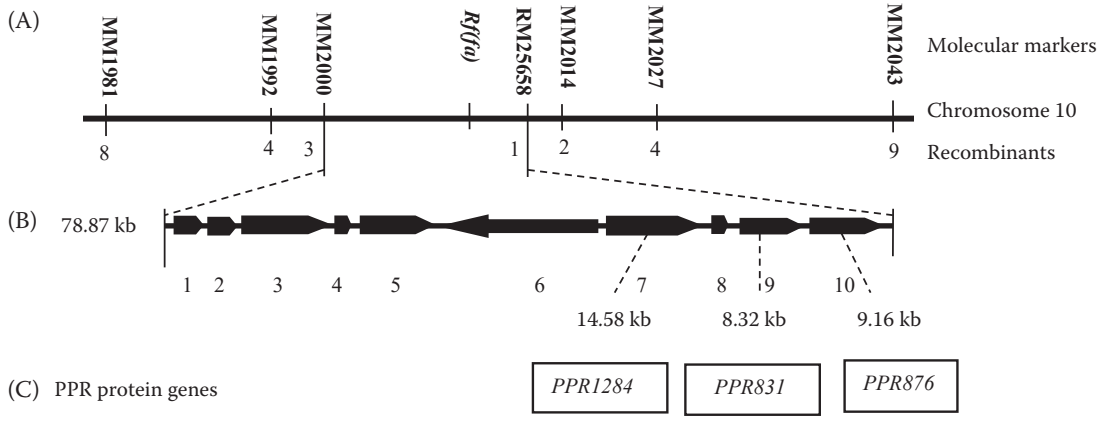

Figure 4. Further fine mapping of the target gene *Rf(fa)*: the position of the target gene *Rf(fa)* on Chromosome 10 (A), *Rf(fa)* was finally mapped between MM2000 and RM25658, a range of 78.87 kb, with 10 predicted genes (Gene 1~Gene 10) (B), Gene 7, Gene 9 and Gene 10 were PPR protein genes, and one of genes *PPR1284* was predicted as the candidate gene of *Rf(fa)* (C)

https://doi.org/10.17221/49/2022-CJGPB

Table 2. Genes and their functions of the final mapping region

Gene sequence No.	Genome size (bp)	Total exon length (bp)	Possible protein type	Reference (NCBI Sequence ID)
Gene 1	2 451	738	protochlorophyllide reductase B	XP-015614945.1
Gene 2	2 046	597	protochlorophyllide reductase B	XP-015614945.1
Gene 3	13 984	3 630	uncharacterized protein LOC4349005 isoform X2	XP-015614617.1
Gene 4	2 814	255	XS domain containing protein-like	BAD26426.1
Gene 5	3 846	2 844	retrotransposon protein	AAX96828.1
Gene 6	9 647	6 333	retrotransposon protein	ABB47352.2
Gene 7*	14 587	3 855	PPR protein; protein Rf1, mitochondrial isoform X1	XP-015614273.1; ABC42330.1
Gene 8	4 547	654	membrane-associated protein	AAL58263.1
Gene 9	8 322	2 496	PPR protein; protein Rf1	XP-015614275.1; BAD13709.1
Gene 10	9 164	2 631	PPR protein; membrane-associated protein	AAL58282.1; ABC42330.1

*the candidate gene of *Rf(fa)*: *PPR1284*

ponbare and the wild rice species was different. According to the gene prediction by Softberry Software (http://www. softberry.com), there were 28 genes based on the wild rice genome, and 23 genes based on Nipponbare genome. Within the final fine mapping region, there were only 9 genes for Nipponbare and 10 genes for the wild rice. (3) 9311-9 and 9311 are isogenic lines with similar nuclear background, but 9311-9 possesses the sequence introduced from the wild rice with *Rf(fa)*. The mapping region between MM2000 and RM25658 of 9311-9 was 25.8 kb longer than that of 9311. Compared with 9311-9, 9311 had discontinuous large-fragment deletions. Zhang et al. (2016) compared the genomic sequence differences between Minghui 63 and Zhenshan 97A, uncovering surprising structural differences resulting from inversions, translocations, presence/absence variations, and segmental duplications, suggesting that there exists extensive sequence differences between rice varieties.

Gene structure and function of the final mapping region. Within the mapping region (78.87 kb), all the predicted genes contained introns. Bioinformatics analysis showed that the protein sequences of the predicted genes could partially match the sequences of some functional proteins such as protochlorophyllide reductase B, XS domain containing protein-like, retrotransposon protein, PPR protein, membrane-associated protein etc. (Table 2; Table S1 in ESM).

Transcriptional level analysis of candidate genes. Using 10 pairs of RT-PCR primers (Table 1), we obtained 10 DNA fragments (Figure 5) corresponding to the 10 predicted genes. Through TA cloning and sequencing, we found that the full-length RT-PCR sequence of Gene 2 (287 bp) was identical with the mRNA sequence $_{(3\text{-}289)}$ of Gene 2 (Figure 6A); that

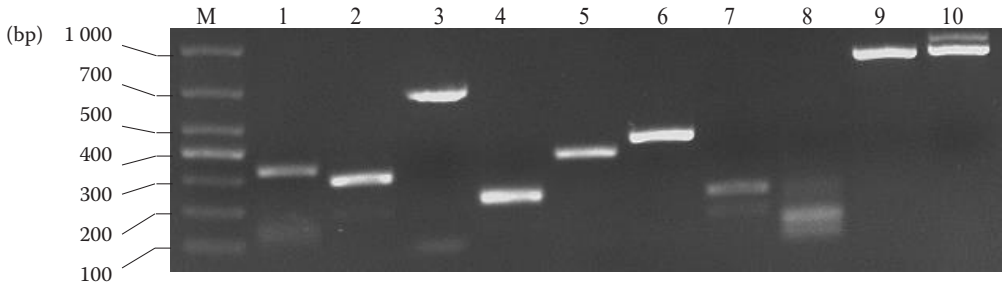

Figure 5. RT-PCR fragments of 10 predicted genes
M – DL1000 DNA marker; 1~10 – fragments for predicted Gene 1~ Gene 10

https://doi.org/10.17221/49/2022-CJGPB

of Gene 3 (647 bp) was identical with the mRNA sequence $_{(1689–2335)}$ of Gene 3 (Figure 6B), while that of Gene 7 (201 bp) was consistent with Gene 7 genomic sequence $_{(6537–6631)}$ (L_1) and genomic sequence $_{(9764–9887)}$ (L_2) respectively (Figure 6C), not completely consistent with the predicted mRNA sequence. Except for the three genes, the RT-PCR sequences for the rest seven genes either had no significant similarity or were not completely consistent with the corresponding genomic sequence or predicted mRNA sequences, suggesting that these RT-PCR sequences might be sequences of genes outside the final mapping region. Our result also indicated that during the young panicle development, only a small number of genes within the final mapping region were selectively expressed, which is in line with the general cognition.

Prediction for candidate gene(s) of *Rf(fa).* According to the bioinformatics analysis of the genes within the final mapping region, the proteins of three predicted genes were partially similar to several PPR proteins in sequence, and therefore were considered to be PPR protein genes, which encoded proteins of 1 284, 831 and 876 amino acid residues, hence named *PPR1284*, *PPR831* and *PPR876*, respectively (Figure 4). But at transcriptional level, only one of the PPR protein genes (*PPR1284*) was expressed. Although two other genes (Gene 2 and Gene 3) were expressed at RNA level, they were less likely to be the target gene, taking into consideration the function of restoring genes for hybrid rice. Therefore, *PPR1284* was most probably the candidate gene of *Rf(fa)*. Till now, hybrid rice restoring genes identified across the world are mainly located on Chromosome 10 (Liu

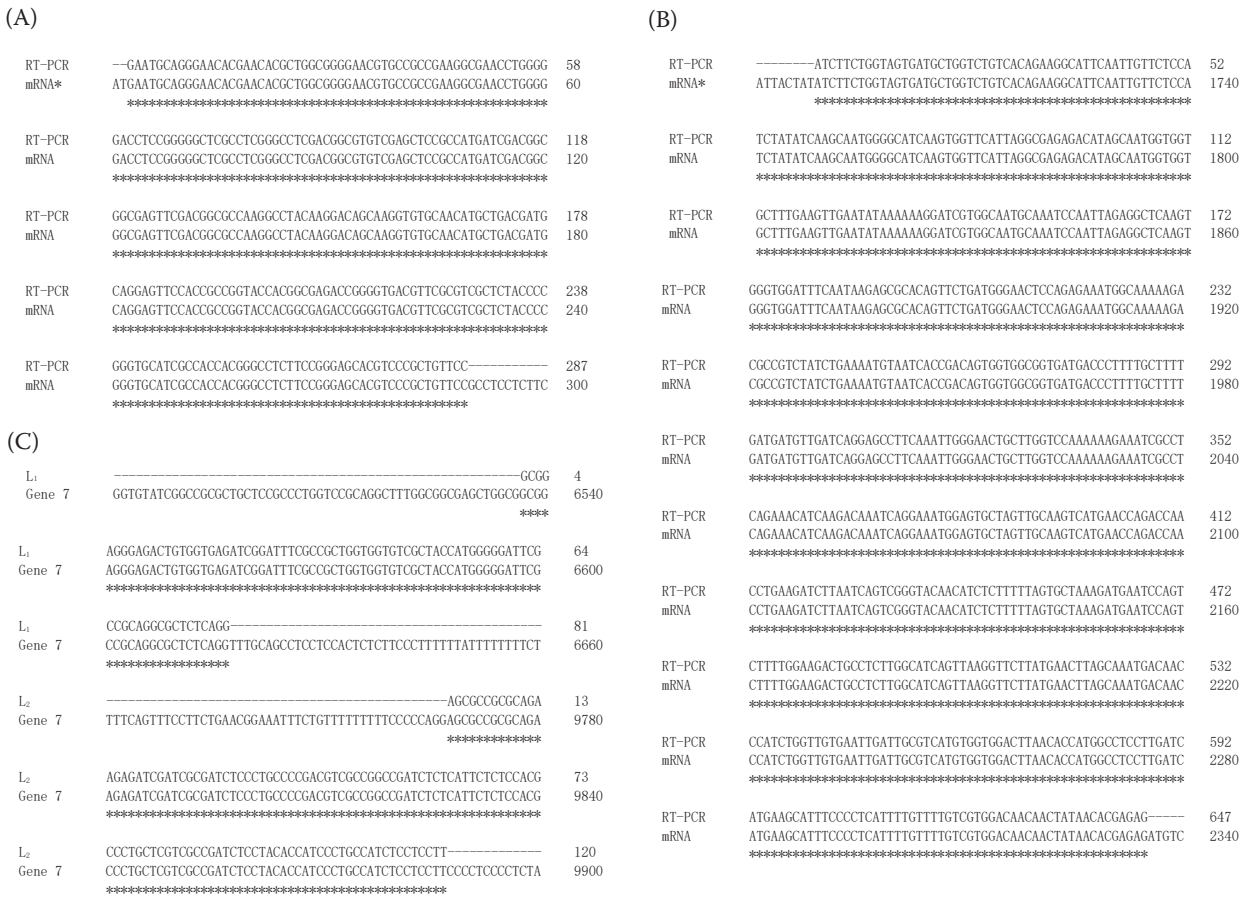

Figure 6. Comparison of three gene RT-PCR sequences with their corresponding mRNA or genomic sequences of Gene 2, Gene 3 and Gene 7: comparison of the RT-PCR sequence with the mRNA sequence for Gene 2 (A), comparison of the RT-PCR sequence with the mRNA sequence for Gene 3 (B), comparison of the RT-PCR sequence with the genomic sequence for Gene 7 (C)

L_1 – RT-PCR sequence (1-81); L_2 – RT-PCR sequence (82–201); *here all the U bases in the mRNA sequence were replaced by T

https://doi.org/10.17221/49/2022-CJGPB

et al. 2004; Huang et al. 2012; Luo et al. 2013; Keisuke et al. 2016; Zhang et al. 2022), including *Rf1a*, *Rf1b*, *Rf4*, *Rf5*, *Rf98* and *Rf18(t)*, the rest are on the other chromosomes: *Rf2*, *Rf3*, *Rf6*, *Rf17* and *Rf18(t)* were on Chromosome 2 (Eitabashi et al. 2011), Chromosome 1 (Zhang et al. 1997), Chromosome 8 (Huang et al. 2012), Chromosome 4 (Fujii & Toriyama 2009) and Chromosome 1 (Zhang et al. 2022) respectively. *Rf(fa)* was also located on Chromosome 10. Accordingly, rice male sterile restorers might originate from a common ancestral gene, which was later replicated and inserted into adjacent positions on the same chromosome, even into the other chromosomes. Later, these genes further turned into highly homologous genes through mutation, but were functionally different from each other.

About two months after this manuscript was submitted, Jiang et al. (2022) published online a paper entitled "Fujian cytoplasmic male sterility and the fertility restorer gene *OsRf19* provide a promising breeding system for hybrid rice". Their cloned *OsRf19* turned to be our predicted candidate gene Gene 7 for *Rf(fa)* within our final mapping region in this study. They used RICE6K microarray analysis along with an F_2 population and a BC_2F_2 population to map the target gene, while we used a small F_2 population and a large BC_1F_1 population, combined with the expression analysis of genes within the final mapping region to complete the mapping and the subsequent prediction of the candidate gene. Their result proved that our mapping and the prediction of the candidate gene for *Rf(fa)*/*OsRf19* is correct and accurate. They and we used different materials and methods in the mapping of the target gene, but obtained the same mapping result, which provided important reference for later researchers. What is particularly important is that we provided several SSR molecular markers tightly linked to *Rf(fa)*/*OsRf19* other than those they provided, which will be very useful to the development and application of CMS-FA hybrid rice. They analyzed the origin of the target gene, but did not investigate the genomic sequence differences of the previous mapping region between the wild rice and cultivated rices. We performed RT-PCR analysis of all the genes in the final mapping region to help predict the candidate gene for *Rf(fa)*/*OsRf19*, and their result demonstrated the effectiveness of this method for the candidate gene prediction. In addition, we discussed the size of mapping segregation populations required for the preliminary and fine mapping of a gene. Therefore our study is not just a confirmation of Jiang et al. (2022), but further provides new information.

CONCLUSION

The target gene *Rf(fa)* was finally mapped to an area of 78.87 kb on Chromosome 10 of rice, flanked by two SSR molecular markers MM2000 and RM25658. Within this mapping region, the expression of three genes at transcriptional level was detected, including Gene 2, Gene 3 and Gene 7 (*PPR1284*), and that of the other genes at transcriptional level was not. Thus the PPR protein gene *PPR1284* was predicted as the candidate gene of *Rf(fa)*. There existed extensive genomic sequence divergence between the wild rice and the cultivated rice.

REFERENCES

Chiu X.L., Wang H.F., Wang N.Y., Li Y., Liang K.J. (2013): Genetic effects analysis of male sterile cytoplasm and nucleo-cytoplasmic interactions in CMS-FA hybrid rice. Journal of Fujian Agriculture and Forestry University (Natural Science Edition), 42: 113–118. (in Chinese)

Cui X., Wise R.P., Schnable P.S. (1996): The *rf2* nuclear restorer gene of male-sterile T-cytoplasm maize. Science, 272: 1334–1336.

Ding J., Qing L., Yidan O., Hailing M., Pingbo Z., Jialing Y., Caiguo X., Xianghua L., Jinghua X., Qifa Z. (2012): A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. The Proceedings of National Academy of Sciences, 109: 2654–2659.

Eitabashi E., Iwata N., Fujii S., Kazama T., Toriyama K. (2011): The fertility restorer gene, *Rf2*, for lead rice-type cytoplasmic male sterility of rice encodes a mitochondrial glycine-rich protein. Plant Journal, 65: 359–367.

Fujii S., Toriyama K. (2009): Suppressed expression of retrograde-regulated male sterility restores pollen fertility in cytoplasmic male sterile rice plants. The Proceedings of the National Academy of Sciences, 106: 9513–9158.

Guo G.Q., Song X.W., Song P.Y., Yin J.Y., He Q., Yuang D.Y., Deng F.H., Yuan L.P. (2016): Morphological characterization and fine mapping of zebra leaf mutant *zebra1349* in rice (*Oryza sativa* L.). Acta Agronomica Sinica, 42: 957–965. (in Chinese)

Hua J., Wang K., Huang W.C., Liu G., Gao Y., Wang J.M., Huang Q., Jia Y.X., Qin X.J., Lei W., Zhu R.S., Li S.Q., Yang D.C., Zhu Y.G. (2012): The rice pentatricopeptide repeat protein RF5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. Plant Cell, 24: 109–122.

https://doi.org/10.17221/49/2022-CJGPB

Huang W.C., Hu J., Yu C.C., Huang Q., Wan L., Wang L.L., Qin X.J., Ji Y.X., Zhu R.S., Li S.Q., Zhu Y.G. (2012): Two non-allelic nuclear genes restore fertility in a gametophytic pattern and enhance abiotic stress tolerance in the hybrid rice plant. Theoretical and Applied Genetics, 124: 799–807.

Huang W.C., Yu C.C., Hua J., Wang L.L., Dan Z.W., Zhou W., He C.L., Zeng Y.F., Yao G. X., Qi J.Z., Zhang Z.H., Zhu R.S., Chen X.F., Zhu Y.G. (2015): Pentatricopeptide-repeat family protein RF6 functions with hexokinase 6 to rescue rice cytoplasmic male sterility. The Proceedings of the National Academy of Sciences, 112: 14984–14989.

Jiang H.C., Lu Q., Qiu S.Q., Yu H.H., Wang Z.J., Yu Z.C., Lu Y.R., Wang L., Xia F., Wu Y.Y., Li F., Zhang Q.L., Liu G., Song D.D., Ma C.L., Ding Q., Zhang X.B., Zhang L., Zhang X.T., Li X., Zhang J.W., Xiao J.H., Li X.H., Wang N.Y., Ouyang Y.D., Zhou F.S., Zhang Q.F. (2022): Fujian cytoplasmic male sterility and the fertility restorer gene *OsRf19* provide a promising breeding system for hybrid rice. The Proceedings of the National Academy of Sciences, 119: e2208759119.

Keisuke I., Tomohiko K., Kinya T. (2016): A gene encoding pentatricopeptide repeat protein partially restores fertility in RT98-Type cytoplasmic male-sterile rice. Plant and Cell Physiology, 57: 2187–2193.

Li Y., Zhang M., Yang X.F., Lin C.F., Duan Y.L., Wang N.Y. (2016): Fine mapping of a fertility restoring gene for a new CMS hybrid rice system. Molecular Breeding, 36:141.

Liu X.Q., Xu X., Tan Y.P., Li S.Q., Hu J., Huang J.Y., Yang D.C., Li Y.S., Zhu Y.G. (2004): Inheritance and molecular mapping of two fertility-restoring loci for Honglian gametophytic cytoplasmic male sterility in rice (*Oryza sativa* L.). Molecular Genetics and Genomics, 271: 586–594.

Luo D.P., Xu H., Liu Z.L., Guo J.X., Li H.Y., Chen L.T., Fang C., Zhang Q.Y., Bai M., Yao N., Wu H., Ji C.H., Zheng H.Q., Chen Y.L., Ye S., Li X.Y., Zhao X.C., Li R.Q., Liu Y.G. (2013): A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. Nature Genetics, 45: 573–578.

Murray M.G., Thompson W.F. (1980): Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research, 8: 4321–4325.

Ponnuswamy R., Singh A.K., Raman M.S., Subbarao L.V., Neeraja C.N. (2020): Conversion of partial restorer Swarna into restorer by transferring fertility restorer *Rf* gene(s) through marker assisted back cross breeding (MABB) in rice. Scientific Reports, 10: 1101.

Shinya F., Tomohiko K., Yukihiro I., Soichi K., Kinya O. (2014): A candidate factor that interacts with RF2, a restorer of fertility of lead rice-type cytoplasmic male sterility in rice. Rice, 7: 21.

Sota F., Kinya T. (2009): Suppressed expression of RETROGRADE-REGULATED MALE STERILITY restores pollen fertility in cytoplasmic male sterile rice plants. The Proceedings of the National Academy of Sciences, 106: 9513–9518.

Tan Y.N., Sun X.W., Yuan D.Y., Sun Z.Z., Yu D., He Q., Duan M.J., Deng H.F., Yuan L.P. (2015): Identification and fine mapping of green-revertible chlorina gene *grc2* in rice (*Oryza sativa* L.). Acta Agronomica Sinica, 41: 831–837. (in Chinese)

Tautzs D. (1989): Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Research, 17: 6463–6471.

Tomohiko K., Kinya T. (2014): A fertility restorer gene, *Rf4*, widely used for hybrid rice breeding encodes a pentatricopeptide repeat protein. Rice, 7: 28.

Virmaniss S. (1988): Current status of analysis and symbols for male-sterile cytoplasms and fertility-restoring genes. Rice Genetic Newsletter, 59: 15.

Wang H.F., Wang N.Y., Li Y., Liang K.J., Chiu X.L., Zhou W.Y. (2010): Genetic effects and heterosis analysis on yield-related traits of CMS-FA hybrid rice. Scientia Agricultura Sinica, 43: 230–239. (in Chinese)

Wang N.Y., Liang K.J., Li Y., Wang Y., Wang H.F., Chiu X.L. (2008a): Inheritance of restorer gene for CMS-FA hybrid rice. Acta Agronomica Sinica, 34: 1929–1937. (in Chinese)

Wang N.Y., Liang K.J., Li Y., Wang Y., Wang H.F., Chiu X.L., Wei X.Y., Zhang Y.Y., Ke B., Zhao J.Y. (2008b): Screening parent resources of hybrid rice in new cytoplasm male sterile system (CMS-FA). Acta Agronomica Sinica, 34: 1549–556. (in Chinese)

Wang Z.H., Zou Y.J., Li X.Y., Zhan Q.Y., Chen L.T., Wu H., Su D.H., Chen Y.L., Guo J.X., Luo D., LongY.M., Zhong Y., Liu Y.H. (2006): Cytoplasmic male sterility of rice with Boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. Plant Cell, 18: 676–687.

Xie H.W., Peng X.J., Qian M.J., Cai Y.C., Ding X., Chen Q.S., Cai Q.Y., Zhu Y.L., Yang N.A., Cai Y.H. (2018): The chimeric mitochondrial gene *orf182* causes non-pollen-type abortion in Dongxiang cytoplasmic male-sterile rice. Plant Journal, 95: 715–726.

Yuan L.P. (1966): The cytoplasmic male sterility in rice. China Science Bulletin, 17: 185–188. (in Chinese)

Zhang G., Lu Y., Bharaj T.S., Virmani S.S., Huang N. (1997): Mapping of the *Rf-3* nuclear fertility- restoring gene for WA cytoplasmic male sterility in rice using RAPD and RFLP markers. Theoretical and Applied Genetics, 94: 27–33.

Zhang H.G., Li X.X., Xu Z.P., Wan Z.H., Wang R.X., Zhao X.Q., Tu G.L., Liang G.H., Gu M.H., Tang S.Z. (2022). Precise genetic mapping of *Rf18(t)*, a new fertility restorer

https://doi.org/10.17221/49/2022-CJGPB

gene from 'Nipponbare' for wild abortive cytoplasmic male sterility in rice (*Oryza sativa* L.). Theoretical and Applied Genetics, 135: 2687–2698.

Zhang J.W., Chen L.L., Xing F., David A., Kudrna, Yao W., Dario Copetti, Mu T., Li W.M., Song J.M., Xie W.B., Lee S.H., Jayson Talag L., Shao An Y., Zhang C.L., Ouyang Y.D., Sun S., Jiao W.B., Lv F., Bogu Du, M.Z Luo, Maldonado C.E., Goicoechea J.L., Xiong L.Z., Wu C.Y., Xing Y.Z., Zhou D.X., Yu S.B., Zhao Y., Wang G.W., Yeisoo Yub, Luo Y.J., Zhou Z.W., Padilla Hurtado B.E., Danowitz A., Wing R., Zhang Q.F.(2016): Extensive sequence divergence between the reference genomes of two elite *indica* rice varieties Zhenshan 97 and Minghui 63. The Proceedings of the National Academy of Sciences, 113: E5163–E5171.

Zhang Q.U., Liu Y.G., Zhang G.Q., Mei M.T. (2002): Molecular mapping of the fertility restorer gene *Rf-4* for WA cytoplasmic male sterility in rice. Acta Genetica Sinica, 29: 1001–1004. (in Chinese)

Zhang Z., Deng Y., Tan J., Hu S., Yu J., Xue Q. (2007): A genomewide microsatellite polymorphism database for the *indica* and *japonica* rice. DNA Research, 14: 37–45.

Received: June 20, 2022
Accepted: September 9, 2022
Published online: October 10, 2022