

Development and validation of candidate gene-specific markers for the major fertility restorer genes, Rf4 and Rf3 in rice

K. Pranathi \cdot B. C. Viraktamath \cdot C. N. Neeraja \cdot S. M. Balachandran \cdot A. S. Hari prasad \cdot

P. Koteswara Rao \cdot P. Revathi \cdot P. Senguttuvel \cdot S. K. Hajira \cdot C. H. Balachiranjeevi \cdot

S. Bhaskar Naik · V. Abhilash · M. Praveen · K. Parimala · S. R. Kulkarni · M. Anila ·

G. Rekha \cdot M. B. V. N. Koushik \cdot B. Kemparaju \cdot M. S. Madhav \cdot S. K. Mangrauthia \cdot

G. Harika \cdot T. Dilip \cdot R. R. Kale \cdot V. Vishnu Prasanth \cdot V. Ravindra Babu \cdot R. M. Sundaram

Received: 27 June 2016 / Accepted: 3 October 2016 - Springer Science+Business Media Dordrecht 2016

Abstract Two major nuclear genes, $Rf3$ and $Rf4$, are known to be associated with fertility restoration of wild-abortive cytoplasmic male sterility (WA-CMS) in rice. In the present study, through a comparative sequence analysis of the reported putative candidate genes, viz. $PPR9-782-(M,I)$ and $PPR762$ (for $Rf4$) and SF21 (for Rf3), among restorer and maintainer lines of rice, we identified significant polymorphism between the two lines and developed a set of PCR-based codominant markers, which could distinguish maintainers from restorers. Among the five markers

Electronic supplementary material The online version of this article (doi[:10.1007/s11032-016-0566-8\)](http://dx.doi.org/10.1007/s11032-016-0566-8) contains supplementary material, which is available to authorized users.

- K. Pranathi · B. C. Viraktamath · C. N. Neeraja ·
- S. M. Balachandran · A. S. Hari prasad ·
- P. Koteswara Rao · P. Revathi · P. Senguttuvel ·
- S. K. Hajira C. H. Balachiranjeevi S. Bhaskar Naik -
- V. Abhilash M. Praveen K. Parimala -
- S. R. Kulkarni M. Anila G. Rekha M.
- B. V. N. Koushik · B. Kemparaju · M. S. Madhav ·
- S. K. Mangrauthia · G. Harika · T. Dilip ·
- R. R. Kale V. Vishnu Prasanth V. Ravindra Babu -R. M. Sundaram (\boxtimes)
- Crop Improvement Section, Indian Institute of Rice Research, Rajendranagar, Hyderabad 500030, India e-mail: rms_28@rediffmail.com

developed targeting the polymorphisms in PPR9- 782- (M,I) , the marker **RMS-PPR9-1** was observed to show clear polymorphism between the restorer $(n = 120)$ and maintainer lines $(n = 44)$ analyzed. Another codominant marker, named RMS-PPR762 targeting PPR762, displayed a lower efficiency in identification of restorers and maintainers, indicating that $PPR9-782-(M,I)$ is indeed the candidate gene for $Rf4$. With respect to $Rf3$, a codominant marker, named RMS-SF21-5 developed targeting SF21, displayed significantly lower efficiency in identification of restorers and non-restorers as compared to the Rf4 specific markers. Validation of these markers in a $F₂$ mapping population segregating for fertility restoration indicated that Rf4 has a major influence on fertility restoration and $Rf3$ is a minor gene. Further, the functional marker RMS-PPR9-1 was observed to be very useful in identification of impurities in a seed lot of the popular hybrid, DRRH3. Interestingly, when RMS-PPR9-1 and RMS-SF21-5 were considered in conjunction with analysis, near-complete, marker– trait co-segregation was observed, indicating that deployment of the candidate gene-specific markers both $Rf4$ and $Rf3$, together, can be helpful in accurate identification of fertility restorer lines and can facilitate targeted transfer of the two restorer genes into elite varieties through marker-assisted breeding.

Keywords Fertility restoration $-Rf4 - Rf3 - WA$ CMS - Gene-specific markers - Hybrid seed purity

K. Pranathi and R. M. Sundaram have contributed equally to this work.

Introduction

Breeding rice for higher yield remains the key priority for developing nations such as India, which needs to produce \sim 125 million tonnes of rice by 2030 to feed its burgeoning population. Hybrids in rice have yield superiority of about 15–20 % over the best commercial inbred varieties under similar conditions (Virmani [1996\)](#page-13-0) and large-scale adoption of hybrid rice production is one of the feasible options to meet the food security challenges in India. Hybrid based on wildabortive cytoplasmic male sterility (WA-CMS) system has been extensively used in commercial rice hybrids production in most of the Asian countries including India (Lin and Yuan [1980](#page-12-0); Virmani and Wan [1988\)](#page-13-0), and so far, 75 hybrids based on WA-CMS system have been released for commercial cultivation in India (AS Hariprasad, personal communication). The utility of the CMS lines in hybrid rice breeding is determined by the availability of characterized and effective fertility restoration lines. A total of 17 alleles for fertility restoration have been identified in rice, and all except rf17 are dominant in rice. Among these, at least two genes, viz. $Rf3$ (located on chromosome 1) and Rf4 (located on chromosome 10), are known to control fertility restoration of WA cytoplasm (Zhang et al. [1997;](#page-13-0) Yao et al. [1997](#page-13-0)). Various attempts have been made to fine-map and characterize the candidate genes underlying Rf4 and Rf3 (Ahmadikhah and Karlov [2006](#page-12-0); Sheeba et al. [2009](#page-12-0); Ngangkham et al. [2010;](#page-12-0) Balaji et al. [2012](#page-12-0)). Ngangkham et al. ([2010\)](#page-12-0) proposed that a gene encoding a pentatricopeptide repeat (PPR) motif-containing protein, named PPR3, located on the long arm of chromosome 10 is the candidate gene for $Rf4$, while recently, another study (Tang et al. [2014](#page-13-0)) identified another candidate, PPR9- 782 (M,I) , located in the same region as *PPR3* as the candidate for $Rf4$ gene. With respect to $Rf3$, Balaji et al. ([2012\)](#page-12-0) reported that a gene, named SF21, encoding a pollen-specific protein to be putative candidate for the gene.

In WA-CMS-based hybrid breeding (also called three-line system of hybrid rice breeding), identification of potential restorers among the diverse rice germplasm lines is of significant importance, as genetically diverse restorer lines can be helpful in breeding hybrids with higher magnitude of heterosis. The traditional method of identifying restorers by breeders involves test crossing the prospective lines with selected WA-CMS lines and evaluating the F_1 progenies for pollen and spikelet fertility. Lines with progenies showing >70 % pollen and spikelet fertility are then designated as restorers (Govinda Raj and Virmani [1988](#page-12-0)). Molecular mapping of Rf3 and Rf4 can reduce the time and effort involved in identification of fertility restorer lines (Sattari et al. [2007;](#page-12-0) Sheeba et al. [2009\)](#page-12-0). Further, molecular markers specific for Rf3 and $Rf4$ can aid in targeted transfer of the two Rf genes into elite genetic backgrounds and also facilitate accurate estimation of genetic impurities in hybrid seed lots (Nandakumar et al. [2004;](#page-12-0) Sundaram et al. [2008](#page-13-0)). Many markers have been developed for Rf4 (Ahmadikhah and Karlov [2006](#page-12-0); Ngangkham et al. [2010](#page-12-0); Balaji et al. [2012](#page-12-0)), and a few have been developed for Rf3 (Nas et al. [2003](#page-12-0)). However, these markers display limited efficiency in accurate identification of restorers, as all of them are linked markers and not specific for the putative candidate genes underlying either $Rf3$ or Rf4. The present study was carried out with the objective to analyze the sequence polymorphism in the genomic region underlying the reported candidate genes for Rf3 and Rf4, develop candidate genespecific, PCR-based codominant markers, validate them among a large set of known maintainer and restorer lines and a mapping population segregating for the trait of fertility restoration and finally demonstrate the utility of the candidate gene-specific marker in accurate identification of impurities in seed lot of a commercial rice hybrid.

Materials and methods

Plant materials

The plant materials in the study included a total of 120 restorer and 44 non-restorer lines (i.e., maintainers) of indica-type rice for WA-CMS cytoplasm (Table [1](#page-2-0)), which were used for validation of the gene-specific markers developed for $Rf3$ and $Rf4$. The developed markers were also validated in a segregating population consisting of $1252 \mathrm{F}_2$ individuals derived from the cross between the WA-CMS line, IR58025A and the restorer line, KMR3R, which were phenotyped for spikelet fertility. A set of 71 wild rice lines (Supplementary Table 1) was analyzed for their amplification pattern with respect to the gene-specific marker for Rf4. In addition, a seed lot of the popular rice hybrid

Table 1 continued

Table 1 continued

continued

DRRH3 consisting of 400 seeds was also included for analysis of efficiency of gene-specific marker for Rf4 in accurate identification of genetic impurities. All the plant materials utilized in the study were collected from Hybrid Rice Section of ICAR-Indian Institute of Rice Research (ICAR-IIRR), Hyderabad.

Analysis of gene sequences of Rf3 and Rf4

The candidate genes $PPR9-782-(M,I)$ (Tang et al. [2014;](#page-13-0) Kazama and Toriyama [2014](#page-12-0)) and PPR762 (Balaji et al. 2012) reported to be specific for $Rf4$ on chromosome 10 were considered for sequence analysis. The reported restorer sequences (PPR9-782- M and PPR9-782-I) and non-restorer gene sequences (PPR9-409 and PPR9-782-ZH) of PPR9 gene (Tang et al. [2014\)](#page-13-0) were downloaded from NCBI/GenBank public database. The coordinates of PPR9-782- M gene were identified in Nipponbare, a japonica cultivar from (19,287,680 to 19,295,473 bp; Pseudo molecule 6.1, [http://rice.plantbiology.msu.edu/](http://rice.plantbiology.msu.edu/pseudomolecules) [pseudomolecules\)](http://rice.plantbiology.msu.edu/pseudomolecules) using BioEdit tool version 7.0.9 (Hall [2007\)](#page-12-0). Using ClustalW multiple sequence alignment tool (Higgins et al. [1994](#page-12-0)), two functional restorer sequences and two non-restorer sequences were compared to identify different polymorphic regions (Supplementary Figure 1). Further, a 25-kb region upstream and a 25-kb region downstream of PPR9-782-M on chromosome 10 of Nipponbare (19,290,587–19,340,587 bp) and chromosome 10 of indica cultivar, 93–11 (17,749,895–17,799,895 bp) were also aligned using ClustalW tool in order to identify the polymorphic regions in the vicinity of the candidate gene. Similar sequence analysis was performed for another reported candidate gene PPR762 specific for Rf4 (Balaji et al. [2012\)](#page-12-0). The reported amplicon sequences of DRCG-Rf4-14 marker (Balaji et al. [2012\)](#page-12-0) targeting PPR762 in restorer and nonrestorer sequences were also considered for polymorphism analysis.

With respect to $Rf3$, a pollen-specific protein, $SF21$, located on chromosome 1 was identified earlier by fine-mapping analysis to be the putative candidate gene (Balaji et al. [2012\)](#page-12-0). SF21 gene sequence (LOC_Os01g09670) was downloaded from Gramene/NCBI public database. The coordinates of SF21 gene were identified on chromosome 1 of Nipponbare, japonica cultivar (Pseudo molecule 6.1, [http://rice.plantbiology.msu.edu/pseudomolecules\)](http://rice.plantbiology.msu.edu/pseudomolecules)

and in indica cultivar 93–11 (Beijing Rice Information System [http://rice.genomics.org.cn/rice\)](http://rice.genomics.org.cn/rice) using BioEdit tool version 7.0.9 (Hall [2007](#page-12-0)). A 10-kb region upstream and 10-kb region downstream of SF21 gene from both japonica (4,917,224–4,992,046 bp) and indica (5,350,773–5,371,596 bp) were aligned using ClustalW alignment tool (Higgins et al. [1994](#page-12-0)) (Supplementary Figure 2) to identify polymorphic regions in the vicinity of the SF21 gene.

Primer designing and PCR analysis of the developed markers

The different polymorphic regions identified within the PPR9-782-M and PPR-782-I gene and also in the vicinity of gene (i.e., within 50 kb on either side) were targeted for designing of five PCR-based codominant markers specific for Rf4. Another codominant marker specific for *Rf4* was designed targeting the polymorphism within PPR762 gene. In addition, different polymorphic regions identified based on alignment between SF21 sequences from japonica and indica were targeted for designing of five PCR-based codominant markers specific for Rf3. All these primer pairs were designed using Primer 3 online tool [\(http://](http://bioinfo.ut.ee/primer3-0.4.0) bioinfo.ut.ee/primer3-0.4.0), and the primer sequences of these markers specific to $Rf3$ and $Rf4$ are listed in Table 2. In addition, the reported SSR markers RM6100 (specific for Rf4) (Singh et al. [2005](#page-13-0)), DRCG-Rf4-14 (specific for Rf4) (Balaji et al. [2012](#page-12-0)) and DRRM-Rf3-10 (specific for Rf3) (Balaji et al. [2012\)](#page-12-0) were also considered for analysis.

The total genomic DNA was isolated from young, healthy leaves of all the restorer lines, maintainer lines and individuals of the segregating F_2 mapping population by following the method of Dellaporta et al. [\(1983](#page-12-0)). The isolated DNA was used for PCR amplification with the codominant markers developed in the study. PCR was performed in $20 \mu l$ reaction volumes containing 1X PCR buffer [10 mM Tris-HCI (pH 8.3),

50 mM KCI, 1.5 mM MgCL₂, 0.01 % (v/v) gelatin], 30–50 ng of template DNA, 5 pmol of each primer, $200 \mu M$ (each) deoxyribonucleotide and 1 unit of Taq polymerase (Merck, India). PCR conditions included an initial denaturation step at $94 °C$ for 5 min, followed by 30 cycles of 94 \degree C for 30 s, 55 \degree C for 30 s and 72 °C for 1 min and a final extension at 72 °C for 7 min. All amplified products were resolved in 2–3.5 % agarose gels (Lonza Inc., USA) along with 100-bp molecular marker (Merck, India). The codominant markers that showed clear polymorphism between restorers and maintainers were validated in the F_2 segregating populations. The scores 1, 2 and 3 were given to codominant markers for parent 1 type (P1) and parent 2 type (P2) and heterozygous (F1). The segregation of codominant markers in the $F₂$ population was studied by Chi-square test for the Mendelian segregation ratio 1:2:1 as outlined by Gomez and Gomez ([1984\)](#page-12-0).

Spikelet fertility analysis

About 20-day-old seedlings of F_2 individuals were transplanted in the field. At reproductive stage of growth, just before flowering, the panicles of main tiller and two side tillers of each individual plant were bagged with a paper bag to prevent cross-pollination. The seed set in each panicle was counted, and spikelet fertility was determined according to Sheeba et al. [\(2009](#page-12-0)). All the plants in the population were classified into four classes based on spikelet fertility percentage, namely fertile (more than 71 % spikelet fertility), partially fertile $(31–70\%)$, partially sterile $(1–30\%)$ and sterile (0 %).

Analysis of impurities in a seed lot of DRRH3 using Rf4-specific codominant marker

Four hundred seedlings of the popular rice hybrid DRRH3 from a seed lot were planted in a grow-out plot in the experimental farm of ICAR-Indian Institute of Rice Research, Hyderabad, India, during wet season 2015. DNA was isolated from 20-day-old seedlings of the 400 coded plants, individually as per the procedure of Zheng et al. [\(1995](#page-13-0)). Genotyping of the 400 seedlings was done using the Rf4-specific codominant marker, RMS-PPR9-1, which exhibited polymorphism among the female (APMS6A) and male (RPHR-1005) parents of DRRH3. The genotype

inferred from the marker profile was compared with the phenotype at maturity to verify the results derived from marker analysis with grow-out test (as described in Yashitola et al. [2002](#page-13-0) and Sundaram et al. [2008](#page-13-0)).

Sequencing of PCR fragments

Amplified PCR product of RMS-PPR9-1 marker from KMR3R and IR58025A was gel-purified (Wizard $\mathbb{S}V$ PCR clean up kit, Promega), cloned into pDrive cloning vector (Qiagen, USA) and sequenced using an ABI Prism 3700 automated DNA sequencer (PerkinElmer, Wellesley, MA) as per the procedure suggested in Rajendrakumar et al. ([2007\)](#page-12-0). Homology search was performed by BLASTN algorithm (Altschul et al. [1990](#page-12-0)) through the National Center for Biotechnology Information [\(http://www.ncbi.nlm.](http://www.ncbi.nlm.nih.gov/blast) [nih.gov/blast\)](http://www.ncbi.nlm.nih.gov/blast), and the amplicon sequences from IR58025A and KMR3R were aligned using the software ClustalW to validate the *in-del* polymorphisms which were identified through sequence analysis of PPR9 genomic regions.

Results

Development and validation of candidate genespecific markers for $Rf3$ and $Rf4$

The sequence analysis of candidate gene PPR9-782- M, PPR9-782-I (which are specific for $Rf4$) and the non-restorer sequences PPR9-409 and PPR9-782-ZH revealed the presence of three major in-dels within the gene (Supplementary Figure 1). These include a 42-bp in-del, identified in the first intronic region, a 105-bp in-del and a 1476-bp in-del identified within second exonic region. Targeting each of these in-del polymorphisms, codominant markers were designed and validated. Two other major in-dels were also identified in the upstream region of PPR9-782-M gene and targeted for development of codominant markers. Out of the five codominant markers specific for in-dels within PPR9-782-M or in its vicinity, three markers, viz. a marker targeting 42-bp in-del polymorphism within PPR9 gene, i.e., RMS-PPR9-1, two codominant markers, viz. RMS-PPR9-4, RMS-PPR9-5 targeting polymorphisms in the upstream region of PPR9-782-M gene displayed clear polymorphism between IR58025A and KMR3R (Fig. [1;](#page-7-0) Table [3](#page-7-0)).

Table 3 Expected amplification sizes of the markers developed in the

study

Fig. 1 Amplification pattern of markers developed targeting the candidate genes for Rf4, viz. PPR9-78-M and PPR762: a RMS-PPR9-1 (targeting in-del within PPR9-782-M); b RMS-PPR762 (targeting PPR762); c RMS-PPR9-4 marker (targeting PPR9-782-M); d RMS-PPR9-5 marker (targeting PPR9-782-

M); e RMS-SF21-5 marker (targeting SF21). L1 indicates 100-bp ladder, L2 indicates 50-bp ladder, 1 indicates IR58025A, 2 indicates IR58025B, 3 indicates KMR3R, and 4 indicates KRH2 in the figure

S. no.	Primer name	Expected PCR amplicon size (bp)	Expected PCR product size (bp) in		
			Restorer	Non-restorer	Hybrid
1	RMS-PRR9-1	114/159	114	159	114,159
$\overline{2}$	RMS-PPR9-2	447/1923	447	1923	447,1923
3	RMS-PPR9-3	365/470	365	470	365,470
$\overline{4}$	RMS-PPR9-4	129/160	129	160	129,160
5	RMS-PPR9-5	178/360	178	360	178,360
6	RMS-PPR762	280/385	280	385	280/385
7	RMS-SF21-1	183/131	183	131	183,131
8	RMS-SF21-2	312/415	312	415	312,415
9	RMS-SF21-3	196/165	196	165	196,165
10	RMS-SF21-4	101/113	101	113	101.113
11	RMS-SF21-5	172/127	172	127	172,127

The analysis of restorer and non-restorer amplicon sequences of DRCG-RF4-14 marker targeting PPR762 gene revealed existence of a 105-bp in-del polymorphism. Targeting this, a codominant marker RMS-PPR762 was designed and validated. RMS-PPR762 showed clear polymorphism between IR58025A and KMR3R (Fig. 1; Table 3). Thus, a total of four polymorphic markers were designed and validated in this study targeting the putative candidate genes for Rf4 (i.e., PPR9-782-M and PPR762).

In addition to $Rf4$, the putative candidate gene for Rf3 (another fertility restorer gene for WA-CMS), viz. SF21 (Balaji et al. [2012](#page-12-0)), was analyzed through comparative sequence analysis of restorer and nonrestorer genotypes and five major in-del polymorphisms were identified in the vicinity of gene. Targeting each of these, a codominant marker was designed and validated. However, only one codominant marker, RMS-SF21-5, displayed clear polymorphism between the WA-CMS lines IR58025A and the restorer line, KMR3R (Fig. [1](#page-7-0); Table [3\)](#page-7-0).

Marker–trait co-segregation analysis

The candidate gene-specific markers for Rf4 and Rf3 which have shown clear polymorphism between IR58025A and KMR3R were analyzed for their cosegregation with the trait of fertility restoration in a F_2 population derived from the cross IR58025A/KMR3R (Supplementary Table 2). All the codominant markers displayed a Mendelian segregation ratio of 1:2:1 in the $F₂$ mapping population and the candidate gene-specific marker for Rf4, RMS-PPR9-1 was observed to be significantly associated with the trait at $P < 0.01$. The markers RM6100 (Singh et al. [2005](#page-13-0)) and DRCG-Rf4- 14 (Balaji et al. [2012\)](#page-12-0) were also observed to be associated with trait phenotype, but to a lesser extent. The earlier reported marker DRCG-Rf4-14 (Balaji et al. [2012](#page-12-0)) and the marker RMS-PPR762 developed in this study, targeting the same 105-bp polymorphism in PPR762 gene, displayed identical association with the trait phenotype, but at a slightly lesser level of association as compared to RMS-PPR9-1 targeting PPR9-782-M. RMS-PPR762 showed clear and robust resolution of the restorer-specific and non-restorerspecific alleles when compared to the earlier designed marker, DRCG-Rf4-14. With respect to Rf3, the earlier reported SSR marker DRRM-Rf3-10 (Balaji et al. [2012](#page-12-0)) and RMS-SF21-5, the marker developed in this study, displayed same level of association with the trait phenotype (Supplementary Table 2) with the newly designed marker showing clear resolution of alleles as compared to DRRM-Rf3-10 (Supplementary Figure 3).

Assessment of prediction efficiency of the markers targeting Rf4 and Rf3

To validate the efficiency of these markers in accurately predicting the fertility restoration trait, they were analyzed with a set of 120 known restorers and 44 known non-restorers. The selection efficiency of the candidate gene-specific markers, RMS-PPR9-1 and RMS-PPR762 developed in this study, was 91 and 82 %, respectively (Supplementary Table 3; Fig. [2](#page-9-0)).

As expected, the earlier reported marker, DRCG-Rf4- 14, and the newly designed marker, RMS-PPR762, displayed same selection efficiency of 82 %, as they targeted the same polymorphism. The selection efficiency of candidate gene-specific marker for Rf3, RMS-SF21-5 and the earlier reported marker DRRMRf3-10 in identification of restorers and nonrestorers was identical (i.e., 57 %; Supplementary Table 3; Fig. [3](#page-10-0)). The combined selection efficiency of the best markers for Rf4 and Rf3, viz. RMS-PPR9- $1 + \text{RMS-SF21-5}$, was as high as 94 %. Particularly, the candidate gene-specific marker for Rf4, RMS-PPR9-1, was observed to show polymorphism among all the male and female parents of commercial rice hybrids based on WA-CMS system analyzed in this study (Supplementary Figure 4). When these markers (viz. RMS-PPR9-1, RMS-SF21-5) were analyzed in a set of wild rice accession belonging to O. nivara and O. rufipogon (Supplementary Table 1), it was observed that many O. nivara accessions showed the presence of restoring allele with respect to Rf4.

Utility of gene-specific markers for $Rf4$ in detection of impurities in hybrid/parental seed lots

The candidate gene-specific marker for Rf4 locus, RMS-PPR9-1, was deployed for identification of impurities in a seed lot of the hybrid DRRH3. With the help of the marker, a total of seven impurities were identified in the seed lot (Supplementary Figure 5), and a perfect correlation was observed between the marker analysis data and grow-out test (GOT) data.

Discussion

Wild-abortive (WA)-type CMS-based hybrids contribute significantly to the total rice cultivated area worldwide. Inheritance of fertility restoration for the WA-CMS system has been extensively investigated and two major loci, Rf4 and Rf3 are known to control the trait (Young and Virmani [1984;](#page-13-0) Li and Yuan [1986](#page-12-0); Virmani et al. [1986](#page-13-0); Govinda Raj and Virmani [1988](#page-12-0); Bharaj et al. [1991](#page-12-0); [1995](#page-12-0); Teng and Shen[1994](#page-13-0)). However, efforts to delineate the candidate genes underlying $Rf4$ and $Rf3$ are very limited. Recently, Ngangkham et al. [\(2010](#page-12-0)) and Balaji et al. ([2012\)](#page-12-0) identified that PPR3 and PPR762 are the putative

Fig. 2 Amplification pattern of RMS-PRR9-1, the candidate gene-specific marker for Rf4 (targeting in-del within PPR9-782- M) in a set of restorers and non-restorers (i.e., maintainers) in the

candidate genes for $Rf4$, while two independent groups cloned and characterized another putative candidate gene $PPR9-782(M,I)$, for $Rf4$ loci (Tang et al. 2014 ; Kazama and Toriyama [2014](#page-12-0)). According to the report of Tang et al. ([2014\)](#page-13-0), there are diverse functional Rf4/ rf4 alleles based on their donor source, PPR9-782- M allele from MH63 and *PPR9-782-I* from IR24 and two types of non-functional $rf4$ alleles $(rf4-i$ from *indica*) and *PPR9-782-ZH* $(rf4-i$ from *iaponica*). With respect to *Rf3*, a putative candidate gene, SF21 has been identified (Balaji et al. [2012](#page-12-0)).

In the present study, we analyzed the sequences of the above mentioned putative candidate genes, which have been earlier implicated with $Rf4$ and $Rf3$ controlled fertility restoration, identified sequence polymorphisms within the candidate genes and targeting these polymorphic regions, designed codominant markers and validated them in a mapping population and also in a large set of restorers and non-restorers lines. Based on marker–trait co-segregation analysis and analysis of selection efficiency of markers, we confirmed the candidacy of PPR9-782- M gene to be specific for $Rf4$. Our study is the first report on development of the candidate gene-specific above lanes, IR58025A (WA-CMS) and KMR3R (restorers) are used to standard checks

marker, named RMS-PPR9-1 targeting PPR9-782-M, and PPR9-782-I gene specific for Rf4 and another candidate gene-specific marker named RMS-SF21-5 targeting $SF21$ gene specific for $Rf3$. Among these two candidate gene-specific markers developed in this study, RMS-PPR9-1, specific for Rf4 has displayed higher selection efficiency of 91 % in terms of identification of all the known major restorer lines (Supplementary Table 2), as compared to the RMS-SF21-5 marker, which is specific for Rf3 showing only 57 % selection efficiency. These findings support the general understanding that a good restorer would possess Rf4 gene alone or Rf4 gene along with Rf3 gene, while lines possessing $Rf₃$ alone might not be good restorers. Thus, Rf4 has a stronger influence on the trait than $Rf3$ as observed earlier by several groups (Yao et al. [1997](#page-13-0); Sattari et al. [2008;](#page-12-0) Cai et al. [2013,](#page-12-0) [2014\)](#page-12-0). However, a few exceptions were also found in this study. Two of the known restorer lines IR66 and IR40750R were observed to possess only Rf3 and not Rf4 and another two known restorers PNR 3158 and AYT 21 do not possess both $Rf3$ and $Rf4$. The possible explanation could be that IR66, IR40750R may not be very good restorers and/or may not possess

Fig. 3 Amplification pattern of RMS-SF21-5, the candidate gene-specific marker for Rf3 (targeting in-del upstream of SF21) in a set of restorers and non-restorers (i.e., maintainers) in the

PPR9 gene (both PPR9-782-M and PPR9-782-I functional alleles) as RMS-PPR9-1 targets PPR9-782 (M,I) and might possess novel loci other than $Rf3$ and Rf4 for fertility restoration, as Kazama and Toriyama [\(2014](#page-12-0)) reported that other fertility restoration genes could be associated with restoration of WA-CMS.

The process of screening for the trait of fertility restoration is laborious and time-consuming as it involves test crossing with a set of WA-CMS lines followed by evaluation of the F_1 s for pollen and spikelet fertility. Molecular markers targeting the candidate gene associated with the trait are more efficient in accurate identification of restorers among rice germplasm (Sheeba et al. [2009\)](#page-12-0). Recently, our group reported development of a functional marker, targeting the candidate gene, WA352 for WA-CMS trait (Pranathi et al. [2016\)](#page-12-0). However, functional markers for the fertility restoration trait were not available, when this study was initiated and most of the markers available were either linked markers or markers targeting nonvalidated putative candidate genes. To develop candidate gene-specific markers for fertility restoration trait,

above lanes, IR58025A (WA-CMS) and KMR3R (restorer) are used to standard checks

we first attempted to identify candidate genes for Rf4 and Rf3. In a recent study, Tang et al. [\(2014\)](#page-13-0) delineated Rf4 locus to a 137-kb region on chromosome 10 and identified three candidate genes, out of which PPR9- 782-M derived from an elite restorer line Minghui 63 (MH63, with $Rf3$ and $Rf4$) and PPR9-782-I from IR24 was confirmed as a causal gene through complementation assay. Further, the action of Rf4 on WA352 (orf 352) was confirmed by RNA blot analysis. The same study reported two in-del markers (M19288, with a 23-bp in-del and M19280 with a 6-bp in-del). However, it was observed that M19288 displays a dominant fashion of amplification, while M19280 amplified polymorphic fragments from restorers and maintainers in our study. However, the reported primer binding sites (F primer binding site-19,280,871 bp in japonica) of M19280 in-del marker was observed to be not located within the candidate gene *(PPR9* with genomic region from 19,287,680 to 19,295,473 bp) and the 6-bp in-del was located at a distance of 6.8 kb upstream from the gene. Further, as the marker targeted a 6-bp deletion, the polymorphism detected by it was not robust and required a higher percentage of agarose gels $(>3.5 \%)$ for discrimination of the restorer and non-restorer alleles and could not identify many known restorer lines (data not shown), and hence, the marker may not be useful in routine breeding programs. Kazama and Toriyama [\(2014\)](#page-12-0) identified the same gene (i.e., PPR9- 782- M , I through fine-mapping of Rf4 locus that corresponded to a 213-kb region of Nipponbare genome in IR24 cultivar and demonstrated that the fertility restoration is controlled sporophytically. Interestingly, the sequence annotation study of reported candidate genes specific for Rf4, PPR9-782(M,I) and PPR3 (Ngangkham et al. [2010\)](#page-12-0) in Nipponbare genome shows that both genes encode the same 782 amino acidcontaining protein (Os10g0495200 or LOC Os10g35240), where as another reported putative candidate gene, PPR762 encodes a different protein with 762 amino acids (BAD08213; protein sequence alignment file as Supplementary Figure 6). Even though we could observe a 105-bp polymorphism between restorers and non-restorers with respect to PPR9-782-M, on critical analysis of restorer and nonrestorer sequences of the three putative candidate genes reported for Rf4 (mentioned above), we identified that the 105-bp deletion is present in all the three putative candidates PPR762, PPR9-782-(M,I) and PPR3 (alignment file as Supplementary Figure 7). Thus, this polymorphic region may not be unique to a particular candidate gene and hence is not be amenable for development of candidate gene-specific marker. We have identified a unique in-del region of 42 bp within the candidate gene (i.e., PPR9-782-M) (Supplementary Figure 1). Targeting this in-del polymorphism, we designed and validated a codominant marker named RMS-PPR9-1. The marker RMS-PPR9-1 displayed very significant association with the trait phenotype (of fertility restoration) and unequivocally distinguishes almost all the major restorer lines from the nonrestorers of indica rice type (Supplementary Table 2). Thus, RMS-PPR9-1 marker targeting a unique 42-bp in -del within PPR9-782(M,I) gene can be considered as the ideal candidate gene-specific marker for Rf4. Further, we validated the 42-bp in-del polymorphism through sequencing of RMS-PPR9-1 candidate genespecific marker amplicons from IR58025A (WA-CMS) and popular restorer KMR-3R (restorer) lines (Supplementary Figure 8).

Another loci known to be controlling the trait of fertility restoration in WA-CMS system is $Rf3$. Using RAPD and RFLP markers $Rf3$ was earlier mapped on

chromosome 1 (Yao et al. [1997;](#page-13-0) Zhang et al. [1997](#page-13-0)). Till now, there are no reports on cloning and characterization of the candidate gene(s) controlling Rf3 loci. Balaji et al. [2012](#page-12-0) reported a putative candidate, a pollen-specific protein (SF21) encoding gene to be specific for $Rf3$. Targeting the major deletion in the upstream region of SF21 gene, RMS-SF21-5, a codominant, gene-specific marker, was designed and validated in our study. The newly developed marker, RMS-SF21-5, and the earlier reported SSR marker, DRRM-Rf3-10 (Balaji et al. [2012\)](#page-12-0), displayed the same selection efficiency, as they target the same candidate gene, SF21. However, the newly designed marker RMS-SF21-5 was more robust, showing clear polymorphism as compared to DRRM-Rf3-10 (Supplementary Figure 3). Interestingly, when both the gene-specific markers RMS-PPR9-1 (specific for Rf4) and RMS-SF21-5 (specific for $Rf3$), used in conjunction, displayed increased selection efficiency of 94 % as compared to deploying them alone and also as compared to earlier reported markers for fertility restoration trait. The gene-specific markers developed in the study, notably RMS-PPR9-1 has higher efficiency in identifying all true F_1 s hybrids in WA-CMS system (Supplementary Figure 4) and also highly efficient in detection of impurities in hybrid seed lots (Supplementary Figure 5), which was clearly demonstrated in this study through analysis of genetic impurities in a seed lot of the popular hybrid, DRRH3. When the marker RMS-PPR9-1 (specific for Rf4) was validated among 71 Indian accessions of O. nivara and O. rufipogon (Pranathi et al. [2016](#page-12-0)), many accessions displayed amplification of the restorerspecific allele and interestingly, most of the wild rice accessions, which possess wild-abortive cytoplasm and are still fertile, displayed Rf4-specific allele with respect to RMS-PPR9-1 marker indicating the possibility of coevolution of WA-CMS and fertility restoration traits in a few Indian wild rice accessions of O. nivara and O. rufipogon.

In conclusion, through the present study, we identified significant in-del polymorphisms within and around each putative candidate gene specific for Rf4 and Rf3 loci. Through marker–trait co-segregation analysis and higher selection efficiency of genespecific marker targeting in-del polymorphism specific to PPR9-782-M, we confirmed the association of earlier reported gene PRR9-782 (M,I) with Rf4 loci (Tang et al. [2014\)](#page-13-0) and we report the first gene-specific markers for major fertility restoration loci, Rf4 and $Rf3$. The deployment of gene-specific marker for $Rf4$ (RMS-PPR9-1) and another gene-specific marker for Rf3 (RMS-SF21-5) in conjunction displayed higher selection efficiency compared to utilization of genespecific marker alone. Further validation of genespecific markers in F_2 population and germplasm established major influence of Rf4 than Rf3 on the trait. The gene-specific markers, particularly RMS-PPR9-1 marker, can facilitate marker-assisted selection and targeted transfer of $Rf4$ gene into elite backgrounds and also provide highly accurate, rapid detection of impurities in hybrid seed lots. Further efforts are necessary for characterization of candidate gene(s) for Rf3 loci.

Acknowledgments K. Pranathi would like to thank Department of Science and Technology (DST), Government of India, for the INSPIRE fellowship awarded for Ph.D. studies. The authors would also like to thank the Indian Council of Agricultural Research and Department of Biotechnology, Government of India, for the generous funding support for the research work presented in the study.

References

- Ahmadikhah A, Karlov GI (2006) Molecular mapping of the fertility-restoration gene Rf4 for WA-cytoplasmic male sterility in rice. Plant Breed 125:363–367
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
- Balaji S, Vemireddy LR, Srikanth B, Dharika N, Sambasiva Rao KRS, Hemanth Kishore V, Sundaram RM, Viraktamath BC, Subhakara Rao I, Ramesha MS, Neeraja CN (2012) Fine mapping of Rf3 and Rf4 fertility restorer loci of WACMS of rice (Oryza sativa L.) and validation of the developed marker system for identification of restorer lines. Euphytica 187:421–435
- Bharaj TS, Bains SS, Sidhu GS, Gagneja MR (1991) Genetics of fertility restoration of 'Wild Abortive' cytoplasmic male sterility in rice (Oryza sativa L.). Euphytica 56:199–203
- Bharaj TS, Virmani SS, Khush GS (1995) Chromosomal location of fertility restoring genes for wild abortive cytoplasmic male sterility using primary trisomics in rice. Euphytica 83:169–173
- Cai J, Liao QP, Dai ZJ, Zhu HT, Zeng RZ, Zhang ZM, Zhang ZM (2013) Allelic differentiations and effects of the Rf3 and Rf4 genes on fertility restoration in rice with wild abortive cytoplasmic male sterility. Biol Plant 57:274–280
- Cai J, Dai ZJ, Zhu HT, Zeng RZ, Zhang ZM, Ma TF, Zhang GQ (2014) Comparitive analysis of fertility restoaration genes for WA, Y, and DA cytoplasmic male sterility in rice. Biol Plant 58:241–246
- Dellaporta RP, Wood J, Hicks JD (1983) A plant DNA mini preparation: version II. Plant Mol Biol Rep 1:19–21
- Gomez KA, Gomez AA (1984) Statistical procedures for agricultural research, 2nd edn. Wiley, New York
- Govinda Raj K, Virmani SS (1988) Genetics of fertility restoration of WA type cytoplasmic male sterility in rice. Crop Sci 28:787–792
- Hall T (2007) BioEdit version 7.0.9. Carlsbad: Computer Program and documentation, Ibis Biosciences. [http://www.](http://www.mbio.ncsu.edu/BioEdit/bioedit.html) [mbio.ncsu.edu/BioEdit/bioedit.html](http://www.mbio.ncsu.edu/BioEdit/bioedit.html)
- Higgins D, Thompson J, Gibson Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680
- Kazama T, Toriyama K (2014) A fertility restorer gene, Rf4, widely used for hybrid rice breeding encodes a pentatricopeptide repeat protein. Rice 7:28
- Li YC, Yuan LP (1986) Genetic analysis of fertility restoration in male sterile lines of rice. In: IRRI (ed) Rice genetics. Hybrid rice/proceedings of the IRG symposium. IRRI, Manila, pp 617–632
- Lin SC, Yuan LP (1980) Hybrid rice breeding in China. In: IRRI (ed) Innovative approaches to rice breeding. International Rice Research Institute, Manila, pp 35–51
- Nandakumar N, Singh AK, Sharma RK, Mohapatra T, Prabhu KV, Zaman FU (2004) Molecular fingerprinting of hybrids and assessment of genetic purity of hybrid seeds in rice using microsatellite markers. Euphytica 136:257–264
- Nas TMS, Casal CL, Li Z, Virmani SS (2003) Application of molecular markers for identification of restorers. Rice Genet Newsl 20:69–71
- Ngangkham U, Parida SK, De S, Anand Raj Kumar K, Singh AK, Singh SK, Mohapatra T (2010) Genic markers for wild abortive (WA) cytoplasm based male sterility and its fertility restoration in rice. Mol Breeding 26(2):275–292
- Pranathi K, Viraktamath BC, Neeraja CN, Balachandran SM, Hariprasad AS, Koteswara Rao P, Kulkarni SR, Senguttuvel P, Hajira SK, Balachiranjeevi CH, Bhaskar Naik S, Abhilash V, Anila M, Mahadevaswamy HK, Rekha G, Madhav MS, Revathi P, Harika G, Dilip T, Kemparaju B, Sundaram RM (2016) Comparative analysis of sequences of mitochondrial genomes of wild abortive male sterile (WA-CMS) and male fertile lines of rice, development of functional markers for WA-CMS trait and their use in assessment of genetic purity of seeds of WA-CMS lines. Mol Breeding 36:21
- Rajendrakumar P, Biswal AK, Balachandran SM, Ramesha MS, Viraktamath BC, Sundaram RM (2007) A mitochondrial repeat specific marker for distinguishing wild abortive type cytoplasmic male sterile rice lines from their cognate isogenic maintainer lines. Crop Sci 47:207–211
- Sattari M, Kathiresan A, Gregorio GB, Hernandez JE, Nas TM, Virman SS (2007) Development and use of a two-gene markeraided selection system for fertility restorer genes in rice. Euphytica 153:35–42
- Sattari M, Kathiresan A, Glenn B, GregorioS Virmani S (2008) Comparative genetic analysis and molecular mapping of fertility restoration genes for WA, Dissi, and Gambiaca cytoplasmic male sterility systems in rice. Euphytica 160:305–315
- Sheeba NK, Viraktamath BC, Sivaramakrishnan S, Gangashetti MG, Khera P, Sundaram RM (2009) Validation of

molecular markers linked to fertility restorer gene(s) for WA-CMS lines of rice. Euphytica 167:217–227

- Singh AK, Mahapatra T, Prabhu KV, Singh VP, Zaman FU, Mishra GP, Nandakumar N, Joseph M, Gopalakrishnan S, Aparajita G, Tyagi NK, Prakash P, Sharma RK, Shab US, Singh SK (2005) Application of molecular markers in rice breeding: progress at IARI. In: Advances in marker assisted selection workshop. Trainee's manual, Handouts and references
- Sundaram RM, Naveenkumar B, Biradar SK et al (2008) Identification of informative SSR markers capable of distinguishing hybrid rice parental lines and their utilization in seed purity assessment. Euphytica 63:215–224
- Tang H, Luo D, Zhou D, Zhang Q, Tian D, Zheng X, Chen L, Liu YG (2014) The rice restorer Rf4 for wild-abortive cytoplasmic male sterility encodes a PPR protein that functions in reduction of WA352 transcripts. Mol Plant 7:1497–1500
- Teng LS, Shen ZT (1994) Inheritance of fertility restoration for cytoplasmic male sterility in rice. Rice Genet Newsl 11:95–97

Virmani SS (1996) Hybrid rice. Adv Agron 57:377–462

Virmani SS, Wan BH (1988) Development of CMS lines in hybrid rice breeding. In: IRRI (ed) Hybrid rice. International Rice Research Institute, Manila, pp 103–114

- Virmani SS, Raj KG, Casal C, Dalmacio RD and Aurin PA (1986) Current knowledge of and outlook on cytoplasmicgenetic male sterility and fertility restoration in rice. In: IRRI (ed) Rice genetics proceedings of the international rice genetic symposium IRRI, Manila, pp 633–648
- Yao FY, Xu CG, Yu SB, Li JX, Gao YJ, Li XH, Zhang Q (1997) Mapping and genetic analysis of two fertility restorer loci in the wild-abortive cytoplasmic male sterility system of rice (Oryza sativa L). Euphytica 98:183–187
- Yashitola J, Thirumurugan T, Sundaram RM, Naseerullah MK, Ramesha MS, Sarma NP, Sonti RV (2002) Hybrids using microsatellite and STS markers. Crop Sci 42:1369–1373
- Young JB, Virmani SS (1984) Inheritance of fertility restoration in a rice cross. Rice Genet Newsl 1:102–103
- Zhang Q, Bharaj TS, Virmani SS, Huang H (1997) Mapping of the Rf3 nuclear fertility restoring gene for WA cytoplasmic male sterility in rice using RAPD and RFLP markers. Theor Appl Genet 94:27–33
- Zheng KL, Subudhi PK, Dmingo J, Magpantag G, Huang N (1995) Rapid DNA isolation for marker assisted selection in rice breeding. Rice Genet Newsl 12:255–257