Received: 16 October 2017 | Accepted: 6 January 2018

DOI: 10.1111/pbr.12569

ORIGINAL ARTICLE

WILEY Plant Breeding

Fine mapping of *Rf2*, a major locus controlling pollen fertility restoration in sorghum A_1 cytoplasm, encodes a PPR gene and its validation through expression analysis

Praveen Madugula | Anurag G. Uttam | Vilas A. Tonapi | Madhusudhana Ragimasalawada

Marker-assisted selection Laboratory, ICAR-Indian Institute of Millets Research (IIMR), Rajendranagar, Hyderabad 500 030, India

Correspondence
Madhusudhana Ragimasalawada, Marker-assisted selection Laboratory, ICAR-Indian Institute of Millets Research (IIMR), Rajendranagar, Hyderabad 500 030 India.
Email: madhu@millets.res.in

Communicated by: Rod Snowdon

Abstract

Sorghum is one of the pioneering cereal crops where cytoplasmic male sterility (CMS) was successfully exploited for mass production of F_1 hybrid seed. Mapping genes for fertility restoration (*Rf*) is an important aspect of understanding the molecular basis of fertility restoration in crop plants. In this study, we fine-mapped a fertility restoration locus, *Rf2* of sorghum reported earlier (Jordan, Mace, Henzell, Klein, & Klein, 2010), involving two F_2 populations (296A × RS29 and 296A × DSV1) and newly developed SSR markers delimited *Rf2* locus to 10.32-kb region on chromosome 2. The *Rf2* locus was tightly linked with two new SSRs, MS-SB02-3460 (0.14 cM) and MS-SB02-3466 (0.75 cM) on both sides, and hosted only one gene (*Sobic.002G057050*) of PPR gene family. Another new SSR marker developed in the study, MS-SB02-37912, forms the part of PPR gene and could act as a perfect marker in marker-assisted breeding for fertility restoration involving *Rf2* in sorghum breeding. The strong involvement of *Sobic.002G057050* gene in fertility restoration was supported through RNA expression analysis.

KEYWORDS

A_1 cytoplasm, fine mapping, PPR gene, *Rf2* fertility restoration locus, sorghum

1 | INTRODUCTION

Sorghum is one of the pioneering cereal crops where cytoplasmic-genetic male sterility (CMS) is successfully exploited for mass production of F_1 hybrid seed. Sorghum hybrids, with their superiority for grain yield (by 50%–60%) over traditional landraces/pure line varieties, have boosted grain production and profitability all over the world. In India, sorghum production remained static until 1960s with very low grain yields ranging from 400–480 kg/ha. Since 1970s, sorghum production increased dramatically due to the cultivation of F_1 hybrids exploiting heterosis, the commercial exploitation of which became possible following the discovery and use of CMS system (Stephens & Holland, 1954). The commercial success of a F_1 grain hybrid depends chiefly on pollen fertility restoration for full seed set and higher grain yield. Of the available sorghum cytoplasmic sterility systems, A_1 (*milo*) CMS is primarily exploited across the globe. Pollen fertility restoration on A_1 has been reported to be controlled by one or two major genes with few modifiers (Arunkumar, Biradar, & Salimath, 2004; Maunder & Pickett, 1959; Murty & Gangadhar, 1990).

The CMS phenotype is a maternally inherited trait caused by incompatibility between the nuclear and cytoplasmic genomes (Ruiz & Daniell, 2005). Classical genetic studies have reported that CMS is associated with aberrant recombination in the mitochondrial genome (Conde, Pring, Schertz, & Ross, 1982; Das et al., 2010; Eckardt, 2006; Hanson & Bentolila, 2004; Lee, Muthukrishnan, Sorensen, & Liang, 1989; Levings & Pring, 1976; Luo et al., 2013; Pring, Conde, & Schertz, 1982; Smith & Chowdhury, 1989) and fertility was restored by nuclear-encoded fertility restorer (*Rf*) genes in a number of crop species like rice, soybean, maize, sunflower (Ahmadikhah & Karlov, 2006; Dong et al., 2012; Gabay-Laughnan, Chase, Ortega, & Zhao, 2004; Yue, Vick, Cai, & Hu, 2010) including sorghum (Jordan

et al., 2010, 2011; Klein et al., 2001; Praveen, Suneetha, Umakanth, Patil, & Madhusudhana, 2015).

A number of nuclear genes restoring fertility have been cloned in various crop species, and the majority of these genes have been shown to encode pentatricopeptide repeat (PPR) proteins (Bentolila, Alfonso, & Hanson, 2002; Chen & Liu, 2014; Hölzle et al., 2011; Hu et al., 2012; Klein et al., 2005; Koizuka et al., 2003; Wang et al., 2006), which act by specifically suppressing the expression of sterility-causing mitochondrial transcripts (Dahan & Mireau, 2013). So far in sorghum, four *Rf* loci for A_1 cytoplasm have been mapped and the putative candidate genes identified as members of the PPR gene family (Jordan et al., 2010, 2011; Klein et al., 2001; Praveen et al., 2015). The PPR protein family is characterized by the signature motif of a degenerate 35-amino acid repeat often arranged in tandem arrays of 2–27 repeats per peptide (Small & Peeters, 2000). They are thought to be RNA-binding proteins predicted to be targeted to either mitochondria or chloroplast involved in post-transcriptional processes (Lurin et al., 2004).

In this study, we fine-mapped a *Rf* locus on the sorghum chromosome SBI-02 using a $F_{2:3}$ population on 296A × RS29 cross and confirmed the locus using recessive class analysis (RCA) of a small-scale F_2 population of 296A × DSV1. The locus corresponds to *Rf2* previously localized to a 236-kb region of chromosome SBI-02 that was predicted to encode 31 genes; of these genes, only one gene was from the PPR family and this gene showed strong homology to the Rf1 fertility restorer gene in rice (Jordan et al., 2010). This study confirmed the involvement of *Rf2* as a major gene in fertility restoration on a genetically different CMS line. We fine-mapped the *Rf2* locus to a 10.32-kb region on chromosome 02 with only one candidate PPR gene *Sobic*.002G057050 (reported as *Sb02g04810* by Jordan et al. (2010)) flanked by newly developed SSR markers. One of the newly developed SSR marker, MS-SB02-37912, was a part of the PPR gene and could act as a perfect marker for marker-assisted breeding for restorer development in sorghum.

2 | METHODS

2.1 | Genetic material

Mapping population consisted of 235 F_2 plants of 296A × RS29 (P1) cross. 296A is an A_1 cytoplasmic male-sterile line widely used in sorghum hybrid industry in India. Five commercial rainy season hybrids, viz. CSH9, CSH10, CSH11, CSH12 and CSH13 were bred using this CMS line as seed parent. RS29 is the restorer parent of hybrid CSH13 (296A × RS29) developed from SC108-14E × SPV126. Both 296A and RS29 were developed at the ICAR-Indian Institute of Millets Research (IIMR), Hyderabad, India.

2.2 | Phenotyping for fertility restoration

The F_1 (P1:296A × RS29) was planted during the post-rainy season of 2012 at the research farm of IIMR. Before anthesis, F_1 plants were selfed using paper bags to obtain F_2 seeds. During the post-rainy season of 2013, F_2 plants (P1F_2:296A × RS29) were planted as single hill plants in 3-m row plots with a spacing of 15 cm (between plants) and 60 cm (between rows). Plants were protected from insect pests like sorghum shoot fly (*Atherigona soccota*) and stem borer (*Chilo partellus*) using chemical control measures. Standard agronomic practices were followed to raise good crop. Each F_2 plant was labelled, and at their boot-leaf stage, panicle was covered with paper bag to ensure seed set only under self-pollination. At about 25–30 days after anthesis, each F_2 panicle was observed for its seed set, and plants were classified as fertile (selfed seed) or male-sterile (<1% seed set) as suggested (Klein et al., 2001).

To ascertain the genotype of the *Rf* locus of each fertile F_2 plants, selfed seed from each fertile F_2 plant was planted as F_3 family row during post-rainy season of 2014. F_3 progeny from each row were selfed (min. of 20 plants) prior to flowering. F_3 progeny rows with full fertile plants were concluded to have come from F_2 plants homozygous for *Rf* locus (*RfRf*) while those with segregation for sterility/fertility were from F_2 plants with heterozygous *Rf* locus (*Rfrf*).

2.3 | DNA extraction and PCR

Genomic DNA from each labelled F_2 plant was extracted following CTAB extraction method (Murray & Thompson, 1980). The quality of the DNA was analysed by 1% agarose gel electrophoresis, and concentrations were normalized to 25–30 ng/μl. Veriti PCR system (M/s Applied Biosystem, USA) was used for PCR amplification with following buffer composition and thermal profile: about 2 μl (i.e., 50–100 ng) of template DNA was added to the walls of well in a PCR plate. The master mix consisted of 2 mM forward primer, 2 mM reverse primer, 1.25 mM dNTPs (M/s Bangalore Genei Pvt. Ltd.), 1U *Taq* DNA polymerase (M/s Bangalore Genei Pvt. Ltd.), 1× PCR buffer (Tris with 1.5 mM $MgCl_2$), and the total reaction volume was made up to 10 μl using sterile distilled water. Then, the master mix (8.0 μl) was dispensed to the PCR plate containing 2 μl template DNA. The thermal profile followed was as follows: initial denaturation at 94°C for 5 min followed by 35 cycles of denaturation for 1 min at 94°C, primer annealing for 30 s, primer extension for 30 s at 72°C and final extension at 72°C for 10 min. Successfully amplified products were loaded in 4% agarose gel for high-resolving products, and in 6% PAGE gel for low-resolving products.

2.4 | Marker and linkage analysis

To map *Rf* locus involved in the restoration of fertility in the cross of the study (296A × RS29), SSRs of Xtxp, Xgap, Xcup, SB and other marker series available in public domain (Bhattramakki, Dong, Chhabra, & Hart, 2000; Brown et al., 1996; Schloss et al., 2002; Yonemaru et al., 2009) covering roughly the sorghum genome were tested for polymorphism between parental lines (296A and RS29). To identify indicative linkage between marker and *Rf* locus, we followed bulked segregant analysis (BSA) (Michelmore, Paran, & Kesseli, 1991). In BSA, two DNA bulks were prepared using equal amounts

TABLE 1 List of *Rf* loci linked markers used for polymorphism in sorghum cross, 296A × RS29

Locus	CMS	LG	Marker	Polymorphism#	Physical position (bp)	Forward primer	Reverse primer	References
Rf1	A_1	8	Xtxp18*	P	50413700	ACTGTCTAGAACAAGCTGCG	TTGCTCTAGCTAGGCATTTC	Klein et al. (2001, 2005)
			SB4558	M	50606899	CACTACTCGATCTGTCTGCAAGCAA	GTCGTTTTCCTCGGAAAGGAAACT	
			Xtxp400*	M	50656630	CTAAGAACCGACGCGTGTATAGT	GCATCTATCTTCACTCCGATTCT	
			Xisep1236	M	50677610	GGTGACCACTGACGAGATGA	AAACACCAGCTGCTCCAAGT	
			Xtxp406*	M	50689478	GGCCTGAATCTCAGTGTTAAG	AGTTGCCTGCTTCGACACTT	
			SB4560	M	50804880	GCTTCACCTCGAAAGCTCAGACTC	TGACATCCATGTGCTCCCAGTAGT	
			Xtxp63	M	1022590	CCAACCGCGTCGCTGATG	GTGGACTCTGTCGGGGCACTG	
Rf2	A_1	2	Xtxp96	M	1022597	GCTGATGTCATGTTCCCTCAC	CATTCGTGGACTCTGTCGG	Jordan et al. (2010)
			Xtxp197	M	1449033	GCGTCAATTAATCCAAACAGCCTC	GAGTTCCTATTCCCGTTCATGGTGAT	
			Xtxp80	M	3849624	GCTGCACTGTCCTCCCACAA	CAGCAGGCGATATGGATGAGC	
			Xtxp24	M	3910670	CCATTGAGCTTCTGCTATCTC	TTCTAAGCCCACCGAAGTTG	
			Xtxp297	M	4229981	GACCCATATGTGGTTTAGTCGCAAAG	GCACAATCTTCGCCTAAATCAACAAT	
			Unnhsbm152	M	4357799	TTATAGCAGGGGCATCATCC	CATGGAGCAGAAGGTTCCAC	
			Xsdre75	M	4357760	AGAGGCAGCAAAGCGAGAC	ACTGGTGGGAGTCCGTGTAG	
			msbCIR223	P	4654159	CGTTCCAATGACTTTTCTTC	GCCAATGTGGTGTGATAAAT	
			Unnhsbm179	M	4761964	GACGTCAGCAACTGCAAGG	GCTAGCATCCGATCATCACA	
			Xtxp84	M	4850892	CCGATCAGCACACCAG	GTACTAGGTCCAATCCAGC	
			Xtxp211	M	4991669	TCAACGGCCAATGATTTCTAAC	AGGTTGCGAATAAAAGGTAATGTG	
			Xtxp50	M	5076938	TGATGTTGTTACCCTTCTGG	AGCCTATGTATGTGTTCGTCC	
			S11(CNL170)	P	5333906	GAGGGCGTACAGGAAGAACA	CCGAGAAGGACTTGGTGAAG	
			Xtxp654*	-	5428973	TGAGCTGTTAGGGTTTCTCCA	ATGCCTGTGAGCTCTTCCTTC	
			Xisep730	M	5492857	ACCACCACCACCACAACTC	CTCTTGCCCTTGATGGTGAT	
			Unnhsbm1431	M	5492857	GACCACCACCACCACAACTC	AGGAAGATGCTTCCGTCCTC	
			SB1027	M	5494305	GTGCAGCAGTGTAGCAGAGGAAGA	AAGGAGGAAGGAAGAAGAGGAGCA	
			Unnhsbm338	M	5614426	GCGAGACGGTGGAGACTATC	CCCACAAATGACAAACCAAA	
			SB1028	P	5664702	ACTTACGCCTGCAAAGCTACCATC	AAGATCTCACCCAAGCTCCTCCTT	
			Xtxp616*	P	5664933	GCAGACAAGATCTCACCCAAG	GCATTTCTTTCCTGCAATGAC	
			Xtxp304	P	5700017	ACATAAAAGCCCCTCTTC	CTTTCACACCCTTTATTCA	
Rf5	A_1 and A_2	5	Xtxp65	M	1907404	CACGTCGTCACCAACCAAC	GTTAAACGAAAGGGAAATGGC	Jordan et al. (2011)
			SB05-2429989*	-	2430209	TGACTGGTCAACAACGAGGA	CTCTCCCGTGCATGTACTCA	
			Xnhsbm1084	P	2659490	CATTTCACATTCAAGGTCATGG	ACATTTATGGGTGCGTGCTT	

(Continues)

TABLE 1 (Continued)

Locus	CMS	LG	Marker	Polymorphism[#]	Physical position (bp)	Forward primer	Reverse primer	References
			Xnhsbm1100	M	3313831	CGTGAATGAATGAACGAACG	GAGAGCAGAGGGGTAACTGC	
			Xnhsbm1102	P	3477174	CTGAGAACCGGGTCAGTAGC	AAGGCCGAACTTGTAACAGC	
			Xnhsbm1103	P	3558762	GAGTTTAAAACCGCGGACAG	CCATCGCCAACTCCACTACT	
			Xnhsbm1104	M	3645061	CCCTAGTCCACCACCAAGAG	TCCCTCTCCCATATCTCTCTCA	
			SB05-3693459*	-	3645062	CCCTAGTCCACCACCAAGAG	CTCCCATATCTCTCTCAGCACA	
			Xtxp303	M	5686159	AATGAGGAAAATATGAAACAA GTACCAA	AATAACAAGCGCAACTATATG AACAATAAA	
Rf6	A1 and A2	4	Xnhsbm1142	M	351479	CATCCATATCCATGGCAACA	AATGAAGGTGGAGAGGACGA	Praveen et al. (2015)
			SB2385	M	354575	ATCCTCTTCCTCCTCCTCCACCT	TGGTTTGGCTTGTTTACATGTTGC	
			Xnhsbm1143	M	354557	GGTGATGGTGAAGCTTGGTT	AGCCTCCCCTCCATTAAAGA	
			Xnhsbm1144	P	364279	TCCAATGGAATGCTGAATCC	CGCACTTTTCCTTCCTAAAAA	
			SB2386	P	358946	GGCGGTAGGTGTAAAAAGGAAGGA	GCATGCCCTACGACTCTTGTGTCT	
			Unnhsbm66	M	375052	ATACTGCCCACTTGGTTTCG	GGCAACAAGCAGTTGGAAAT	
			SB2387*	M	375093	AAGTTTTGTCACCCGTGCAGATTT	AAGGTCAGTAGCTCGCATGATTCC	
			SB2388*	M	418039	AATTATGGATGCATGGAGCAAAGC	GATGGAGGATCGAGTCACCAAAAC	
			Xnhsbm1157	M	559821	GCTCCCAAATCTCGATAACAA	ACTGGTTTAGGCGGATCCTT	
			Unnhsbm215	M	638023	ACGGACACGGACACGAAG	ACAGGGAGGAGAGGAAGGAA	
			Xnhsbm1173	P	2056932	GAGATCCAGTCGCTGCTAGG	GAACGAACGAACCACCACTT	

P, Polymorphic; M, Monomorphic.

[#]Polymorphism between parents 296A and RS29.

*Reported linked SSR markers.

of genomic DNA from 15 fully fertile (fertile bulk) and 15 fully sterile (sterile bulk) plants of F_2 population. The two bulks (fertile and sterile) and the parents (296A, RS29) were surveyed with polymorphic SSRs.

Positive markers (likely to be linked to the *Rf* locus) from BSA assay and other nearby SSR markers were used for genotyping all the 235F_2 plants. The linkage relationship between markers and *Rf* locus was analysed using Joinmap 3.0 (Ooijen & Voorrips, 2001) to determine the marker order and the map location of the *Rf* locus in the target region on the chromosome. In map construction, *Rf* locus was inserted as a marker based on the phenotypic data for its homozygous (*RfRf*), heterozygous (*Rfrf*) and sterility (*rfrf*) in F_2 and F_3 generations.

With the aim to narrow down the marker interval carrying the *Rf* locus, additional 23 microsatellite markers were developed around the *Rf* locus using the sorghum genome sequence (Paterson et al., 2009) (http://www.phytozome.org). MISA (MIcroSAtellite) Perl script capable of identifying both perfect and compound microsatellites (http://pgrc.ipk-gatersleben.de/misa/) was used to identify the compound microsatellites. Primer pairs flanking microsatellites were designed using Primer3 software (http://frodo.wi.mit.edu/primer3/). The key parameter sets for primer design were as follows: primer length 18–24 bp with 20 bp as the optimum; PCR product size 100–300 bp; optimum annealing temperature 60°C; GC content 35%–60% with 50% as the optimum.

2.5 | Validation of Rf locus

To validate linkage between the *Rf* locus and its linked markers, we further studied two F_2 populations (P2F_2:296A × DSV1; P3F_2: 296A × M35-1) involving the same CMS line, 296A. Restorer DSV1 was developed at the University of Agricultural Sciences, Dharwad, Karnataka, India, and M35-1 is a popular post-rainy sorghum variety widely cultivated in the Indian states of Maharashtra and Karnataka. F_2 populations (P2F_2:183 plants and P3F_2:231 plants) were planted during post-rainy season of 2014 at IIMR. Phenotyping for sterility/fertility reaction and genotyping activities were carried out similar to P1F_2 analysis as detailed above.

Positive markers (linked to the *Rf* locus) from P1F_2 were used for BSA assay in both P2F_2 and P3F_2. Positive markers were used for recessive class analysis (RCA) in P2F_2 only since we did not find any positive marker in BSA of P3F_2. Thirty-nine sterile plants out of 183 P2F_2 plants were assayed with positive markers. RCA using the extreme sterile class was shown to be a reliable approach for mapping fertility restoration locus (Jing, Li, Yi, & Zhu, 2001; Praveen et al., 2015; Yao et al., 1997; Zhang et al., 1994). Recombination frequencies between a linked marker and *Rf* locus were calculated by the maximum-likelihood estimator (Allard, 1956) assuming that all the extremely steriles from both the populations were homozygous at the targeted *Rf* locus.

An additional analysis was conducted to validate the presence of the *Rf* locus identified using BSA and RCA. All the 183 P2F_2 plants were genotyped with positive markers that were likely to be linked with *Rf* locus. Single-marker analysis using linear regression was done. The fertility data in F_2 scored as 1 (fertile) and 2 (sterile) were used for conducting regression analysis and it was performed using the Minitab (http://www.minitab.com) software to detect the association of markers with *Rf* locus. The coefficient of determination (R^2) was used as a measure of the magnitude of association.

2.6 | Fine mapping of Rf locus

For fine mapping of identified *Rf* locus in 296A × RS29, an additional large F_2 population with 1,262 plants were grown during the post-rainy season of 2014. Phenotyping for sterility/fertility reaction and genotyping activities were carried out similar to P1F_2 analysis as detailed above.

2.7 | Quantitative Real-time PCR (qRT-PCR) analysis

Samples of leaf, inflorescence and stem were collected from 296A, RS29, CHS13 (hybrid) at booting stage and immediately frozen in liquid nitrogen for RNA isolation. Total RNA was isolated using SV Total RNA Isolation System (Promega), and the quality of the RNA was assessed using Nanodrop® ND1000 spectrophotometer (Thermo Scientific, USA). Approximately 1 μg of total RNA from each sample was used as template for the first-strand cDNA synthesis, using Verso cDNA Synthesis Kit (Themo Fisher Scientific USA).

Sequence of putative candidate gene *Sobic.002G057050* was obtained from Phytozome (https://phytozome.jgi.doe.gov), and gene-specific primer IIMR-Rf2-Ex2 (forward primer: AATGTCAGCACC-TATGGAAGTATG, reverse primer: ATATGATGATTAGGTTGCACTC CA) was designed using Primer 3.0 (http://simgene.com/Primer3). qRT-PCR was performed using LightCycler 96 (Roche, Switzerland), in a final volume of 25 μl, containing 12.5 μl of Maxima SYBR Green/ROX qPCR Master Mix (2×) (Thermo Scientific, USA) with 500 nM each of forward and reverse primers and 30 ng of the cDNA samples. The real-time PCR cycling conditions included a pre-incubation at 50°C for 2 min and denaturation at 95°C for 10 min followed by 40 cycles of denaturation at 95°C for 15 s and annealing and extension at 60°C for 1 min. qRT-PCR was performed as three biological replicates. Samples were run in duplicates on the same plate along with controls. The data were analysed using the LightCycler 96 SW 1.1 Software (Roche, Switzerland) with default baseline and threshold. The relative expression levels of genes were calculated using the $2^{-\Delta CT\Delta CT}$ method (Livak & Schmittgen, 2001), which represents the difference of CT between the control products and the target gene products. Recommended ubiquitin expression was taken as control.

2.8 | Sequencing and data analysis

Thirteen sequencing primers IIMR-Rf2-S1 to IIMR-Rf2-S13 were designed (Table S1) for amplifying whole gene, and CMS line 296A and both the restorer lines RS29 and DSV1 were PCR-amplified

using these primers. The amplicons were purified with GeneJET PCR Purification Kit (Thermo Scientific EU) and cloned into TA cloning vector using Puregene Gene-TA Cloning Kit (Genetix, USA). After successful cloning and transformation, positive colonies were identified using blue-white colony screening method using Xgal-IPTG in medium. The plasmids were isolated from three positive colonies of each fragment using GeneJET Plasmid Mini Prep Kit (Thermo Scientific EU). The sequencing of inserted fragments was done using universal primer, M13 (Sanger sequencing), by Integrated DNA Technologies, Inc, USA. Homology searches were performed by BLASTn algorithm (Altschul et al., 1997) (http://www.ncbi.nlm.nih.gov/blast). DNA sequences were aligned using the Clustal Omega software (Sievers & Higgins, 2014). Using coding sequence (CDS) of BTx623 (www.phytozome.com) as reference, CDS of 296A, RS29 and DSV1 was identified and analysed using Clustal Omega (Figure S1). Further, peptide sequences of three parents obtained from CDS were compared using Clustal Omega (Figure S2).

3 | RESULTS

3.1 | Segregation of F_2 plants

F_1 plants of all the three crosses (P1, P2 and P3) were fully fertile under selfing suggesting that fertility restoration was dominant over sterility. The $P1F_2$ population segregated into 172 fertile and 63 sterile plants that well fitted with 3:1 ratio (χ^2 (3:1) = 0.41; *p*-value = .522), and $P2F_2$ population segregated into 144 fertile and 39 sterile plants and fitted with 3:1 ratio (χ^2 (3:1) = 1.328; *p*-value = .249). Similarly, the $P3F_2$ population segregated in 172 fertile and 59 sterile, which also confirmed 3:1 ratio (χ^2 (3:1) = 0.036; *p*-value = .849). Thus, the chi-square values of $P1F_2$, $P2F_2$ and $P3F_2$ indicated single gene control of fertility restoration.

In the second F_2 population of 296A × RS29 used for fine mapping of *Rf* locus with 1262 F_2 plants, 957 were fertile and 305 were sterile and fitted with 3:1 ratio (χ^2 (3:1) = 0.466; *p*-value = .495).

3.2 | Markers linked to *Rf* gene

Of the previously mapped *Rf* loci linked markers tested (Table 1) for polymorphism between the parents, 296A and RS29, Xnhsbm1144, Xnhsbm1173 and SB2386 on chromosome 4; Xtxp406 on chromosome 8; msbCIR223, Xtxp304, Xtxp616, SB1028 and SB1031 on chromosome 2; and Xnhsbm1084, Xnhsbm1102 and Xnhsbm1103 on chromosome 5 were polymorphic and were therefore used for BSA. BSA with extreme individuals revealed that of 12 polymorphic markers tested, markers Xtxp304, Xtxp616, SB1028 on chromosome 2 were positive suggesting linkage with *Rf* locus.

3.3 | Linkage map construction

To construct genetic map of *Rf* locus, five polymorphic markers (msbCIR223, Xtxp304, Xtxp616, SB1028, SB1031) were used to construct the local map of *Rf* locus. Linkage analysis indicated that *Rf* locus lies between msbCIR223 (5.33 cM) and SB1028 (0.93 cM). To fine map *Rf* locus to a smaller region, 22 new SSRs were developed from the genomic sequence between msbCIR223 and SB1028 targeting *Rf* locus (Table 2). Of them, two markers (MS-SB02-3460 and MS-SB02-3466) were polymorphic. Recombinant screening with these markers showed that the *Rf* was delimited between the new markers, MS-SB02-3460 (0.14 cM) and MS-SB02-3466 (0.75 cM) (Figure 1).

Towards validation of the presence of an *Rf* locus in the genomic region on chromosome 2, 39 completely sterile plants from $P2F_2$ (296A × DSV1) were assayed individually with each of the polymorphic markers from BSA. Chi-square tests revealed that the segregation of the three genotypes at six linked loci deviated significantly from the expected 1:2:1 ratio (if unlinked), while chi-square test for markers from other chromosomes (Xnhsbm1195 on SBI-04 and Xtxp406 on SBI-08) gave a good fit of 1:2:1 indicating no linkage between these markers and the *Rf* locus (Table 3). At all the linked loci, the genotypes homozygous for the alleles from male-sterile parent, 296A was significantly in excess and the other two genotypes were in great deficiency. Additional single-marker assay with six positive markers and two unlinked markers using all the 183 fertile and sterile individuals of $P2F_2$ also confirmed the existence of the *Rf* locus in the region on chromosome 2 (Table 4). Single-marker analysis indicated significant linkage of positive markers with the restorer locus while the unlinked markers failed to show any linkage. The recombination frequencies between the linked markers and the target *Rf* locus was calculated using 39 steriles of $P2F_2$ using the maximum-likelihood estimator (Table 5). The *Rf* was found to be located very close to SSR marker SB02-3460 with no recombination showing perfect co-segregation with phenotypic data. Thus, BSA, RCA and single-marker analyses of linked and unlinked markers clearly established the presence of an *Rf* locus on chromosome 2 validating the results of P1 (296A × RS29).

4 | THE CANDIDATE GENES IN THE 10.32KB TARGET REGION

There are two annotated genes, *Sobic.002G057050* (Chr02: 5546273..5550944: 4.672 kb) and *Sobic.002G057200* (Chr02:5552957..5555693: 2.737 kb) in the 10.32-kb region (Figure 2), according to the available sequence annotation database (http://www.phytozome.org). *Sobic.002G057050* is functionally annotated as PPR repeat while *Sobic.002G057200* is described as Targeting protein for Xklp2. Sequence analysis of *Sobic.002G057050* showed 60 SNPs between the three parents. Forty-four SNPs were observed between 296A and RS29, 43 SNPs were found between 296A and DSV1 while 22 SNPs were observed between RS29 and DSV1. Furthermore, it also detected a 3-bp deletion in CDS of 296A at 1,283–1,285 position. The changes in the bp also reflected in the change in amino acid sequence. Multiple sequence alignment was performed with the deduced amino acid (aa) sequence of the *Sobic.002G057050* from three parents using the Clustal Omega

TABLE 2 Details of new SSRs designed (MS series) for fine mapping of *Rf2* locus on chromosome 2

S.no	Marker	Gene-based marker	Gene	Polymorphism	Physical position (bp)	Forward primer	Reverse primer
1	Xtxp80	Yes	Sobic.002G039400	M	3839152	GCTGCACTGTCCTCCCACAA	CAGCAGGCGATATGGATGAGC
2	Xtxp297	No		M	4219583	GACCCATATGTGGTTTAGTCGCAAAG	GCACAATCTTCGCCTAAATCAACAAT
3	Xtxp211	No		M	4982191	TCAACGGCCAATGATTTCTAAC	AGGTTGCGAATAAAAGGTAATGTG
4	Xtxp50	Yes	Sobic.002G053350	M	5065678	TGATGTTGTTACCCTTCTGG	AGCCTATGTATGTGTTCGTCC
5	Xtxp654	No		-	5425187	TGAGCTGTTAGGGTTTCTCCA	ATGCCTGTGAGCTCTTCCTTC
6	MS-SB02-3454	No		M	5535990	TGCCATGTTGCTGGATGCTA	ATGCTAGTGTCTGCGTATGC
7	MS-SB02-3455	No		M	5537961	ATCATCTACACCGCGGTCAC	CTCGTGTGTGTGTGTCTGGG
8	MS-SB02-3456	No		M	5538125	CCACAAGGCCACAAACAGC	TGAGGAGAGTCAGCGGAGAG
9	MS-SB02-3464	Yes	Sobic.002G056900	M	5540794	TCCCCAATTTCCTGAGCTGT	TCCTCTTCCAAAGCCTTGCG
10	MS-SB02-3457	Yes	Sobic.002G056900	M	5541142	ATCATGAAGAGTGCCACCCC	AGACAAGACAAGACACGACTCA
11	MS-SB02-3458	Yes	Sobic.002G056900	M	5541142	ATCATGAAGAGTGCCACCCC	TGATTGTTTCCAACACCTCGA
12	MS-SB02-3459	No		M	5544668	CCTTAAGTGGCCCTAATGAGGT	CCGAGTTTTAGGCCGTCCAT
13	MS-SB02-3460	No		P	5545212	CCAAGCACGTGATGAGTCCT	GTGTGGGAGATTGGGATCGG
14	Sobic.002G057050				5546273..5550944		
15	MS-SB02-3465	Yes	Sobic.002G057200	M	5554864	ACCTACCCTTTTCCTCCACA	ACCGCACTAACCAAGCAAGA
16	MS-SB02-3466	Yes	Sobic.002G057200	P	5555530	TTCGACGAAACTCTGCCTGC	AGCTTGGCTTGGGTTTTCCT
17	Xtxp616	Yes	Sobic.002G058500	P	5659089	GCAGACAAGATCTCACCCAAG	GCATTTCTTTCCTGCAATGAC
18	Xtxp304	No		P	5694187	ACATAAAAGCCCCTCTTCT	CTTTCACACCCTTTATTCA

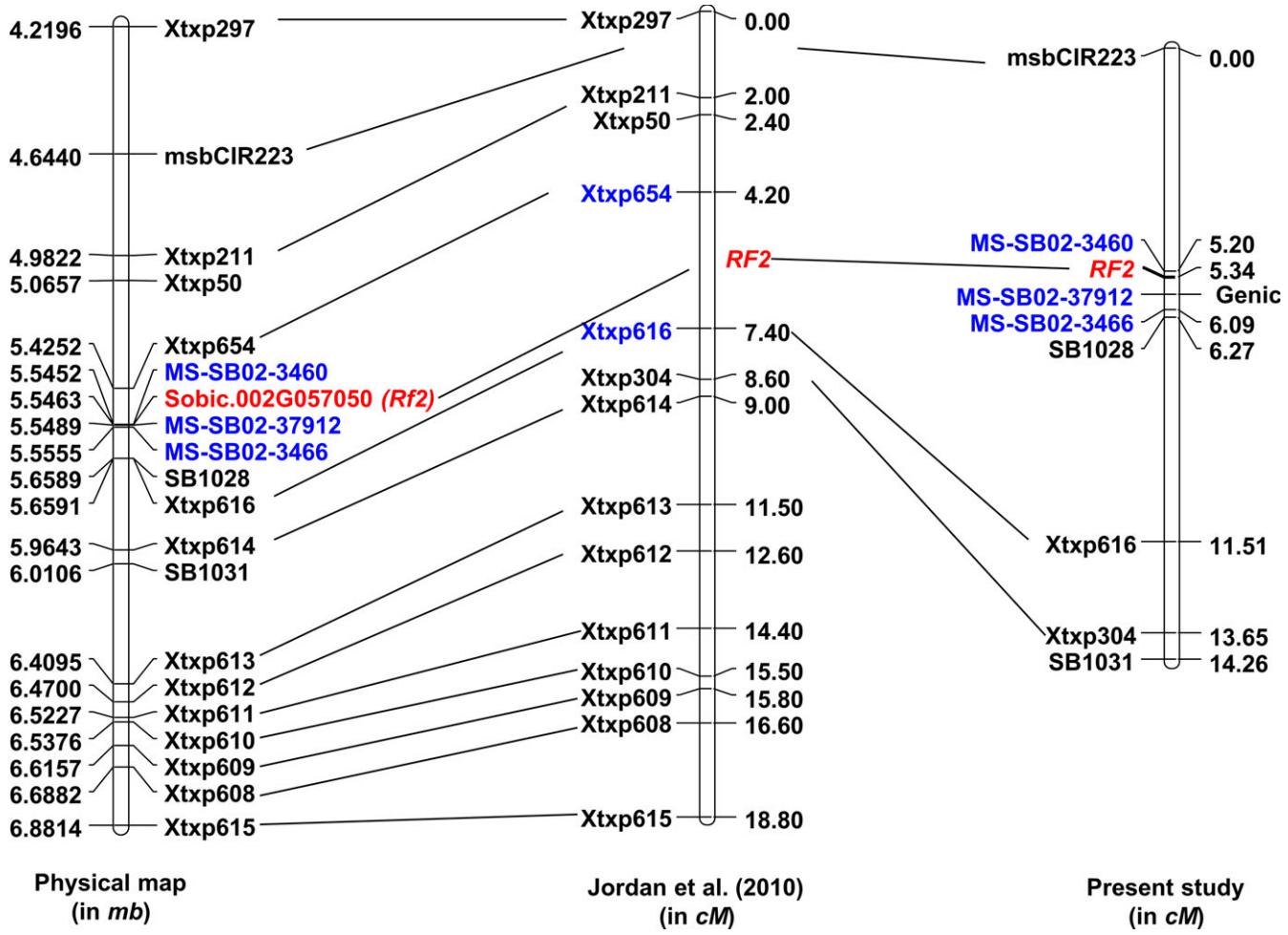

FIGURE 1 Map position of Rf2 locus in sorghum in reference to physical and linkage maps [Colour figure can be viewed at wileyonlinelibrary.com]

program. As shown in Figure S2, aa acid sequence of *Sobic.002G057050* protein shares high levels of identity except at 8 aa positions.

Using qRT-PCR, the expression levels of *Sobic.002G057050* were studied in leaf, stem and inflorescence tissues of male-sterile parent (296A), restorer parent (RS29) and their hybrid (CSH13). As shown in Figure 3, the *Sobic.002G057050* gene was highly expressed in the inflorescence, but was hardly detected in the leaf and stem during the preheading stage in all the three parents. Interestingly, around 12× increase of *Sobic.002G057050* expression was observed in inflorescence tissue of restorer parent (RS29) compared to sterile parent 296A, while gene expression was midway (5×) in hybrid (296A × RS29 = hybrid CSH13) as compared to male-sterile parent (296A).

5 | DISCUSSION

Fertility restoration is an important trait in F_1 hybrid as it determines seed set and grain yield. Unravelling the genetics of fertility restoration and location of *Rf* genes for a CMS system is an important plant breeding activity as it enhances the efficiency of selection for restorer parents through marker-assisted breeding. Several CMS systems have been described in sorghum (Schertz, Sotomayor-Rios, & Torres-Cardona, 1989), of which A_1 (*milo*) system is primarily used in sorghum hybrid industry across the globe. The advantage of A_1 is that the frequency of fertility restorers among germplasm lines is high compared to alternative CMS systems (Bosques-Vega, Sotomayor-Rios, Torres-Cardona, Hepperly, & Schertz, 1989; Reddy, Rai, Sarma, Kumar, & Saxena, 2003; Senthil, Ramasamy, & Fazlullah Khan, 1998), which in turn provides greater choice of selecting a better and diverse restorer for hybrid development. The genetics of fertility restoration on A_1 CMS system is fairly simple and controlled by one to two *Rf* genes (Arunkumar et al., 2004; Maunder & Pickett, 1959; Murty & Gangadhar, 1990). In the present study, F_2 plants of three crosses segregated with a monogenic ratio of 3 (fertile): 1 (sterile), which suggested existence of single gene difference in the expression of restoration of fertility between the fertile parents (RS29, DSV1 and M35-1) and the recessive sterile parent, 296A. Thus, this gene shows a similar pattern of inheritance to all previously identified *Rf* genes including this loci as previously mapped by Jordan et al. (2010).

Cytoplasmic male sterility (CMS) is exploited for mass production of hybrid seed in a wide variety of crops such as maize, sorghum,

TABLE 3 Deviation of marker segregation from 1:2:1 ratio in 39 sterile individuals of 296A × DSV1F_2 population

SSR Primer	LG	N1 (9.75)	N2 (19.5)	N3 (9.75)	chi-square	*p*-value
Linked						
msbCIR223	2	2	2	35	87.256	1.13E-19
MSSB02-3460	2	0	0	39	117.00	3.92E-26
MSSB02-3466	2	1	2	36	94.2308	3.45E-21
SB1028	2	3	2	34	80.6923	3.01E-18
Xtxp616	2	3	2	34	80.6923	3.01E-18
Xtxp304	2	3	3	33	74.0769	8.21E-17
Unlinked						
Xnhsbm1195	4	9	20	10	0.071	.962
Xtxp406	8	7	22	10	1.10256	.576

N_1, Sterile individuals with restorer band (B) expected if unlinked is 10; **N_2**, number of sterile individuals heterozygous for the SSR bands from both restorer and male-sterile parents (H) expected if unlinked is 20; **N_3**, sterile individuals with male-sterile band (A) expected if unlinked is 10.

TABLE 5 Recombination frequencies of 39F_2 sterile individuals of 296A × DSV1 with linked markers

SSR Primer	LG	N1	N2	R	C	Rec
msbCIR223	2	2	2	2	0.051	5.128
MS3460	2	0	0	0	0.000	0.000
MS3466	2	1	2	1.5	0.038	3.846
SB1028	2	3	2	2.5	0.064	6.410
Xtxp616	2	3	3	3	0.077	7.692
Xtxp304	2	3	4	3.5	0.090	8.974

N_1, Number of sterile individuals homozygous for the SSR band from the fertile restorer parent band (B). N_2, Number of sterile individuals heterozygous for the SSR bands from both restorer and male-sterile parent (H). R, Number of recombinants ($R = N_1 + N_2/2$). C, Recombination frequency C = R/N where *N* is the total no of sterile plants surveyed. Rec., % Recombination.

rice, sunflower, cotton, rapeseed, carrot. CMS is a maternally inherited trait and is often associated with unusual open-reading frames (ORFs) found in mitochondrial genomes, and in many instances, male fertility can be restored specifically by nuclear-encoded, fertility restorer (*Rf*) genes (Schnable & Wise, 1998). To date in sorghum, five *Rf* loci have been described (Jordan et al., 2010, 2011; Klein et al., 2001; Praveen et al., 2015; Wen et al., 2002) along with their chromosomal locations. The first major gene described to restore fertility on A1 cytoplasm, Rf1, was mapped on SBI-08 (Klein et al., 2001). Sequence analysis of the Rf1 locus identified a PPR protein that co-segregated with the fertility restoration phenotype (Klein et al., 2005). The Rf2 on A1 was mapped on SBI-02 along with 31 candidate genes (Jordan et al., 2010). Of the two Rf3 and Rf4 loci associated with fertility restoration in A3 cytoplasm, Rf4 was mapped on to SBI-01 using AFLP-derived CAPS markers (Wen et al., 2002). Rf5 and Rf6 were found to restore male fertility on both A1 and A2 CMS systems (Jordan et al., 2011; Praveen et al., 2015) and were localized on SBI-05 and SBI-04 respectively with PPR genes as candidate *Rf* genes.

In this study, we mapped a major fertility restoration (*Rf*) locus of A1 CMS delimited to a region of 10.32 kb on chromosome 2 and was tightly linked with newly developed two SSRs, MS-SB02-3460 and MS-SB02-3466. While comparing the map positions of the *Rf* with previous ones, we found that the present *Rf* locus was previously reported as Rf2 between Xtxp654- and Xtxp616-flanked markers (Jordan et al., 2010). However, in their study, Rf2 locus represented a genomic region of 236.219 kb on chromosome 2, which hosted 31 genes including a PPR gene family member, Sb02g004810 (Jordan et al., 2010). However, further efforts to narrow down Rf2 locus using new markers to identify the probable *Rf* candidate gene of the 31 genes were unsuccessful as mapping parents were highly homologous at this region (Jordan et al., 2010). Therefore, fine mapping of Rf2 locus using new markers and diverse mapping parents was necessary to pin point the gene involved in fertility restoration of Rf2 locus. The gene Sb02g004810 (Sbi.v1.4) identified, by Jordan et al. (2010), is the same gene identified in this

TABLE 4 Single-marker analysis of linked and unlinked polymorphic markers for *Rf2* locus in 296A x DSV1F2 population

Locus	LG	F mean	S Mean	MS effect	Ms error	*df* error	F	*p*	R^2 (%)
Linked markers									
msbCIR223	2	2.59	1.15	63.93	0.30	181	207.00	.000	52.48
MS3460	2	2.59	1.00	78.29	0.25	181	303.83	.000	62.14
MS3466	2	2.51	1.12	58.92	0.33	181	176.78	.000	48.43
SB1028	2	2.51	1.17	54.64	0.35	181	155.23	.000	45.09
Xtxp616	2	2.56	1.23	54.66	0.38	180	140.49	.000	42.64
Xtxp304	2	2.57	1.28	51.10	0.40	180	126.22	.000	39.88
Unlinked markers									
Xnhsbm1195	4	2.19	2.25	0.11	0.72	181	0.16	.688	0.00
Xtxp406	8	2.15	2.30	0.67	0.72	181	0.92	.330	0.00

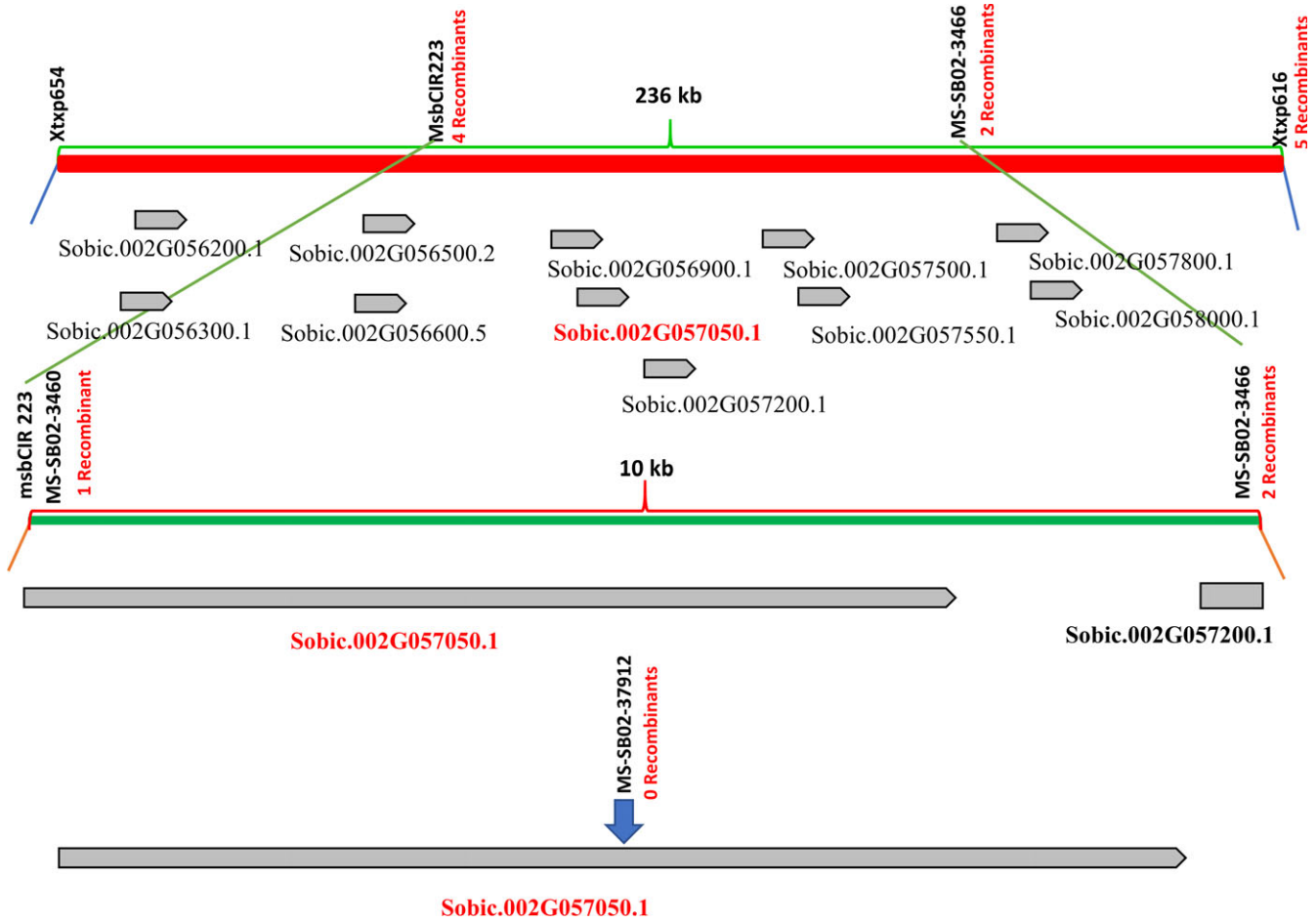

FIGURE 2 A fine-scale, high-resolution map of the Rf2 locus on SBI-02 indicating involvement of PPR gene Sobic.002G057050 in fertility restoration [Colour figure can be viewed at wileyonlinelibrary.com]

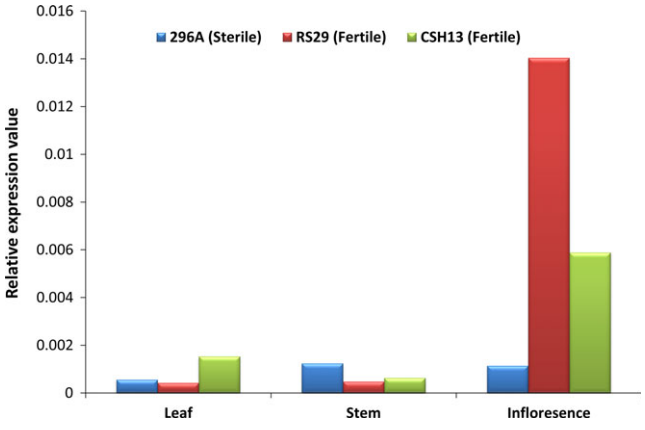

FIGURE 3 Expression of PPR gene Sobic.002G057050. Graph showing differential expression of PPR gene Sobic.002G057050 in three different tissues (leaf, stem and inflorescence) of three sorghum genotypes. 296A (sterile), RS29 (Restorer) and CSH13 (F1 hybrid of 296A × RS29) [Colour figure can be viewed at wileyonlinelibrary.com]

study and had strong evidence to support it. A BLASTn search using the nucleotide sequence of the gene, Sb02g004810, had shown 100% identity to Sobic.002G057050 (v3.1.1). Therefore, we treated Sobic.002G057050 as the refined sequence of the PPR gene, Sb02g004810. The gene identified by Jordan et al. (2010) came from SC170C, a Zera zera sorghum. RS29 has the pedigree SC108-14E/SPV126 is also a Zera zera type from the same region of Ethiopia. It is likely that the alleles in SC170 and RS29 are the same, as RS29 is a derivative of SC108.

PPR proteins are RNA-binding proteins with a range of essential functions in post-transcriptional processes (including RNA editing, RNA splicing, RNA cleavage and translation) within mitochondria and chloroplasts (Schmitz-Linneweber & Small, 2008). PPR proteins harbour between two and 30 PPR motifs, and their tandem alignment allows the modular recognition of specific RNA sequences, generally function in chloroplast or mitochondria (Fujii, Sato, & Shikanai, 2013). Most fertility restoration genes reported so far in a range of crop species belong to PPR gene family (Chen & Liu, 2014) except for aldehyde dehydrogenase (Rf2) in maize (Cui, Wise, & Schnable, 1996), glycine-rich protein (Rf2) in rice (Itabashi, Iwata, Fujii, Kazama, & Toriyama, 2011), unknown protein (Rf17) in rice (Fujii & Toriyama, 2009) and a yeast Oma1-like peptidase (Rf1) in sugar beet (Matsuhira et al., 2012). *Rf* locus tends to occur in clusters of PPR genes as observed in species including rice (Kato et al., 2007), maize (Xu

TABLE 6 Details of new *Sobic.002G057050* genic SSRs

S.no	Marker	Gene-based marker	Gene	Polymorphism	Physical position (bp)	Forward primer	Reverse primer
1	MS-SB02-37907	No		M	5545212	CCAAGCACGTGATGAGTCCT	GGTGGGAAATGCTGCTCACA
2	MS-SB02-37908	No		M	5545468	CTGTGAGCAGCATTTCCCAC	GATGAACTGCTCCAGCTCCC
	Sobic.002G057050				5546273..5550944		
3	MS-SB02-37909	Yes	Sobic.002G057050	M	5546445	CAATGGACAAGGCCGAGGAG	TAAGTGGCGGCATTGGGTTT
4	MS-SB02-37911	Yes	Sobic.002G057050	M	5548527	CCCCAGAGGCCTACATGATG	GCAGTCCCAGCCATTCAGTT
5	**MS-SB02-37912**	**Yes**	**Sobic.002G057050**	**P**	**5548850**	**GCCAGAGCAATCGTGACACA**	**CCAGGAAGGGCGGTGGTCG**
6	MS-SB02-37913	Yes	Sobic.002G057050	M	5549467	GTTTACGCCCTCAACCAGCT	ATCTGTCCTCAGCCCTGTCT
7	MS-SB02-37914	Yes	Sobic.002G057050	M	5549856	GCCGAGAAGGTCTTTCGACA	CCATAGGTGCTGACATTGGGT
8	MS-SB02-3461	No		M	5552599	GCCACCATTACCCAGGGAAA	TCTTTCCCCATTGCACTGCA
9	MS-SB02-37915	Yes	Sobic.002G057200	M	5553344	GGATGTTGGCCCTCGATCTT	CACACCTCTGAAGCAGCAGA
10	MS-SB02-3462	Yes	Sobic.002G057200	M	5553697	TACAAATGGCCAGAGCAGGG	GAGACGTTCCCCACCTGTTT
11	MS-SB02-3463	Yes	Sobic.002G057200	M	5553906	AAACAGGTGGGGAACGTCTC	CTGCTGTCCGTTTGCTGGT
12	MS-SB02-37916	Yes	Sobic.002G057200	M	5554752	GGCAACATGAAGAACACCCA	ACTGATCGTTTCCAAAACCTCG
13	MS-SB02-3465	Yes	Sobic.002G057200	M	5554864	ACCTACCCTTTTCCTCCACA	ACCGCACTAACCAAGCAAGA

et al., 2009), Petunia (Bentolila et al., 2002), Brassica (Brown et al., 2003), Arabidopsis (Lurin et al., 2004). In sorghum, five PPR genes, viz. Sobic.008G099900, Sobic.008G100400, Sobic.008G147400, Sobic.008G149200, Sobic.008G163400, are linked within 12 Mb of Rf1 locus. Similarly, Rf5 locus hosts Sobic.005G027600, Sobic.005G027840, Sobic.005G028200, Sobic.005G028300 and Sobic.005G030850 genes in a region of 0.32 Mb. Rf2 locus contains tightly linked five PPR genes, viz. Sobic.002G054100, Sobic.002G054200, Sobic.002G055300, Sobic.002G057050 and Sobic.002G059700 with in 0.58 Mb genomic region. Kato et al. (2007) postulated that PPR-containing duplicated genes have arisen through evolutionarily recent gene duplication events similar to those that have been observed with disease resistance loci. Recent pan-genomic analysis of crops has shown that copy number variation can occur between individuals of the same species particularly in areas of gene duplication (Springer et al., 2009; Wang et al., 2015). With most fertility restoration genes being PPR gene family members, it can be postulated that the Rf2 restoration phenotype in sorghum is conditioned by one or multiple PPR genes residing within the locus.

Refinement of Rf2 locus in the present study using new SSR markers in a different genetic cross was highly successful as it narrowed down to a small genomic region of 10.32 kb on chromosome 2 delimited by two SSR markers. The region between these two SSRs, MS-SB02-3460 and MS-SB02-3466, hosted only two genes, *Sobic.002G057050* and *Sobic.002G057200*. *Sobic.002G057050* is functionally annotated as PPR repeat while Sobic.002G057200 is described as Targeting protein for Xklp2. Thus, only one gene of PPR family, *Sobic.002G057050* is available in the marker intervals. The gene *Sobic.002G057050* structurally characterized with two exons (exon1:1268 bp and exon2:1996 bp) and one intron (1408 bp). The online program MitoProt II v1.101 (Claros & Vincens, 1996) was used to predict the subcellular localization of both proteins, of which only PPR protein was found to target mitochondria (p = .94–.96). We also used Pfam 27.0 (http://pfam.sanger.ac.uk/) to detect the conserved motifs of both gene proteins, of which only the *Sobic.002G057050* was detected with eight PPR motifs. The gene has high homology to the Rf1 protein of rice (*LOC_Os10g35240*), maize (*GRMZM2G393935*) and foxtail millet (*Seita.9G204800*) *Rf* genes with same functional annotation.

To additionally prove *Sobic.002G057050* as the most probable *Rf* gene, we designed 13 SSRs (Table 6) covering the gene, of which SSR marker MS-SB02-37912 (forward primer—GCCAGAGCAA TCGTGACACA and reverse primer—CCAGGAAGGGCGGTGGTCG) present within the PPR gene was polymorphic between the parents. Our analysis involving MS-SB02-37912 in a large population of 1262F_2 plants (of 296A × RS29) showed perfect co-segregation of marker data with phenotypic sterility/fertility data without any recombinational events. Further validation of *Rf* gene came from qRT-PCR analysis of *Sobic.002G057050* wherein its expression was significantly high in inflorescence compared to vegetative tissues like leaf and stem (Figure 3). As expected, its expression in restorer parent RS29 (*RfRf*) was highest compared to CMS parent 296A (rfrf),

and it was moderate in hybrid CSH13 (Rfrf). The SSR MS-SB02-37912 is Rf2 gene-based perfect marker, which could be used in the marker-assisted breeding of restorer line development on A1 CMS in sorghum. Three base pair deletions (one aa), and other transitions and transversions in the CDS region of the CMS parent 296A and restorers RS29, DSV1 in *Sobic.002G057050*, could be the reason for malfunctioning of the protein in the CMS parent.

Of the three F_2 crosses we studied, the Rf2 locus was found to operate only in two (296A × RS29 and 296A × DSV1). The M35-1, a restorer on 296A (296A × M35-1), might possess a novel *Rf* locus as all the previously reported *Rf* loci (Rf1, Rf5 and Rf6 that restore fertility on A1 cytoplasm) linked markers did not show any linkage with fertility restoration in this population. This is similar to involvement of different genes in the fertility restoration of single WA-CMS in rice (Kazama & Toriyama, 2014; Pranathi et al., 2016). Further studies are required to map this new *Rf* locus involved in the restoration of fertility on A1.

6 | CONCLUSIONS

The present study facilitated the fine mapping of Rf2 locus of sorghum and identified a PPR gene *Sobic.002G057050* as the most probable fertility restorer gene confirming the earlier report on Sb02g004810 as a strong candidate (Jordan et al., 2010). Fine mapping of Rf2 locus creates a basis for cloning of this gene and understanding its regulation. In hybrid breeding exploiting CMS, it is vital to select sterile inbreds with complete sterility, and the restorer lines with highest restoration ability. The availability of genic SSR marker, MS-SB02-37912, facilitates the efficient screening of germplasm lines to identify promising fertility restorers on A_1 CMS as conventional test-crossing is resource- and time-consuming, thereby limiting the selection gain in the restorer pool. The marker is also useful to carry out marker-assisted pyramiding of *Rf* genes into different hybrid restorer parents. Compared with earlier reported Rf2 linked markers (Jordan et al., 2010), new SSR markers of the study are more closely linked and are highly efficient for marker-assisted breeding. Like in rice (Garg et al., 2006), MS-SB02-37912 could be employed for seed purity test of sorghum hybrid seed as an alternative to conventional grow-out test, which is time-consuming.

ACKNOWLEDGMENTS

We thank the Director, ICAR-IIMR, Rajendranagar, Hyderabad, for the facilities to undertake this study. We sincerely thank Dr. M Sheshu Madav, ICAR-IIRR, for extending the facility for gene expression analysis, and M Bhaskar for the support in conducting field experiments. We sincerely acknowledge the Anonymous reviewers for their suggestions which improved the quality of the manuscript.

CONFLICT OF INTERESTS

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTION

RM designed the experiments; PM and AU conducted the experiments; RM analysed the data and wrote the manuscript. VA gave suggestion in manuscript preparation. All authors approved the final version of the manuscript.

ETHICAL STANDARDS

The authors declare no ethical standards have been violated in the course of the study.

ORCID

Madhusudhana Ragimasalawada http://orcid.org/0000-0001-7316-0986

REFERENCES

Ahmadikhah, A., & Karlov, G. (2006). Molecular mapping of the fertility-restoration gene Rf4 for WA-cytoplasmic male sterility in rice. *Plant Breeding*, *125*, 363–367. https://doi.org/10.1111/j.1439-0523.2006.01246.x

Allard, R. (1956). Formulas and tables to facilitate the calculation of recombination values in heredity. *Hilgardia*, *2*, 235–278. https://doi.org/10.3733/hilg.v24n10p235

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic acids research*, *25*, 3389–3402. https://doi.org/10.1093/nar/25.17.3389

Arunkumar, B., Biradar, B., & Salimath, P. (2004). Inheritance of fertility restoration on milo and maldandi sources of male sterility in rabi sorghum [Sorghum bicoJor (L.) Moench]. *The Indian Journal of Genetics and Plant Breeding*, *64*, 325–326.

Bentolila, S., Alfonso, A. A., & Hanson, M. R. (2002). A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. *Proceedings of the National Academy of Sciences*, *99*, 10887–10892. https://doi.org/10.1073/pnas.102301599

Bhattramakki, D., Dong, J., Chhabra, A. K., & Hart, G. E. (2000). An integrated SSR and RFLP linkage map of Sorghum bicolor (L.) Moench. *Genome*, *43*, 988–1002. https://doi.org/10.1139/g00-074

Bosques-Vega, A., Sotomayor-Rios, A., Torres-Cardona, S., Hepperly, P., & Schertz, K. (1989). Maintainer and restorer reactions with A1, A2, and A3 cytoplasms of lines from the sorghum conversion program. Miscellaneous publication MP-Texas Agricultural Experiment Station (USA).

Brown, G. G., Formanová, N., Jin, H., Wargachuk, R., Dendy, C., Patil, P., … Landry, B. S. (2003). The radish Rfo restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. *The Plant Journal*, *35*, 262–272. https://doi.org/10.1046/j.1365-313X.2003.01799.x

Brown, S., Hopkins, M., Mitchell, S., Senior, M., Wang, T., Duncan, R., … Kresovich, S. (1996). Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [Sorghum bicolor (L.) Moench]. *Theoretical and Applied Genetics*, *93*, 190–198. https://doi.org/10.1007/BF00225745

Chen, L., & Liu, Y.-G. (2014). Male sterility and fertility restoration in crops. *Annual review of Plant Biology*, *65*, 579–606. https://doi.org/10.1146/annurev-arplant-050213-040119

Claros, M., & Vincens, P. (1996): MITOPROT: Prediction of mitochondrial targeting sequences. Institut für Humangenetik, Technische

Universität, München and Institut für Humangenetik GSF, Forschungszentrum für Gesundheit und Umwelt, Neuherberg. http://ihg. gsf. de/ihg/mitoprot. html.

Conde, M., Pring, D., Schertz, K., & Ross, W. (1982). Correlation of mitochondrial DNA restriction endonuclease patterns with sterility expression in six male-sterile sorghum cytoplasms. *Crop Science*, *22*, 536–539. https://doi.org/10.2135/cropsci1982.0011183X002200030023x

Cui, X., Wise, R. P., & Schnable, P. S. (1996). The rf2 nuclear restorer gene of male-sterile T-cytoplasm maize. *Science*, *272*, 1334. https://doi.org/10.1126/science.272.5266.1334

Dahan, J., & Mireau, H. (2013). The Rf and Rf-like PPR in higher plants, a fast-evolving subclass of PPR genes. *RNA Biology*, *10*, 1469–1476. https://doi.org/10.4161/rna.25568

Das, S., Sen, S., Chakraborty, A., Chakraborti, P., Maiti, M. K., Basu, A., … Sen, S. K. (2010). An unedited 1.1 kb mitochondrial orfB gene transcript in the wild abortive cytoplasmic male sterility (WA-CMS) system of *Oryza sativa* L. subsp. indica. *BMC Plant Biology*, *10*, 39. https://doi.org/10.1186/1471-2229-10-39

Dong, D., Li, Z., Yuan, F., Zhu, S., Chen, P., Yu, W., … Li, B. (2012). Inheritance and fine mapping of a restorer-of-fertility (Rf) gene for the cytoplasmic male sterility in soybean. *Plant Science*, *188*, 36–40. https://doi.org/10.1016/j.plantsci.2012.02.007

Eckardt, N. A. (2006). Cytoplasmic male sterility and fertility restoration. *The Plant Cell Online*, *18*, 515–517. https://doi.org/10.1105/tpc.106.041830

Fujii, S., Sato, N., & Shikanai, T. (2013). Mutagenesis of individual pentatricopeptide repeat motifs affects RNA binding activity and reveals functional partitioning of Arabidopsis PROTON gradient regulation3. *The Plant Cell*, *25*, 3079–3088. https://doi.org/10.1105/tpc.113.112193

Fujii, S., & Toriyama, K. (2009). Suppressed expression of Retrograde-Regulated Male sterility restores pollen fertility in cytoplasmic male sterile rice plants. *Proceedings of the National Academy of Sciences*, *106*, 9513–9518. https://doi.org/10.1073/pnas.0901860106

Gabay-Laughnan, S., Chase, C. D., Ortega, V. M., & Zhao, L. (2004). Molecular-genetic characterization of CMS-S restorer-of-fertility alleles identified in Mexican maize and teosinte. *Genetics*, *166*, 959–970. https://doi.org/10.1534/genetics.166.2.959

Garg, A., Singh, A. K., Prabhu, K. V., Mohapatra, T., Tyagi, N. K., Nandakumar, N., … Zaman, F. U. (2006). Utility of a fertility restorer gene linked marker for testing genetic purity of hybrid seeds in rice (*Oryza sativa* L.). *Seed Science and Technology*, *34*, 9–18. https://doi.org/10.15258/sst

Hanson, M. R., & Bentolila, S. (2004). Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *The Plant Cell Online*, *16*, S154–S169. https://doi.org/10.1105/tpc.015966

Hölzle, A., Jonietz, C., Törjek, O., Altmann, T., Binder, S., & Forner, J. (2011). A restorer of fertility-like PPR gene is required for 5′-end processing of the nad4 mRNA in mitochondria of Arabidopsis thaliana. *The Plant Journal*, *65*, 737–744. https://doi.org/10.1111/j.1365-313X.2010.04460.x

Hu, J., Wang, K., Huang, W., Liu, G., Gao, Y., Wang, J., … Wan, L. (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. *The Plant Cell*, *24*, 109–122. https://doi.org/10.1105/tpc.111.093211

Itabashi, 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. *The plant journal*, *65*, 359–367. https://doi.org/10.1111/j.1365-313X.2010.04427.x

Jing, R., Li, X., Yi, P., & Zhu, Y. (2001). Mapping fertility-restoring genes of rice WA cytoplasmic male sterility using SSLP markers. *Botanical Bulletin of Academia Sinica*, *42*, 167–171.

Jordan, D. R., Klein, R. R., Sakrewski, K. G., Henzell, R. G., Klein, P. E., & Mace, E. S. (2011). Mapping and characterization of Rf 5: A new gene conditioning pollen fertility restoration in A1 and A2 cytoplasm in sorghum (*Sorghum bicolor* (L.) Moench). *Theoretical and Applied Genetics*, *123*, 383–396. https://doi.org/10.1007/s00122-011-1591-y

Jordan, D., Mace, E. S., Henzell, R., Klein, P., & Klein, R. (2010). Molecular mapping and candidate gene identification of the Rf2 gene for pollen fertility restoration in sorghum [*Sorghum bicolor* (L.) Moench]. *Theoretical and Applied Genetics*, *120*, 1279–1287. https://doi.org/10.1007/s00122-009-1255-3

Kato, H., Tezuka, K., Feng, Y., Kawamoto, T., Takahashi, H., Mori, K., & Akagi, H. (2007). Structural diversity and evolution of the Rf-1 locus in the genus Oryza. *Heredity*, *99*, 516–524. https://doi.org/10.1038/sj.hdy.6801026

Kazama, T., & Toriyama, K. (2014). A fertility restorer gene, Rf4, widely used for hybrid rice breeding encodes a pentatricopeptide repeat protein. *Rice*, *7*, 1.

Klein, R., Klein, P., Chhabra, A., Dong, J., Pammi, S., Childs, K., … Schertz, K. (2001). Molecular mapping of the rf1 gene for pollen fertility restoration in sorghum (*Sorghum bicolor* L.). *Theoretical and Applied Genetics*, *102*, 1206–1212. https://doi.org/10.1007/s001220100575

Klein, R. R., Klein, P. E., Mullet, J. E., Minx, P., Rooney, W. L., & Schertz, K. F. (2005). Fertility restorer locus Rf1 of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. *Theoretical and Applied Genetics*, *111*, 994–1012. https://doi.org/10.1007/s00122-005-2011-y

Koizuka, N., Imai, R., Fujimoto, H., Hayakawa, T., Kimura, Y., Kohno-Murase, J., … Imamura, J. (2003). Genetic characterization of a pentatricopeptide repeat protein gene, orf687, that restores fertility in the cytoplasmic male-sterile Kosena radish. *The Plant Journal*, *34*, 407–415. https://doi.org/10.1046/j.1365-313X.2003.01735.x

Lee, S., Muthukrishnan, S., Sorensen, E., & Liang, G. (1989). Restriction endonuclease analysis of mitochondrial DNA from sorghum with fertile and male-sterile cytoplasms. *Theoretical and Applied Genetics*, *77*, 379–382. https://doi.org/10.1007/BF00305832

Levings, C., & Pring, D. (1976). Restriction endonuclease analysis of mitochondrial DNA from normal and Texas cytoplasmic male-sterile maize. *Science*, *193*, 158–160. https://doi.org/10.1126/science.193.4248.158

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods*, *25*, 402–408. https://doi.org/10.1006/meth.2001.1262

Luo, D., Xu, H., Liu, Z., Guo, J., Li, H., Chen, L., … Yao, N. (2013). A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. *Nature Genetics*, *45*, 573–577. https://doi.org/10.1038/ng.2570

Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., … Small, I. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, *16*, 2089–2103. https://doi.org/10.1105/tpc.104.022236

Matsuhira, H., Kagami, H., Kurata, M., Kitazaki, K., Matsunaga, M., Hamaguchi, Y., … Muramatsu, A. (2012). Unusual and typical features of a novel restorer-of-fertility gene of sugar beet (*Beta vulgaris* L.). *Genetics*, *192*, 1347–1358. https://doi.org/10.1534/genetics.112.145409

Maunder, A., & Pickett, R. (1959). The genetic inheritance of cytoplasmic-genetic male sterility in grain sorghum. *Agronomy Journal*, *51*, 47–49. https://doi.org/10.2134/agronj1959.00021962005100010016x

Michelmore, R. W., Paran, I., & Kesseli, R. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences*, *88*, 9828–9832. https://doi.org/10.1073/pnas.88.21.9828

Murray, M., & Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic acids research*, *8*, 4321–4326. https://doi.org/10.1093/nar/8.19.4321

Murty, U., & Gangadhar, G. (1990). Milo and non-milo sources of cytoplasms in Sorghum bicolor (L.) Moench. III. Genetics of fertility restoration. *Cereal Research Communications*, *18*, 111–116.

vanOoijen, J., & Voorrips, R. (2001). JoinMap Version 3.0: Software for the Calculation of Genetic Linkage Maps. Wageningen: Centre for Plant Breeding and Reproduction Research.

Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., … Poliakov, A. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, *457*, 551–556. https://doi.org/10.1038/nature07723

Pranathi, K., Viraktamath, B. C., Neeraja, C. N., Balachandran, S. M., Hari Prasad, A. S., Koteswara Rao, P., Revathi, P., Senguttuvel, P., Hajira, S. K., Balachiranjeevi, C. H., … Sundaram, R. M. (2016). Development and validation of candidate gene-specific markers for the major fertility restorer genes, Rf4 and Rf3 in rice. *Molecular Breeding*, *36*(10), 145. http://doi.org/10.1007/s11032-016-0566-8

Praveen, M., Suneetha, N., Umakanth, A. V., Patil, J., & Madhusudhana, R. (2015). Inheritance and molecular mapping of Rf 6 locus with pollen fertility restoration ability on A 1 and A 2 cytoplasms in sorghum. *Plant Science*, *238*, 73–80.

Pring, D., Conde, M., & Schertz, K. (1982). Organelle genome diversity in sorghum: Male-sterile cytoplasms. *Crop Science*, *22*, 414–421. https://doi.org/10.2135/cropsci1982.0011183X002200020049x

Reddy, B., Rai, K., Sarma, N., Kumar, I., & Saxena, K. (2003). Cytoplasmic-nuclear male sterility: origin, evaluation, and utilization in hybrid development. In H. K. Jain, & M. C. Kharkwal (Eds.), *Plant breeding: Mendelian to Molecular Approaches*. New Delhi, India: Narosa Publishing House Pvt. Ltd.

Ruiz, O. N., & Daniell, H. (2005). Engineering cytoplasmic male sterility via the chloroplast genome by expression of β-ketothiolase. *Plant Physiology*, *138*, 1232–1246. https://doi.org/10.1104/pp.104.057729

Schertz, K., Sotomayor-Rios, A., & Torres-Cardona, S. (1989). Cytoplasmic-nuclear male sterility–opportunities in breeding and genetics. Proc Grain Sorghum Res and Utility Conf, 175–186, Vol. *16*.

Schloss, S., Mitchell, S., White, G., Kukatla, R., Bowers, J., Paterson, A., & Kresovich, S. (2002). Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. *Theoretical and Applied Genetics*, *105*, 912–920.

Schmitz-Linneweber, C., & Small, I. (2008). Pentatricopeptide repeat proteins: A socket set for organelle gene expression. *Trends in Plant Science*, *13*, 663–670. https://doi.org/10.1016/j.tplants.2008.10.001

Schnable, P. S., & Wise, R. P. (1998). The molecular basis of cytoplasmic male sterility and fertility restoration. *Trends in plant science*, *3*, 175–180. https://doi.org/10.1016/S1360-1385(98)01235-7

Senthil, N., Ramasamy, P., & Fazlullah Khan, A. (1998). Fertility restoration and heterosis involving different cytoplasms in sorghum (*Sorghum bicolor* (L.) Moench) hybrids. *Journal of Genetics and Breeding*, *52*, 339–342.

Sievers, F., & Higgins, D. G. (2014). Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. In: Russell, D. (eds) Multiple Sequence Alignment Methods. Methods in Molecular Biology (Methods and Protocols), *1079*. Humana Press, Totowa, NJ.

Small, I. D., & Peeters, N. (2000). The PPR motif–a TPR-related motif prevalent in plant organellar proteins. *Trends in biochemical sciences*, *25*, 45–47. https://doi.org/10.1016/S0968-0004(99)01520-0

Smith, R. L., & Chowdhury, M. (1989). Mitochondrial DNA polymorphism in male-sterile and fertile cytoplasms of pearl millet. *Crop Science*, *29*, 809–814. https://doi.org/10.2135/cropsci1989.0011183X002900030054x

Springer, N. M., Ying, K., Fu, Y., Ji, T., Yeh, C.-T., Jia, Y., … Rosenbaum, H. (2009). Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. *PLoS genetics*, *5*, e1000734.

Stephens, J., & Holland, R. (1954). Cytoplasmic male-sterility for hybrid sorghum seed production. *Agronomy Journal*, *46*, 20–23. https://doi.org/10.2134/agronj1954.00021962004600010006x

Wang, Y., Xiong, G., Hu, J., Jiang, L., Yu, H., Xu, J., … Xu, J. (2015). Copy number variation at the GL7 locus contributes to grain size diversity in rice. *Nature genetics*, *47*, 944–948. https://doi.org/10.1038/ng.3346

Wang, Z., Zou, Y., Li, X., Zhang, Q., Chen, L., Wu, H., … Luo, D. (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. *The Plant Cell*, *18*, 676–687. https://doi.org/10.1105/tpc.105.038240

Wen, L., Tang, H., Chen, W., Chang, R., Pring, D., Klein, P., … Klein, R. (2002). Development and mapping of AFLP markers linked to the sorghum fertility restorer gene rf4. *Theoretical and Applied Genetics*, *104*, 577–585. https://doi.org/10.1007/s00122-001-0774-3

Xu, X.-B., Liu, Z.-X., Zhang, D.-F., Liu, Y., Song, W.-B., Li, J.-S., & Dai, J.-R. (2009). Isolation and analysis of rice Rf1-orthologus PPR genes co-segregating with Rf3 in maize. *Plant molecular biology reporter*, *27*, 511. https://doi.org/10.1007/s11105-009-0105-4

Yao, F. Y., Xu, C. G., Yu, S. B., Li, J., Gao, Y., Li, X., & 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. https://doi.org/10.1023/A:1003165116059

Yonemaru, J., Ando, T., Mizubayashi, T., Kasuga, S., Matsumoto, T., & Yano, M. (2009). Development of genome-wide simple sequence repeat markers using whole-genome shotgun sequences of sorghum (*Sorghum bicolor* (L.) moench). *DNA Research*, *16*, 187–193. https://doi.org/10.1093/dnares/dsp005

Yue, B., Vick, B., Cai, X., & Hu, J. (2010). Genetic mapping for the Rf1 (fertility restoration) gene in sunflower (*Helianthus annuus* L.) by SSR and TRAP markers. *Plant Breeding*, *129*, 24–28. https://doi.org/10.1111/j.1439-0523.2009.01661.x

Zhang, Q., Shen, B., Dai, X., Mei, M., Maroof, M. S., & Li, Z. (1994). Using bulked extremes and recessive class to map genes for photoperiod-sensitive genic male sterility in rice. *Proceedings of the National Academy of Sciences USA*, *91*, 8675–8679. https://doi.org/10.1073/pnas.91.18.8675

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Madugula P, Uttam AG, Tonapi VA, Ragimasalawada M. Fine mapping of *Rf2*, a major locus controlling pollen fertility restoration in sorghum A_1 cytoplasm, encodes a PPR gene and its validation through expression analysis. *Plant Breed*. 2018;137:148–161. https://doi.org/10.1111/pbr.12569