Research Paper

Fine mapping of Rf5 region for a sorghum fertility restorer gene and microsynteny analysis across grass species

Atsushi Kiyosawa¹, Jun-ichi Yonemaru*², Hiroshi Mizuno², Hiroyuki Kanamori², Jianzhong Wu², Hiroyuki Kawahigashi³ and Kazumi Goto³

¹ Nagano Animal Industry Experiment Station, 10931-1 Kataoka, Shiojiri, Nagano 399-0711, Japan
² Institute of Crop Science, National Agriculture and Food Research Organization, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan
³ Nagano Agricultural Development Public Corporation, Kami-Ina Branch, 3497 Arai, Ina, Nagano 396-8666, Japan

Introduction

Cytoplasmic male sterility (CMS) is widely used to control pollination in the production of commercial F₁ hybrid seed in sorghum. So far, 6 major fertility restorer genes, Rf1 to Rf6, have been reported in sorghum. Here, we fine-mapped the Rf5 locus on sorghum chromosome 5 using descendant populations of a ‘Nakei MS-3A’ × ‘JN43’ cross. The Rf5 locus was narrowed to a 140-kb region in BTx623 genome (161-kb in JN43) with 16 predicted genes, including 6 homologous to the rice fertility restorer Rf1 (PPR.1 to PPR.6). These 6 homologs have tandem pentatricopeptide repeat (PPR) motifs. Many Rf genes encode PPR proteins, which bind RNA transcripts and modulate gene expression at the RNA level. No PPR genes were detected at the Rf5 locus on the corresponding homologous chromosome of rice, foxtail millet, or maize, so this gene cluster may have originated by chromosome translocation and duplication after the divergence of sorghum from these species. Comparison of the sequences of these genes between fertile and CMS lines identified PPR.4 as the most plausible candidate gene for Rf5.

Key Words: sorghum, Rf (fertility restorer) genes, Rf5, fine mapping, microsynteny.
showed that acid insertion (R26) in the protein of sterile plants. A series of candidates of the causal genes have been reported. Recently, we reported the mode of inheritance involved in fertility restoration in seven 2005). PPR13 contains 14 PPRs as well as a C-terminal E motif (Lurin 1960s. CMS lines introduced from the USA were used as breeding for forage use in Japan, but no studies of their restorer sequences indicated a frame-shift mutation with a stop codon (TGA) at 286 bp caused by an 11-bp insertion in the mitochondrial atp9 gene (Jordan 2002). Marker genotyping were the same as in our previous study (Kiyosawa 2020). We used 410 F4 plants and 188 F5 plants for genetic mapping. We used SSR markers (Yonemaru et al. 2009) and indel markers for fine mapping of the Rf5 region (Supplemental Table 1). Genomic DNA extraction, PCR amplification, and marker genotyping were the same as in our previous study (Kiyosawa et al. 2020, Takai et al. 2012, Yonemaru et al. 2015).

BAC sequence analysis of Rf5 region

Bacterial artificial chromosome (BAC) libraries were constructed from young leaves of the CMS maintenance line ‘Nakei MS-3B’ (39 267 clones with an average insert size of 134 kb) and the restorer line ‘JN43’ (30 811 clones, 125 kb). The libraries were prepared by conventional methods, comprising a partial DNA digest with Hind III, size fractionation of high-molecular-weight DNA by pulsed-field gel electrophoresis (CHEF, Bio-Rad Laboratories, Hercules, CA, USA), vector ligation (pIndigo BAC-5, Epicentre Biotechnologies, Madison, WI, USA), and transformation into E. coli strain DH10B. Positive BAC clones covering the Rf5 region were screened from each library using tightly linked DNA markers through PCR amplification, and the identified BACs were shotgun-sequenced (Sasaki et al. 2002, Wu et al. 2003) to provide approximately tenfold sequence coverage.

PCR analysis using PCR markers MS3B Chr05_2420485 and JN43 Chr05_2589446 (Supplemental Table 1) identified three BAC clones containing inserts from the Rf5 region: NaMSB-0088G09 from ‘Nakei MS-3B’ and JN43-0015A16 and JN43-0018F16 from ‘JN43’. BAC sequences were deposited in the DDBJ (Acc. Nos. LC494266 and LC494267).

The reference genome sequence used was Sbicolor_
Fine mapping of Rf5 region in sorghum

v3.0.1_454, derived from ‘BTx623’ (McCormick et al. 2018). The structures of the PPR genes were manually corrected in consideration of frameshifts and splicing junctions. As a result, the structures of six PPR genes of ‘BTx623’ used here (PPR.1_B, PPR.2_B, PPR.3_B, PPR.4_B, PPR.5_B, and PPR.6_B) differed from the structures annotated in Sbicolor v3.1.1_454. The positions and both sets of annotations of the six PPR genes in the Rf5 region are shown in Supplemental Tables 2 and 3. The mitochondrial genome sequence was derived from ‘BTx623’ (Acc. No. DQ 984518).

Next-generation DNA sequencing of restorer and CMS lines

To elucidate the relationship between Rf5 function and sequences, we re-sequenced five restorer lines—‘JN43’, ‘JN290’, ‘SDS7444’, ‘Chohin237Daikoukaku’, and ‘JN503’—and four CMS lines—‘AMP-21’, ‘Nakei MS-3A’, ‘(954149)A’, and ‘MS175 (932233)A’—used in the five parental combinations to detect a QTL in the Rf5 region in the F2 populations (Kiyosawa et al. 2020) by short-read DNB-Seq technology. These data were deposited in the DDBJ Sequence Read Archive under accession number DRA012197. The informatics for the analysis of re-sequencing data was the same as reported previously (Kiyosawa et al. 2020) except for reference genome sequence. To obtain the correct alignments in the region rich in structural variation, we merged two sequence datasets aligned by using two sets of reference sequences, Sbicolor_v3.0.1_454 and Sbicolor_v3.0.1_454, with the Rf5 region replaced by that of ‘JN43’.

Microsynteny analysis of Rf5 locus

Comparing the Rf5 locus of sorghum Chr. 5 (Sbiclor Chr. 05) with homologous chromosomes of rice (Osativa Chr. 11), foxtail millet (Sitalica Scaffold 8), and maize (Zmays Chr. 4), we found no corresponding PPR genes in any of the corresponding chromosomes (Fig. 2). In foxtail millet, homologous PPR genes were located in scaffold 5 (data not shown), not in the corresponding chromosome, scaffold 8. Also, in rice, homologous PPR gene was Rf4 (Os10g0495200, Kazama and Toriyama 2014) in Chr. 10, not in the corresponding chromosome, Chr. 11.

Genome sequence analysis

Since the sorghum genome database was constructed from ‘BTx623’, which lacks Rf5 restoration ability, we isolated and sequenced BAC clones JN43-0015A16 and JN43-0018F16 from a BAC library of the restorer male parent line ‘JN43’ using markers MS3B_Chr05_2420485 (2.425 Mb) and JN43 Chr05_2589446 (2.574 Mb), and BAC clone MS3B-0088G09 from the maintainer line ‘Nakei MS-3B’. Fine mapping defined the length of the Rf5 locus region as 140 kb in ‘BTx623’, 161 kb in ‘JN43’, and 126 kb in ‘Nakei MS-3B’ (Fig. 1).

In each Rf5 region, there were 6 PPR genes predicted in ‘BTx623’ and ‘JN43’ and 5 in ‘Nakei MS-3B’, which had a chimeric PPR gene (PPR.3+4) caused by the fusion of the adjacent genes PPR.3 and PPR.4 (Fig. 1). With this exception, a dot plot with genomic sequences (Supplemental Fig. 2) supported no structural indels, translocations, or inversions involving the PPR genes; that is, the PPR genes were conserved among the 3 cultivars.

PPR.4 of ‘JN43’ is a candidate allele of functional Rf5

The PPR genes of these 3 cultivars had sequence diversity due to indels, fusions, frame shifts, and gain of a stop codon, and thus their protein lengths varied from 32 to 803 aa (Fig. 3). To distinguish each allele, we appended each

Results

Mapping of the Rf5 gene

To narrow down the Rf5 region, we examined 410 F4 plants. The Rf5 locus was mapped to the region between markers InDel_Rf5-5 (2.284 Mb) and InDel_Rf5-17 (2.737 Mb) on Chr. 5 (Fig. 1). Further mapping was conducted by screening 188 F4 plants (F3-78-56-456), which allowed us to localize Rf5 to a 140-kb region between markers Rf5_US2 (2.434 Mb) and Rf5_LS11 (2.574 Mb) (Fig. 1). A search of the sorghum genome sequence annotation database v. 3.1 (https://genome.jgi.doe.gov/portal/) revealed 16 predicted genes in this region (Supplemental Table 2), including 6 tandemly repeated PPR genes (PPR.1–PPR.6; Fig. 1, Supplemental Table 3), one or more of which were considered the Rf5 gene (previously described as 1 of 7 PPR genes in the Rf5 region, Jordan et al. 2011).

Phylogenetic analysis

Alignment and phylogenetic reconstructions were performed using the ‘build’ function of ETE3 v. 3.1.1 software (Huerta-Cepas et al. 2016) as implemented on the GenomeNet website (https://www.genome.jp/tools-bin/etc/). The maximum likelihood tree was inferred in PhyML v. 20160115 software (parameters --piv n e -f m --nclasse 4 -o tlr --alpha e --bootstrap 100; Guindon et al. 2010) and arranged in MEGA7 software (Kumar et al. 2016). The amino acid sequences and alignment of PPR proteins are summarized in Supplemental Fig. 1.
Fig. 1. Predicted PPR genes at \( Rf5 \) locus. The \( Rf5 \) locus was mapped to a ~453-kb region on Chr. 5 between markers InDel_Rf5-5 (2.284 Mb) and InDel_Rf5-17 (2.737 Mb) in the \( F_4 \) population (\( N = 410 \)). It was fine-mapped to a ~140-kb region between markers Rf5_US2 (2.434 Mb) and Rf5_LS11 (2.574) in the \( F_6 \) population (\( N = 188 \)). In the target region, 16 genes (black and white boxes) were predicted from the sorghum BTx623 genome (Sbicolor v3.0.1_454), including 6 pentatricopeptide repeat genes in tandem (black boxes). Grey boxes indicate the position and transcription direction of predicted PPR genes at the \( Rf5 \) locus in the total genomic region obtained by the sequencing of BACs of ‘Nakei MS-3B’ and ‘JN43’ and the corresponding region of Chr. 5 of the reference ‘BTx623’ (Sbicolor v3.0.1_454).

Fig. 2. Microsynteny analysis of \( Rf5 \) region. Sorghum Chr. 5 (Sbicolor Chr. 5) has synteny with rice (Osativa) Chr. 11, foxtail millet (Sitalica) scaffold 8, and maize (Zmays) Chr. 4. Six predicted PPR genes (arrows) exist in sorghum but not in the other species.
gene name with the initial letter of the cultivar name (J = ‘JN43’; B = ‘BTx623’; N = ‘Nakei MS-3B’); for example, PPR.1_B. To identify PPR genes encoding a functional Rf5 protein, we compared amino acid sequences encoded by 6 tandemly repeated PPR genes (PPR.1 to PPR.6) between ‘JN43’ (fertility restorer) and ‘BTx623’ (non-restorer). The sequences of PPR.4_J (fertility restorer), PPR.4_B (non-restorer), and PPR.3+4_N (non-restorer) differed from each other (Table 1). PPR.4_B had 11 aa substitutions compared with PPR.4_J in the PPR motif, among which the T622N substitution may influence the affinity to target RNA (Fig. 4a). There were 127 aa differences between PPR.4_J and PPR.3+4_N (Supplemental Fig. 3); the differences between PPR.4_B and PPR.3+4_N seem to involve a loss of function.

PPR.1 had an I766M substitution between ‘JN43’ and ‘BTx623’ (Supplemental Fig. 1). As the substitution lies in the C-terminal region, which lies outside of the conserved PPR motif, we thought that it does not affect RNA recognition. PPR.3_J encoded 300 aa with only 5 PPR motifs, which suggests that it is not functional as an Rf. Functionally characterized PPR-protein genes for Rf usually contain 11–18 PPR motifs which recognize a specific RNA sequence (Dahan and Mireau 2013). The amino acid sequences of PPR.2, PPR.5, and PPR.6 were identical between ‘JN43’ and ‘BTx623’ (Table 1). These results suggest that PPR.1, PPR.2, PPR.3, PPR.5, and PPR.6 have equivalent functions in ‘JN43’ (restorer) and ‘BTx623’ (non-restorer), and thus are not candidates for the functional Rf5 gene. Thus, PPR.4 is the most plausible candidate for Rf5.

In ‘Nakei MS-3B’ (non-restorer), all PPR genes were drastically changed compared with those in ‘JN43’. PPR.3+4_N had only 84% (676/803) sequence identity with PPR.4_J (Supplemental Fig. 3). The other genes also had lower sequence identity between ‘JN43’ and ‘Nakei MS-3B’ than between ‘JN43’ and ‘BTx623’ (Table 1).

### Discussion

Our previous study showed that Rf5 is the main restorer gene in Japanese CMS lines (Kiyosawa et al. 2020). By detailed mapping of the Rf5 locus, we delimited it to a 140-kb region of chromosome 5 in BTx623 genome (161-kb in JN43), where six PPR genes were predicted. Since the Rf5 locus restores the fertility of A1 cytoplasm lines as well as do the Rf1 and Rf2 loci, the Rf5 gene may have high homology with Rf1 and/or Rf2.

Sorghum branched from other Poaceae about 50 million years ago (Gaut 2002). Within the Panicoideae, sorghum branched from millet about 28 million years ago and from maize about 9 million years ago. Microsynteny analysis of the Rf5 locus detected no PPR genes in rice, foxtail millet, or maize on the corresponding chromosome. These results suggest that the Rf5 region arose more recently after the branching event from maize, and the PPR genes were translocated from another chromosome and duplicated on Chr. 5.

Nine restorers-of-fertility–like (RFL) PPR genes were reported as a cluster together with other tight clusters containing RFLs on Chr. 5 in sorghum (Sykes et al. 2017). This region had 50% (9/18 genes) of the identified RFLs within 554 kb (Sykes et al. 2017), which includes the Rf5 region. This region can be considered a hotspot for recombination at higher rates than expected (Sykes et al. 2017).

### Table 1. Amino acid identity of PPR proteins among ‘BTx623’, ‘Nakei MS-3B’, and ‘JN43’

|       | ‘BTx623’ vs ‘JN43’ | ‘Nakei MS-3B’ vs ‘JN43’ | ‘BTx623’ vs ‘Nakei MS-3B’ |
|-------|-------------------|-------------------------|--------------------------|
| PPR.1 | 99.9% (798/799)    | 3.4% (27/799)           | 3.4% (27/799)            |
| PPR.2 | 100% (803/803)     | 95.6% (768/803)         | 95.6% (768/803)          |
| PPR.3 | 99.7% (299/300)    | --                      | --                       |
| PPR.4 | 98.6% (792/803)    | --                      | --                       |
| PPR.5 | 100% (364/364)     | 63.7% (232/364)         | 63.7% (232/364)          |
| PPR.6 | 100% (757/757)     | 99.1% (750/757)         | 99.1% (750/757)          |

* The ‘Nakei MS-3B’ sequence has a stop codon in the middle, so the sequence is truncated.
which led to expansion of the family through a probable ‘birth-and-death’ process involving diversifying selection (Fujii et al. 2011, Geddy and Brown 2007). RFL genes are highly diverse among species and even among strains of the same species, showing strong signals of diversifying selection (Fujii et al. 2011, Geddy and Brown 2007). In our study, ‘Nakei MS-3B’ had very different PPR genes in the Rf5 region from those in the other two cultivars. The coevolution of fertility-restoring PPR genes with CMS-inducing mitochondrial genes has been described as an arms race between the mitochondrial and nuclear genomes (Touzet and Budar 2004). This is similar to the coevolution of plant genes for leucine-rich-repeat–resistance proteins and rapidly evolving pathogen effectors in plant–pathogen interactions (Dahan and Mireau 2013). The difference between ‘Nakei MS-3B’ and the other two cultivars shows the rapid evolution of PPR genes in this region.

**Diversity of PPR motif of sorghum Rf5 proteins**

There were 19 PPR motifs predicted in PPR.4_J (Fig. 4a). The motifs consist of a repetitive sequence of 35 amino acids (36 in PPR2 and 38 in PPR3). In the PPR, amino acids Y3, L10, C11, G14, F23, M26, G30, and P33 are highly conserved (Fig. 4b). These features are consistent with the characteristics of the P-type PPR motifs (Yagi et al. 2013). Study of recognition between specific amino acids within the PPRs and target RNA sequences identified the amino acids at position 4 and 34 of each PPR motif as important for the recognition of transcripts, and the amino acid at position 1 restricts the accuracy of the interactions between nucleotides and the protein (Barkan et al. 2012, Takenaka et al. 2013, Yagi et al. 2013). According to the recognition codes of the P-type PPR motifs (Yagi et al. 2013), the RNA sequence recognized by the 19 PPR motifs of PPR.4_J was predicted to be 5ʹ- AUCGACAAUGAUUYUCANY-3ʹ.

The sequence of the 17th PPR motif of PPR.4 differed between ‘JN43’ (restorer) and the others (‘BTx623’ and ‘Nakei MS-3B’). In PPR.4_B, the amino acid at position 4 in the 17th PPR motif was changed from T to N, and the conserved 33rd P was changed to Q. In PPR.3+4_N also, the amino acid at position 4 changed from T to N (Supplemental Figs. 1, 3). Thus, both PPR.4_B and PPR.3+4_N had altered PPR motifs, and the nucleotide recognized by these motifs may be important. Two amino acid substitutions in the 17th PPR motif (T to N in the 4th position and
P to Q in the 33rd position) perfectly coincide with those in other fertile and CMS lines (Table 2, Supplemental Fig. 4). These data confirm PPR4 as the most plausible candidate gene for RF5. The change of these PPR motifs may cause the target RNA not to be recognized and result in loss of the restoration of fertility.

In ‘JN43’ with a functional RF5 allele, PPR4 and PPR2 are paired genes which seem to have been generated by segmental duplication (Fig. 3, Supplemental Fig. 5). PPR2 also has 19 PPR motifs and seems intact and functional. However, 30 aa substitutions have already occurred between them (Supplemental Fig. 5). The substitutions N128D, N552D, G622D, D692N (Supplemental Fig. 5) are involved in RNA recognition and suggest these differences in RNA sequence recognition by each PPR protein. Other PPR genes in the RF5 region (PPR1, PPR3, PPR5, and PPR6) have deletions or point mutations making them non-functional (Fig. 3). These results show the rapid evolution of PPR genes in sorghum.

**What is the mitochondrial gene responsible for CMS?**

To find genes responsible for CMS, researchers have compared the whole mitochondrial genome between CMS lines and maintainer lines in maize and wheat. Differences of many indels and chimerism of ORFs have been found (Allen et al. 2007, Liu et al. 2011). However, the causal genes of CMS for RF5 in sorghum have not been identified yet. It is important to clarify the sequences of corresponding mitochondrial genes derived from the CMS line that can be rescued by RF5.

Three independent sequences similar to the RNA sequence recognized by the RF5 protein were found in the mitochondrial genome sequence of ‘BTx623’ (468,628 bp). Two were found in the non-coding region but the other one matched part of the coding region of rps2b, which encodes mitochondrial ribosomal protein 2B. Part of the rps2b RNA sequence (5’-CAAUGAUUCUAAT-3’) partially matched 5’-AUCGACAAUGAUUYUCANY-3’, which the PPR4_J protein recognizes. We consider the mitochondrial rps2b gene as a candidate gene for CMS. The RF5 protein (PPR4_J) may bind to rps2b mRNA, stabilizing it. Interestingly, a maize PPR protein, EMP4 (empty pericarp 4), is necessary to regulate the correct expression of mitochondrial rps2b for seed development and plant growth (Gutierrez-Marcos et al. 2007). This is an example in which the stabilization of the mRNA of a mitochondrial gene by a PPR protein is required for plant development. Since the mitochondrial genome sequence from sorghum CMS lines has not been obtained yet, we used the whole sorghum mitochondrial sequence from ‘BTx623’. ‘ATx623’ (A line corresponding to ‘BTx623’) may have an indel or chimerization of one or more mitochondrial genes. Detailed comparison of the mitochondrial sequence of an A1 cytoplasmic male sterile line, ‘ATx623’ and its maintainer line ‘BTx623’ may provide information on candidate genes for CMS. Interestingly, RF5 restores not only A1 cytoplasmic sterility but also A2 cytoplasmic sterility, and finding mitochondria sequences that interact with the RF5 protein (PPR4_J) and determining their relationship will provide important information for F1 breeding.

### Author Contribution Statement

AK and JY designed the experiments. AK, JY, HKN, JW, HKW, and KG carried out the experiments. AK, JY, HKW, and HM analyzed the data. JY, HKW, and HM wrote the paper. All authors reviewed and approved the final manuscript.

### Acknowledgments

We are indebted to sorghum breeders for the experimental lines and the staff who performed the field work at the Nagano Animal Industry Experiment Station. We gratefully acknowledge Yukari Shimazu and Emi Abe for genotyping mapping populations. This work was supported by grants...
from the Ministry of Agriculture, Forestry and Fisheries of Japan (Rural Biomass Research Project, BCD-1223; Genomics for Agricultural Innovation, SOR-0005).

Literature Cited

Allen, J.O., C.M. Fauron, P. Minx, L. Roark, S. Oddiraju, G.N. Lin, L. Meyer, H. Sun, K. Kim, C. Wang et al. (2007) Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. Genetics 177: 1173–1192.

Barkan, A. and I. Small (2014) Pentatricopeptide repeat proteins in Gutierrez-Marcos, J.F., M. Dal Pra, A. Giulini, L.M. Costa, G. Fujii, S., C.S. Bond and I.D. Small (2011) Selection patterns on Gully, B.S., N. Cowieson, W.A. Stanley, K. Shearston, I.D. Small, A. Bennetzen, J.L., J. Schmutz, H. Wang, R. Percifield, J. Hawkins, A.C. Ban, T., J. Ke, R. Chen, X. Gu, M.H. Tan, X.E. Zhou, Y. Kang, K. Melcher, J.K. Zhu and H.E. Xu (2013) Structure of a PLS-class pentatricopeptide repeat protein provides insights into mechanism of RNA recognition. J Biol Chem 288: 31540–31548.

Barkan, A., M. Rojas, S. Fujii, A. Yap, Y.S. Chong, C.S. Bond and I. Small (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. PLoS Genet 8: e1002910.

Barkan, A. and I. Small (2014) Pentatricopeptide repeat proteins in plants. Annu Rev Plant Biol 65: 415–442.

Bennetzen, J.L., J. Schmutz, H. Wang, R. Percifield, J. Hawkins, A.C. Pontaroli, M. Estep, L. Feng, J.N. Vaughn, J. Grimmwood et al. (2012) Reference genome sequence of the model plant Setaria. Nat Biotechnol 30: 555–561.

Dahan, J. and H. Mireau (2013) The Rf and Rf-like PPR in higher plants, a fast-evolving subclass of PPR genes. RNA Biol 10: 1469–1476.

Fujii, S., C.S. Bond and I.D. Small (2011) Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. Proc Natl Acad Sci USA 108: 1723–1728.

Gaut, B.S. (2002) Evolutionary dynamics of grass genomes. New Phytol 154: 15–28.

Geddy, R. and G.G. Brown (2007) Genes encoding pentatricopeptide repeat (PPR) proteins are not conserved in location in plant genomes and may be subject to diversifying selection. BMC Genomics 8: 130.

Guindon, S., J.F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk and O. Gascuel (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59: 307–321.

Gully, B.S., N. Cowieson, W.A. Stanley, K. Shearston, I.D. Small, A. Barkan and C.S. Bond (2015) The solution structure of the pentatricopeptide repeat protein PPR10 upon binding atpRNA. Nucleic Acids Res 43: 1918–1926.

Gutierrez-Marcos, J.F., M. Dal Pra, A. Giulini, L.M. Costa, G. Gavazzi, S. Cordelier, O. Sellam, C. Tatout, W. Paul, P. Perez et al. (2007) empty pericap4 encodes a mitochondrion-targeted pentatricopeptide repeat protein necessary for seed development and plant growth in maize. Plant Cell 19: 196–210.

Hanson, M.R. and S. Bentolila (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell 16: S154–S169.

Huerta-Cepas, J., F. Serra and P. Bork (2016) ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. Mol Biol Evol 33: 1635–1638.

Jordan, D.R., E.S. Mace, R.G. Henzell, P.E. Klein and R.R. Klein (2010) Molecular mapping and candidate gene identification of the Rf2 gene for pollen fertility restoration in sorghum [Sorghum bicolor (L.) Moench]. Theor Appl Genet 120: 1279–1287.

Jordan, D.R., R.R. Klein, K.G. Sakrewski, R.G. Henzell, P.E. Klein and E.S. Mace (2011) Mapping and characterization of Rf2, a new gene conditioning pollen fertility restoration in A1 and A2 cytoplasm in sorghum (Sorghum bicolor (L.) Moench). Theor Appl Genet 123: 1–14.

Kante, M., H.F.W. Rattunde, B. Nebie, E. Weltzien, B.I.G. Haussmann and W.L. Leiser (2018) QTL mapping and validation of fertility restoration in West African sorghum A1 cytoplasm and identification of a potential causative mutation for Rf2. Theor Appl Genet 131: 2397–2412.

Kazama, T. and K. Toriyama (2014) A fertility restorer gene, Rf4, widely used for hybrid rice breeding encodes a pentatricopeptide repeat protein. Rice (N Y) 7: 28.

Kiyosawa, A., J.-I. Yonemaru, H. Kawahigashi and K. Goto (2020) Analysis of quantitative trait loci for fertility restoration in seven F2 populations derived from sorghum F1 hybrids bred in Japan. Breed Sci 70: 379–386.

Klein, R.R., P.E. Klein, J.E. Mullet, P. Minx, W.L. Rooney and K.F. Schertz (2005) Fertility restorer locus Rf1 of sorghum (Sorghum bicolor (L.) L) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. Theor Appl Genet 111: 994–1012.

Kuhlman, L.C., D.R. Pring, W.L. Rooney and H.V. Tang (2006) Alloplasmic frequency at the Rf3 and Rf4 loci and the genetics of A3 cytoplasmic fertility restoration in converted sorghum lines. Crop Sci 46: 1576–1580.

Kumar, S., G. Stecher and K. Tamura (2016) MEGA7: Molecular Evolutionary Genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33: 1870–1874.

Liu, H., P. Cui, K. Zhan, Q. Lin, G. Zhuo, X. Guo, F. Ding, W. Yang, D. Liu, S. Hu et al. (2011) Comparative analysis of mitochondrial genomes between a wheat K-type cytoplasmic male sterility (CMS) line and its maintainer line. BMC Genomics 12: 163.

Lurin, C., C. Andres, S. Aubourg, M. Bellaoui, F. Bitton, C. Bruyere, M. Caboche, C. Debast, J. Gualberto, B. Hoffmann et al. (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16: 2089–2103.

Manna, S. (2015) An overview of pentatricopeptide repeat proteins and their applications. Biochimie 113: 93–99.

McCormick, R.F., S.K. Truong, A. Sreedasyam, J. Jenkins, S. Shu, D. Sims, M. Kennedy, M. Amirebrahimi, B.D. Weers, B. McKinley et al. (2018) The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. Plant J 93: 338–354.

Ouyang, S., W. Zhu, J. Hamilton, H. Lin, M. Campbell, K. Childs, F. Thibaud-Nissen, R.L. Malek, Y. Lee, L. Zheng et al. (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic Acids Res 35: D883–D887.

Praveen, M., G. Anurag Uttam, N. Suneetha, Av. Umakanth, J.V. Patil and R. Madhusudhana (2015) Inheritance and molecular mapping of Rf6 locus with pollen fertility restoration ability on A1 and A2 cytoplasm in sorghum. Plant Sci 238: 73–80.

Praveen, M., G. Anurag Uttam, V.A. Tonapi, M. Ragimasalawada and R. Snowdon (2018) Fine mapping of A2 fertility restorer locus with pollen fertility restoration ability on A1 and A2 cytoplasm in sorghum. Breed Sci 137: 148–161.

Pring, D., H. Tang, W. Howad and F. Kempken (1999) A unique two-gene gametophytic male sterility system in sorghum involving a possible role of RNA editing in fertility restoration. J Hered 90:
386–393.
Sasaki, T., T. Matsumoto, K. Yamamoto, K. Sakata, T. Baba, Y. Katayose, J. Wu, Y. Niimura, Z. Cheng, Y. Nagamura et al. (2002) The genome sequence and structure of rice chromosome 1. Nature 420: 312–316.
Schmitz-Linneweber, C. and I. Small (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci 13: 663–670.
Schnable, P. and R. Wise (1998) The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci 3: 175–180.
Schnable, P.S., D. Ware, R.S. Fulton, J.C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T.A. Graves et al. (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326: 1112–1115.
Soderlund, C., W. Nelson, A. Shoemaker and A. Paterson (2006) SyMAP: A system for discovering and viewing syntenic regions of FPC maps. Genome Res 16: 1159–1168.
Soderlund, C., M. Bomhoff and W.M. Nelson (2011) SyMAP v3.4: a turnkey synteny system with application to plant genomes. Nucleic Acids Res 39: e68–e68.
Stephens, J.C. and R. Holland (1954) Cytoplasmic male sterility for hybrid sorghum seed production. Agron J 46: 20–23.
Sykes, T., S. Yates, I. Nagy, T. Asp, I. Small and B. Studer (2017) In silico identification of candidate genes for fertility restoration in cytoplasmic male sterile perennial ryegrass (Lolium perenne L.). Genome Biol Evol 9: 351–362.
Takai, T., J. Yonemaru, H. Kaidai and S. Kasuga (2012) Quantitative trait locus analysis for days-to-heading and morphological traits in an RIL population derived from an extremely late flowering F1 hybrid of sorghum. Euphytica 187: 411–420.
Takenaka, M., A. Zehrmann, A. Brennicke and K. Graichen (2013) Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. PLoS One 8: e65343.
Tang, H.V., R. Chang and D.R. Pring (1998) Cosegregation of single genes associated with fertility restoration and transcript processing of sorghum mitochondrial orf107 and orf209. Genetics 150: 383–391.
Tang, H.V. and D.R. Pring (2003) Conversion of fertility restoration of the sorghum IS1112C (A3) male-sterile cytoplasm from two genes to one gene. Joint contribution of the USDA-ARS and the University of Florida Agric. Exp. Stn. published as journal series No. R-09082. Crop Sci 43: 1747–1753.
Touzet, P. and F. Budar (2004) Unveiling the molecular arms race between two conflicting genomes in cytoplasmic male sterility? Trends Plant Sci 9: 568–570.
Wu, J., H. Mizuno, M. Hayashi-Tsugane, Y. Ito, Y. Chiden, M. Fujisawa, S. Katagiri, S. Saji, S. Yoshiki, W. Karasawa et al. (2003) Physical maps and recombination frequency of six rice chromosomes. Plant J 36: 720–730.
Yagi, Y., S. Hayashi, K. Kobayashi, T. Hirayama and T. Nakamura (2013) Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. PLoS One 8: e57286.
Yonemaru, J., T. Ando, T. Mizubayashi, S. Kasuga, T. Matsumoto and M. Yano (2009) Development of genome-wide simple sequence repeat markers using whole-genome shotgun sequences of sorghum (Sorghum bicolor (L.) Moench). DNA Res 16: 187–193.
Yonemaru, J., S.H. Choi, H. Sakai, T. Ando, A. Shomura, M. Yano, J. Wu and S. Fukuoka (2015) Genome-wide indel markers shared by diverse Asian rice cultivars compared to Japanese rice cultivar ‘Koshihikari’. Breed Sci 65: 249–256.