Theor Appl Genet (2011) 123:383–396
DOI 10.1007/s00122-011-1591-y

ORIGINAL PAPER

Mapping and characterization of Rf_5: a new gene conditioning pollen fertility restoration in A_1 and A_2 cytoplasm in sorghum (*Sorghum bicolor* (L.) Moench)

D. R. Jordan · R. R. Klein · K. G. Sakrewski · R. G. Henzell · P. E. Klein · E. S. Mace

Received: 1 September 2010 / Accepted: 31 March 2011 / Published online: 13 April 2011

Abstract With an aim to further characterize the cytoplasmic male sterility–fertility restoration system in sorghum, a major fertility restoration gene was mapped along with a second locus capable of partial restoration of pollen fertility. The major fertility restoration gene, Rf_5, was located on sorghum chromosome SBI-05, and was capable of restoring pollen fertility in both A_1 and A_2 male sterile cytoplasms. Depending on the restorer parent, mapping populations exhibited fertility restoration phenotypes that ranged from nearly bimodal distribution due to the action of Rf_5, to a more normalized distribution reflecting the action of Rf_5 and additional modifier/partial restoration genes. A second fertility restoration locus capable of partially restoring pollen fertility in A_1 cytoplasm was localized to chromosome SBI-04. Unlike Rf_5, this modifier/partial restorer gene acting alone resulted in less than 10% seed set in both A_1 and A_2 cytoplasms, and modified the extent of restoration conditioned by the major restorer Rf_5 in A_1 cytoplasm. In examining the genomic regions spanning the Rf_5 locus, a cluster of pentatricopeptide gene family members with high homology to rice Rf_1 and sorghum Rf_2 were identified as potential candidates encoding Rf_5.

Communicated by H. H. Geiger.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1591-y) contains supplementary material, which is available to authorized users.

D. R. Jordan (✉) · K. G. Sakrewski · R. G. Henzell · E. S. Mace
Agri-Science Queensland, Hermitage Research Station, Yangan Rd, Warwick, QLD 4370, Australia
e-mail: David.R.Jordan@deedi.qld.gov.au

E. S. Mace
e-mail: emma.mace@dpi.qld.gov.au

D. R. Jordan
Queensland Alliance for Agriculture and Food Innovation, Yangan Rd, Warwick, QLD 4370, Australia

R. R. Klein
USDA-ARS, Southern Plains Agricultural Research Center, College Station, TX 77845, USA

P. E. Klein
Department of Horticulture and Institute for Plant Genomics and Biotechnology, Texas A&M University, College Station, TX 77843, USA

Introduction

Cytoplasmic–nuclear male sterility in sorghum was first described by Stephens and Holland (1954) who observed that the interaction of "milo" or A_1 cytoplasm and genes of "kafir" origin produced plants that were male sterile and normal female fertile. This discovery, coupled with the development of male parental lines (restorer or R lines) that carry dominant genes that restore male fertility in hybrid cultivars, permitted the cost-effective exploitation of heterosis via F_1 hybrids in sorghum. Despite the identification of other CMS systems in sorghum (Schertz 1983), the A_1 CMS system is used almost exclusively by commercial hybrid seed production worldwide. Classical genetics studies indicate that fertility restoration in A_1 cytoplasm is complex and controlled by at least two major genes with two or more modifier genes (Maunder and Pickett 1959; Miller and Pickett 1964). These restoration genes and their modifiers are under strong differential selection in hybrid breeding programs with selection for complete restoration of fertility in CMS hybrids by male parents and complete sterility in male sterile cytoplasm in female parents.

One practical implication of this complexity is reluctance on the part of breeders to make crosses between restorers and non-restorers when developing new parental lines (Jordan et al. 2010).This reluctance has resulted in a narrow genetic base among female parental lines.

To date two fertility restoration genes for A_1 cytoplasm have been mapped in sorghum and the causal genes identified as members of the pentatricopeptide repeat (PPR) gene family (Klein et al. 2005; Jordan et al. 2010). PPR proteins are characterised by tandem repeats of a degenerate 35 amino acid motif. This gene family has been predicted to target chloroplasts and mitochondria (Lurin et al. 2004), and PPR proteins are thought to be involved in RNA processing (Schmitz-Linneweber and Small 2008). Several PPR proteins have been shown to act as fertility restorer genes in male sterility systems in a range of crop species (Schmitz-Linneweber and Small 2008).

The A_2 cytoplasm is the only acceptable alternative to the A_1 cytoplasm for commercial hybrid seed production in sorghum (Moran and Rooney 2003) but little is known about the genetic architecture of fertility restoration in A_2 cytoplasm. An investigation of the restoration capacity in A_1 and A_2 cytoplasm of a diverse range of sorghum lines from Sudan indicated that the relationship between fertility restoration in the two cytoplasms was complex (Dahlberg and Madera-Torres 1997). In general, lines with the capacity to restore A_2 cytoplasm were typically also restorers in A_1 cytoplasm whereas lines that restored A_1 cytoplasm were often not restorers in A_2 cytoplasm. These results are in accordance with observations in the Department of Employment, Economic Development and Innovation (DEEDI) sorghum breeding program that crosses between many A_1 restorers and the line QL12 which restores both A_1 and A_2 result in progeny that segregate for the restoration of A_1 and A_2 cytoplasms (Jordan DR, unpublished data).

The cytoplasmic genetic uniformity of commercial sorghum could leave the industry vulnerable if the A_1 cytoplasm was associated with susceptibility to a disease. Such a situation occurred in 1969 and 1970 when the US maize crop was struck by an epidemic of southern corn leaf blight (*Bipolaris maydis*) that caused severe yield reductions. The disease was found to severely affect varieties carrying the Texas cytoplasm (cms-T), which at the time accounted for 85% of the maize planting area (Ullstrup 1972). If a similar disease outbreak was to arise in sorghum the response would necessarily be slower due to the lack of an alternate manual method of male sterilization and the need to backcross female lines into an alternative cytoplasm and identify appropriate restorer lines. Despite the potential utility of A_2 cytoplasm (Moran and Rooney 2003), it is rarely used for commercial sorghum production. The main reason being the relatively low numbers of lines that restore fertility to A_2 cytoplasm. For example, Dahlberg and Madera-Torres (1997) found that in a set of 585 germplasm lines from Sudan 82% restored A_1 cytoplasm but only 28% restored A_2 cytoplasm.

In this study, we report on the mapping of a major fertility restoration locus, *Rf*$_5$, which restores both A_1 and A_2 cytoplasm, and we report on the mapping of a minor gene for partial fertility that incompletely restores fertility in A_1 cytoplasm. A series of informative SSRs are identified to facilitate marker-assisted breeding of new, commercially acceptable parental lines capable of restoring both A_1 and A_2 cytoplasm. Finally, a discussion is presented of those pentatricopeptide (PPR) gene family members that reside within the *Rf*$_5$ locus as plausible candidates for conditioning fertility restoration.

Materials and methods

Genetic stocks

Two hundred and eighteen F_6 recombinant inbred lines (RILs) were developed by single seed descent from a cross between the B line BTx642 (previously known as B35) and the restorer line, QL12. BTx642 is a maintainer or B line, which is male sterile in A_1 cytoplasm (A_1Tx642). QL12 has the pedigree KS19*5/Krish13 (developed by the DEEDI sorghum breeding program). QL12 is able to produce fertile hybrids in both A_1 and A_2 cytoplasm, and crosses between QL12 and other R lines that restore only A_1 cytoplasm have produced progeny with three different restoration phenotypes; (1) restoration of A_1 and A_2 cytoplasm, (2) restoration of A_2 cytoplasm only, and (3) inability to restore A_1 or A_2 cytoplasm.

To permit phenotyping of fertility restoration in F_6 RILs, F_1 hybrids were produced from the majority of the recombinant inbreds by crossing to two different CMS lines, A_1QL33 and A_2B963627-2. Both CMS lines were developed by the DEEDI breeding program with A_1QL33 having A_1 cytoplasm and A_2B963627-2 having A_2 cytoplasm. F_1 hybrids were successfully produced on both females for 199 of the 218 RILs with 212 RIL hybrids being made for A_1QL33 and 205 RIL hybrids for A_2 B963627-2. In addition F_1 hybrid seed was produced for the following five crosses; A_1QL33/QL12, A_2B963627-2/QL12, A_1QL33/BTx642 and A_1Tx642/QL12. An additional five commercial hybrids were included as fertility checks in the trial for a total of 427 entries.

To confirm the QTL results obtained with the RIL population, three F_2 populations were produced by selfing F_1 hybrids of A_1QL33/QL12, A_1Tx642/QL12 and A_2B963627-2/QL12 (see Table 1). In addition to the F_2 populations, a backcross population (BC_1F_1) was produced

Table 1 Description of F_2 and BC_1F_1 populations used for genetic linkage analyses

Population	Pedigree	Population size
F_2 R05460	A_1 QL33/QL12	69
F_2 R05461	A_1 Tx642/QL12	64
F_2 R05462	A_2 B963627-2/QL12	76
BC_1F_1 R07370	A_1 QL33//QL33/QL12	198

by crossing a fertile F_2 plant from the A_1QL33/QL12 population back to A_1QL33 (Table 1).

Field trials

Three separate field trials were conducted to assess fertility restoration phenotype in the various populations detailed above.

Trial 1

The 427 F_1 hybrid combinations (RIL hybrids and checks) were planted (14 November 2005) in 4 m long single row plots in a randomized trial with two replicates at the Hermitage Research Station in Warwick, Queensland. Standard agronomic practices were used and water was adequate for crop growth throughout the entire growing period. Temperatures during flowering were adequate for pollen production. The trial included five widely grown commercial hybrids as checks, 86G87 and Bonus MR from Pioneer Hi-Bred, MR Buster and MR 43 from Pacific Seeds and Liberty from the HSR Group. Prior to flowering three heads from each plot were covered with paper bags to exclude foreign pollen. At ~40 days after the crop flowered, the bags were removed and the percentage seed set on each head was visually rated. In addition, a qualitative visual rating was made of anther fertility of each hybrid at anthesis. This rating was developed by DEEDI and is a 1–9 score based on a visual evaluation of anther size, colour and morphology. Under this scale a rating of 1 was given to hybrids with very small non-pigmented sterile anthers with closed dehiscence pores, while a rating of 9 was given to fully male fertile plants with large pigmented anthers and open dehiscence pores. The characteristics associated with the scores are indicated below:

Score 1 = sterile very small colourless anthers; dehiscence pore absent,
Score 2 = sterile small colourless anthers; dehiscence pore absent,
Score 3 = sterile small but slightly larger again anthers, may be more coloured; dehiscence pore absent,
Score 4 = sterile small, but slight larger again, more coloured anthers; dehiscence pore absent,
Score 5 = sterile medium anthers, coloured; dehiscence pore absent,
Score 6 = partially fertile larger coloured anthers; dehiscence pore may be present in some anthers, pollen absent or present in very small quantities,
Score 7 = partially fertile, larger coloured anthers; dehiscence pore present in some anthers, some visible pollen,
Score 8 = moderately plump coloured anthers; dehiscence pore present in all anthers and pollen present,
Score 9 = plump coloured anthers; dehiscence pore present copious amounts of pollen.

This rating system has been found by DEEDI researchers to be well correlated with seed set and partial fertility in CMS lines in A_1 cytoplasm (Jordan DR, data not shown).

Trial 2

The three F_2 populations described above were planted at Hermitage Research Station on the 21 November 2006. Standard agronomic practices were used and water was adequate for crop growth throughout the entire growing period. Temperatures during flowering were adequate for pollen production. Each population was grown as a single row of spaced plants. Prior to flowering, a single head from each plant was covered with a paper bag to exclude foreign pollen. At ~40 days after the crop flowered, the bags were removed and the percentage seed set on each head was visually rated as described above.

Trial 3

Backcross population R07370 was planted at Hermitage Research Station on the 16 November 2007. Standard agronomic practices were used and water was adequate for crop growth throughout the entire growing period. Temperatures during flowering were adequate for pollen production. The population was grown as a single row of spaced plants. Prior to flowering, a single head from each plant was covered with a paper bag to exclude foreign pollen. At ~40 days after the crop flowered, the bags were removed and the percentage seed set on each head was visually rated.

Statistical analysis

All statistical analysis was conducted with the R package of software (http://www.R-project.org).

Genotype main effects

For trial 1 the 427 F_1 hybrid combinations (RIL hybrids and checks) were planted in a randomized trial with two replicates. The final model contains sources of variation for genotype and replicate effects in standard analysis of variance model. Means for the 427 genotypes were produced using a fixed genotype term and were used in the subsequent QTL analysis. Simple broad sense heritability for the two traits was calculated as the genetic variance over total variance.

Male by female effects

To maximize the power of the genetic design (i.e. a factorial combination of males and females) we fitted an analysis of variance model that allowed for a fixed female effect and a male within female interaction term. For both traits the correlation of male effects with the two females was high (greater than 0.8) permitting the prediction of an average male effect that uses information from both females. These average effects were also used for subsequent QTL analysis.

DNA extraction, genetic map and QTL analysis

Total genomic DNA was extracted from all populations as described by Diversity Arrays Technology (DArT^{TM} P/L, http://www.diversityarrays.com). DArT has been developed for sorghum and has been demonstrated to provide high quality markers for the construction of medium-density genetic linkage maps (Mace et al. 2008, 2009). A subset of ninety-four lines of the RIL population had previously been screened with DArT markers and a map consisting of 117 markers was constructed using the multipoint mapping software (http://www.multipoint.com) (Mace et al. 2009). This map showed strong alignment with a previously developed DArT consensus map of sorghum (Mace et al. 2009).

Composite interval mapping was conducted using the computer program QTL Cartographer for Windows v2.5 (Wang et al. 2010) on the data produced in trial 1. The data consisted of the two traits and two female combinations as well as an average effect calculated from the two female combinations to give a total of six trait × female combinations. Background markers for inclusion in the CIM model were selected by forward stepwise regression for each trait. The five most significant background markers were then used for analysis (default). The 'walking speed' was set at 2 cM and the 'window size' at 10 cM for CIM. A conservative permutation threshold at the 0.01 significance threshold was obtained for each trait using 1,000 permutations. 1-LOD and 2-LOD support intervals were determined, as described by Lander and Botstein (1989). The additive effects and percentage of variation explained (R^2) for all significant QTL were determined at their peak LOD values.

Based on the results of QTL analysis of the RILs in trial 1, SSR markers were identified flanking the QTL region from the consensus map (Mace et al. 2009) and were screened for polymorphism against the parents. Two polymorphic markers (*Xtxp65* and *Xtxp303*) were screened over the three F_2 mapping populations. PCR conditions were as previously detailed (Parh et al. 2008) and were visualised via capillary electrophoresis on a CEQ8800 (Beckman Coulter). Between 0.25 and 1 μL of PCR product was mixed with approximately 25 μL of sample loading solution and 0.3 μL of CEQ 400 size standard (Beckman Coulter) prior to a 35 min run at 6 kV. Fragment sizes were subsequently determined using the Fragment Analysis software (Beckman Coulter). Simple interval QTL analysis was conducted for the seed set trait in the four populations using the computer program QTL Cartographer for Windows, v2.5 (Wang et al. 2010). To further resolve the fertility restoration locus, an additional 10 polymorphic SSR markers distributed across the genomic region containing the QTL were developed by utilizing the sorghum whole genome sequence (http://www.phytozome.net) and SSRIT (Temnykh et al. 2001) and are detailed in the supplementary materials (Table S1). The BC_1F_1 population was screened with the new SSRs, and those markers that further delimited the locus were identified for potential use in marker-assisted selection.

Candidate gene identification

Based on results of the linkage analyses with all populations detailed above, an examination of predicted gene models encoded within the 2-LOD support interval of the fertility restoration locus was conducted to identify potential candidate genes. For this analysis, the current gene annotation provided by JGI (GeneModels SBI v1.4 at http://www.phytozome.net) was utilized and gene homology to rice *Rf1* (LOC_Os10g35240) determined through BLAST analysis (Johnson et al. 2008) based on the deduced amino acid sequence. Based on homology to rice *Rf1*, multiple sequence alignment analysis was conducted using Geneious, v4.7 (Drummond et al. 2009) on the deduced amino acid sequences of selected pentatricopeptide gene family members identified in the fertility restoration locus, in order to (1) determine homology to the proposed candidate for sorghum *Rf2* (Jordan et al. 2010) and sorghum *Rf1* and (2) to determine the relationship between the individual members of the PPR gene cluster

around the sorghum Rf_5 locus. A global alignment with free end gaps was used, with the Blosum62 cost matrix, a gap open penalty of 12 and a gap extension penalty of 3. Following multiple sequence alignment, an unrooted phylogenetic tree was constructed using the neighbor-joining method and the Jukes-Cantor genetic distance model, and the bootstrap resampling method (100 replicates). Following the identification of clusters of pentatricopeptide gene family members within the Rf_5 locus, other pentatricopeptide gene family members within and near the sorghum Rf_1 and Rf_2 loci were identified and phylogenetic analysis performed, using the same methodologies as detailed above.

Results

The seed set percentage and the anther fertility rating for the check hybrids and parental inbred lines are presented in Table 2. The commercial hybrids gave high ratings for the male fertility trait and were characterised by the presence of large coloured fertile anthers and the presence of open dehiscence pores. Seed set percentages in bagged heads of the commercial hybrids varied from 78 to 100%. Data was available for only two parental inbred lines due to poor establishment and/or lack of synchronization of flowering with the rest of the trial (Table 2). The line QL12 that was the source of the A_2 restoration trait gave high seed set and anther fertility ratings. The CMS (female) line A_1QL33 was highly male sterile and produced no seed under bags. Experience from a number of years of sorghum nurseries have shown that the CMS line A_1Tx642 generally produces intermediate anther fertility ratings and may produce low percentages of seed under bags in conditions favouring male fertility, while its male-fertile maintainer line BTx642 has moderate to high fertility and produces high seed sets (Jordan DR, unpublished data). In nurseries the CMS line A_2B963627-2 had a higher anther fertility rating relative to A_1QL33, but has not been observed to produce seed under bags (Jordan DR, unpublished data).

Table 3 shows the population statistics for the BTx642/QL12 RIL population grown in hybrid combination with the two CMS females (A_1QL33 and A_2B963627-2).

Genetic effects for the RILs for both traits in the two genetic backgrounds and overall were highly statistically significant. In general the mean seed set and anther fertility ratings for the hybrids with A_2B963627-2 were higher than those for the hybrids with A_1QL33. The heritability of both traits in the two females was high ranging from 86 to 89%. Figure 1 shows the distribution of seed set percentage in the two sets of RIL hybrids, three F_2 populations and one BC_1F_1 population. Each individual is represented by a column with height being equivalent to the percentage seed set (0–100%). The distribution of the seed set ratings of the RIL hybrids based on the two CMS lines showed a somewhat bimodal distribution. This was particularly the case for hybrids with the hybrids with A_2B963627-2 as the parent where $\sim$75% of the lines had seed sets of >80% or less than 10% (see Fig. 1). While this was suggestive of the segregation of a major gene, the fact that both sets of hybrids also contained many individuals with intermediate seed set percentages, indicated possible segregation of minor (partial) genes for fertility restoration in these populations (Fig. 1).

Correlations between the best linear unbiased estimates for the two male fertility traits in combination with the two females indicated that the phenotypic scores were highly correlated, both within and between females (data not shown). Correlation coefficients ranged from a low of $r = 0.80$ (between seed set for the A_1QL33 hybrids and

Table 2 Seed set percentages and anther fertility ratings for parental inbred lines and check hybrids

Cytoplasm	Type	Genotype	Seed set (%)	Anther fertility rating
–	Parental inbred	QL12	95	9.0
A_1	Parental inbred	A_1 Tx642	–	–
–	Parental inbred	BTx642	–	–
A_2	Parental inbred	A_2 B963627-2	–	–
A_1	Parental inbred	A_1 QL33	0	2
A_2	Parental check hybrid	A_2 B963627-2/QL12	93.8	9.0
A_1	Parental check hybrid	A_1 QL33/QL12	95.4	9.0
A_1	Parental check hybrid	A_1 QL33/BTx642	0.3	3.5
A_1	Commercial hybrid	86G87	78.8	9.0
A_1	Commercial hybrid	MR Buster	86.3	9.0
A_1	Commercial hybrid	Bonus MR	87.5	8.5
A_1	Commercial hybrid	MR 43	90.0	9.0
A_1	Commercial hybrid	Liberty	100.0	9.0

–, data unavailable; 1, fully sterile anthers; 9, fully fertile anthers

Table 3 Population statistics for seed set percentage and anther fertility ratings (1–9 scale with 1 denoting high levels of male sterility and 9 denoting lines which are male fertile) for RIL hybrids grown in trial 1

CMS parent	Statistic	Seed set (%)	Anther fertility rating
A_1 QL33	Mean	43	6.4
	Maximum	100	9
	Minimum	0	2
	Heritability	87	89
A_2 B963627-2	Mean	50	7.3
	Maximum	100	9
	Minimum	0	2
	Heritability	87	86

fertility rating for the A_2B963627-2 hybrids) and a high of $r = 0.95$ (between seed set for the A_1QL33 hybrids and seed set for the A_2B963627-2 hybrids). These results indicate that the anther fertility rating is a good measure of final seed set percentage, and further indicates that the genetic control of the anther fertility rating and seed set are similar in the two genetic backgrounds.

The results of the QTL analysis for the trait-by-female combinations investigated in trial 1 are shown in Fig. 2, and the QTL statistics are summarized in Table 4. The major QTL for fertility restoration detected in the RIL population was located on SBI-05 with flanking markers spanning 1,412,322–3,630,618 bp, as determined via sequence mapped using an in silico mapping strategy analogous to e-PCR, as detailed in Mace and Jordan (2010). The QTL was significant for both phenotypic traits in both females as well as for the average effect indicating that this region is associated with fertility restoration in both A_1 and A_2 cytoplasm. The major QTL on SBI-05 explained between 57 and 69% of the variation in the anther fertility rating and between 42 and 45% of the variation in percent seed set depending on the cytoplasm (A_1 vs. A_2).

A minor QTL for fertility restoration was identified on SBI-04. This region contained a significant QTL for the anther fertility rating in A_1 cytoplasm as well as a QTL for the average effect for the male fertility and seed set percentage. The QTL on SBI-04 explained 10 and 16% of the variation in the average effect of seed set percentage and anther fertility rating, respectively. The fact that the two average effect traits were associated with QTL is probably due to the increased phenotypic precision obtained by combining the data from the two female parents.

Figure 3 shows the mean seed set percentage of hybrids of RIL-by-CMS females classified by QTL allele class. The presence of the QTL fertile allele (genome region derived from male restorer parent) on SBI-05 in the absence of the QTL fertile allele on SBI-04 (genome region derived from female parent) produced quite different average percent seed sets in A_1QL33 versus A_2B963627-2 (66 and 83%, respectively). In combination with the fertile allele at the SBI-05 QTL, the fertile alleles at the SBI-04 QTL increased the percent seed set in A_1QL33 hybrids to 79% (vs. 66% with only the SBI-05 QTL fertile allele), whereas the average seed set of A_2B963627-2 was unchanged by the presence of the fertile allele at the SBI-04 QTL (83%). The presence of the fertile allele at the SBI-04 QTL and sterile allele at the SBI-05 QTL resulted in average seed set percentages of 4 and 7%, respectively, in A_1QL33 and A_2B963627-2 hybrids. The majority of RILs that carried both sterile alleles at the SBI-04 and SBI-05 QTL produced no seed regardless of the female parent.

With the objective of confirming and refining the QTL loci detected for fertility restoration in the RIL population, a series of backcross and F_2 populations in A_1 or A_2 cytoplasms were characterized for seed set percentage. The details of the identified QTL are summarised in Table 5 and Figs. 4 and 5. In all F_2/backcross populations examined, the results confirm the presence of a major QTL for fertility restoration from QL12 on SBI-05 as previously identified in the two RIL hybrid populations, with the estimated dominance effect (D) ranging from 13.7 (R05461) to 49.5 (R05460). The amount of phenotypic variation explained by the QTL varied between 37 and 96% depending on the given population. F_2 populations R05461 (A_1 Tx642/QL12) and R05462 (A_2 B963627-2/QL12) contained many individuals with intermediate fertility (see Fig. 1), which suggests the likely presence of partial fertility genes segregating in these populations as evidenced by the lower R^2 value. This was particularly apparent in F_2 population R05461 where there were no F_2 individuals with zero seed set. In both populations derived from crosses between QL12 and A_1QL33, the R^2 values are very high (96.2–96.4%) suggesting that a single restoration locus on SBI-05 is conditioning the majority of the fertility restoration phenotype. The segregation in seed set percentage phenotypes in Fig. 1 support this hypothesis with both of these populations showing very clear differentiation between fertile and sterile genotypes with the vast majority of genotypes having seed sets greater than 70% or less than 5%. In F_2 population R05460 (A_1QL33/QL12) the average seed set in fertile plants was 95% (Fig. 1). In contrast the seed set in fertile plants in the BC_1F_1 R07370 population (A_2 QL33//QL33/QL12) was 73% (Fig. 1). The difference in the seed set percentage between these two populations may be the result of a range of factors including different environmental conditions in trial 2 compared with trial 3, the possible absence of the QTL on SBI-04 in the individual F_2 plant used to produce the

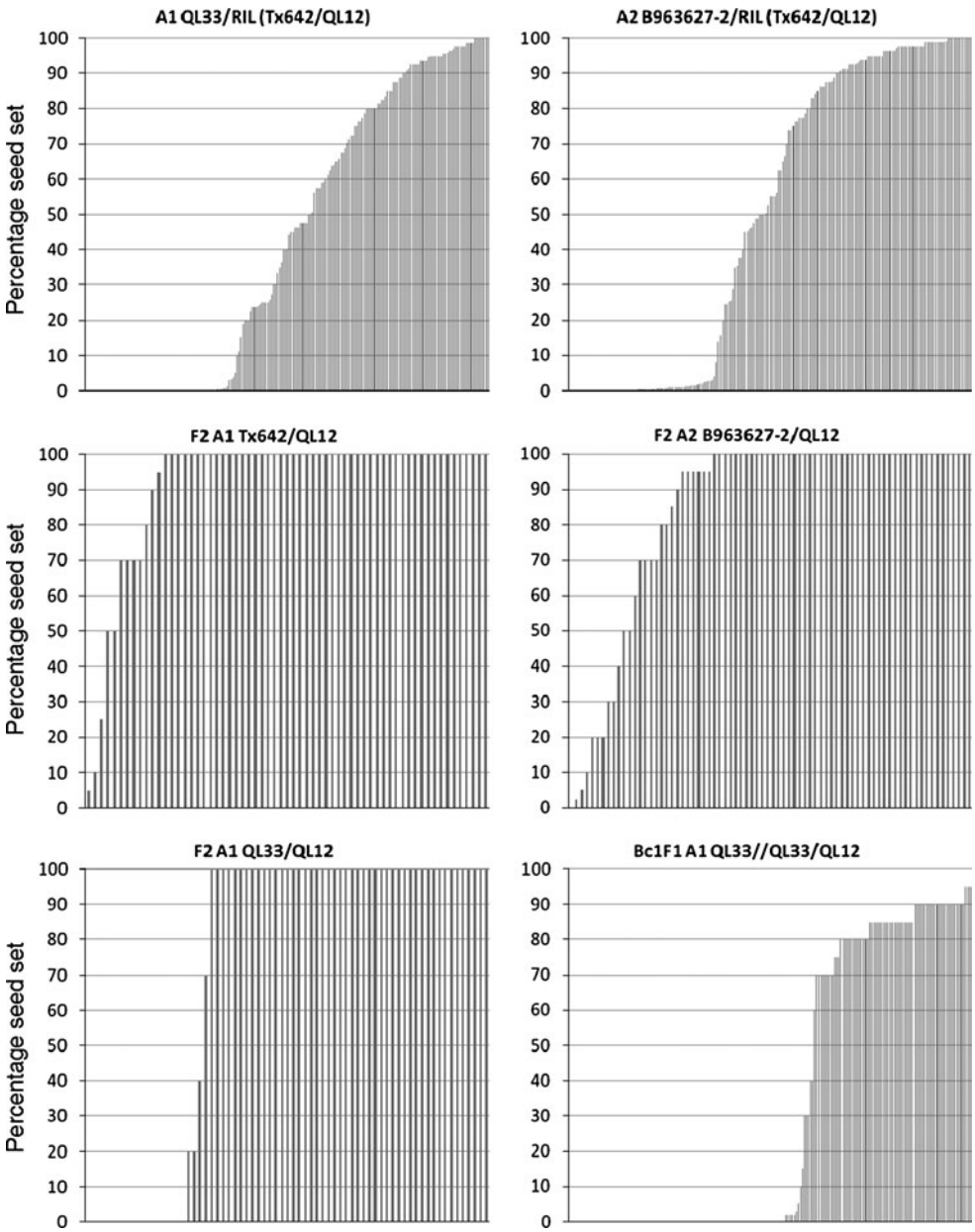

Fig. 1 Distribution of seed set percentage for each of the six mapping populations. Each individual column depicts seed set percentage for a particular genotype. Seed set percentage data from the two RIL populations are presented as means while the raw individual plant data are shown for the three F_2 and one BC_1F_1 populations

population, or the absence of homozygous individuals for the *Rf*$_5$ locus in the backcross population.

Preliminary examination of *Rf*$_5$ candidate genes

We conducted a preliminary examination of potential candidate genes within the *Rf*$_5$ fertility restorer locus on SBI-05. With the exception of *Rf*$_2$ of maize (Cui et al. 1996), all fertility restoration genes identified to date, including the restorer genes for A_1 cytoplasm in sorghum (Klein et al. 2005; Jordan et al. 2010), are members of the PPR gene family; therefore, we focused on this family as potential candidate genes for *Rf*$_5$. The QTL region on SBI-05, defined between markers Sb05-2429989 and Sb05QGM280 (Table 6), spans 584,564 bp and contains 70 predicted open reading frames (ORFs), averaging one gene for every 8.35 and 132 kb per cM. Of the 70 predicted genes in the region, 57 have been annotated as *bona fide* protein coding genes with high confidence support (http://www.phytozome.net/sorghum). In total, seven predicted genes in this region (four with high confidence support) were annotated as encoding full-length proteins belonging to the PPR gene family, all of which showed homology to the rice *Rf*$_1$ gene (Os10g35240, see Table 6). Five of these genes have predicted transcript lengths similar to rice *Rf*$_1$ (2,273–2,774 bp) and are predicted to encode proteins ranging in size from 734 to 924 amino acids. In contrast, genes Sb05g002320 and

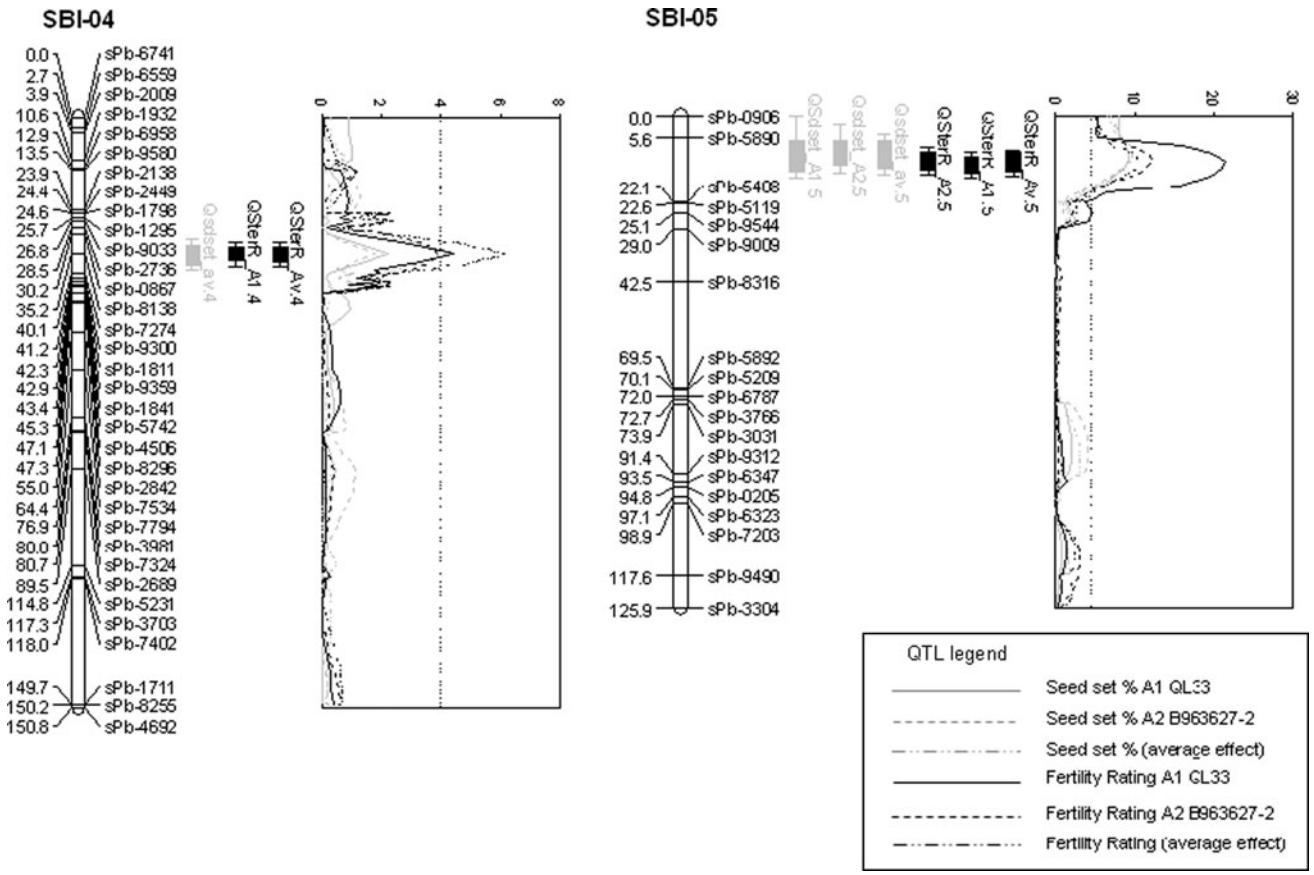

Fig. 2 QTL analysis of fertility restoration related traits in a set of recombinant inbred lines grown in hybrid combination in A_1 and A_2 cytoplasm

Table 4 Summary of QTL analyses for seed set percentage and the anther fertility rating identified from 94 RILs tested in hybrid combination with A_1- and A_2-CMS inbreds, detailing the peak LOD location in base-pairs (bp), marker interval, peak LOD score, significant LOD threshold for each trait, phenotypic variation explained (R^2) and the estimation of the additive effect

Trait	Female	QTL_ID	SBI	Peak LOD position (bp)	Marker interval	LOD	LOD sig. thresh	R^2	Additive effect
Anther fertility rating	A_1 QL33	QSterR_A1.4	04	10,418,853	sPb-0867/sPb-8138	4.45	2.93	12.3	1.129
Seed set (%)	Ave. effect	QSdSet_av.4	04	10,418,853	sPb-0867/sPb-8138	4.34	2.93	10.4	12.56
Anther fertility rating	Ave. effect	QSterR_av.4	04	10,418,853	sPb-0867/sPb-8138	6.16	2.91	16.4	1.322
Anther fertility rating	A_1 QL33	QSterR_A1.5-1	05	2,521,470	sPb-5890/sPb-5408	21.43	2.93	69.5	2.555
Anther fertility rating	A_2 B963627-2	QSterR_A2.5	05	2,521,470	sPb-5890/sPb-5408	12.11	2.93	59.2	1.659
Anther fertility rating	Ave. effect	QSterR_av.5	05	2,521,470	sPb-5890/sPb-5408	10.54	2.91	57.3	1.925
Seed set (%)	A_1 QL33	QSdSet_A1.5	05	1,966,896	sPb-5890/sPb-5408	9.28	3.04	41.9	21.311
Seed set (%)	A_2 B963627-2	QSdSet_A2.5-1	05	1,966,896	sPb-5890/sPb-5408	9.65	2.97	44.8	24.041
Seed set (%)	Ave. effect	QSdSet_av.5	05	1,966,896	sPb-5890/sPb-5408	9.74	2.93	43.1	22.951

Sb05g002606 have predicted transcript lengths of 1,376 and 1,449 bp, respectively, and are predicted to encode for PPR proteins of 456 and 406 amino acids, respectively. These latter two annotated PPR genes are similar in length to the sorghum *Rf*$_1$ (1,697 bp; 566 amino acids) and the putative candidate gene for *Rf*$_2$ (950 bp, 317 amino acids). The seven PPR genes within the *Rf*$_5$ locus show distinct clustering on the physical map, three of them (Sb05g002310, Sb05g002320, Sb05g002330) are adjacent to each other and separated from a second group of two adjacent ORFs (Sb05g002360, Sb05g002370) by two gene models and the final group of two ORFs Sb05g002606 and Sb05g002620, is separated from this group by 26 predicted gene models.

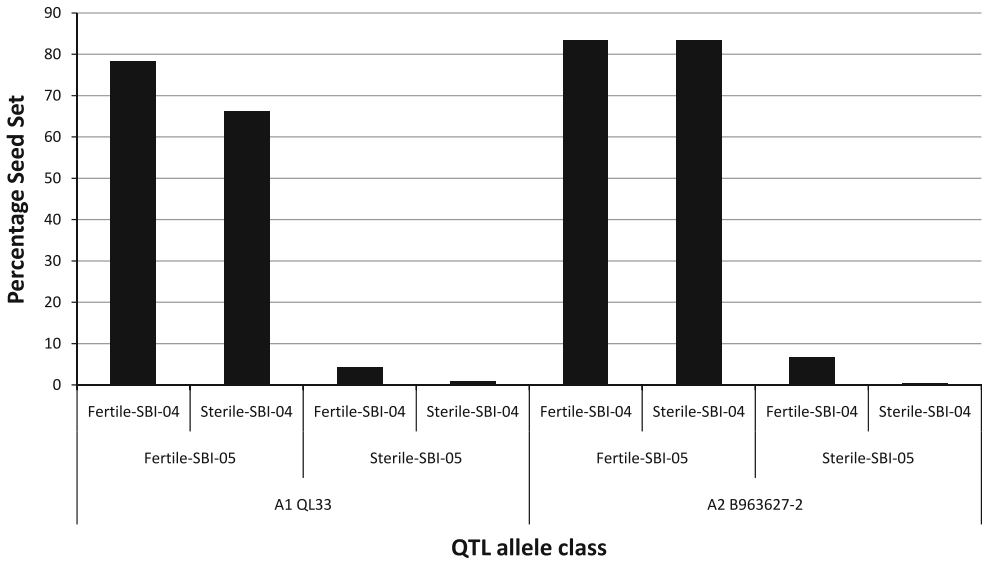

Fig. 3 Mean percentage seed set for RILs tested in hybrid combination in A_1- and A_2-cytoplasm clustered by QTL allele class

Table 5 Summary of QTL identified on SBI-05 in the F_2 and BC_1F_1 populations for pollen fertility restoration in A_1 and A_2 cytoplasms, detailing the marker interval, their physical location in base-pairs (bp), peak LOD score, significant LOD threshold for each population, phenotypic variation explained (R^2), the estimation of additive (A) and dominance effects (D) and the ratio of dominance effect to additive effect for each QTL (d/a)

Population	Pedigree	Locus delimiting markers	Locus size (bp)	Peak LOD position (bp)	LOD	LOD sig. thresh	R^2	A	D
F_2 R05460	A_1 QL33/QL12	SB05-2429989/SB05-3693459	1,263,470	2,689,283	51.9	1.95	96.4	47.9	47.1
F_2 R05461	A_1 Tx642/QL12	*Xtxp65*/Sb05QGM279	1,785,960	2,126,098	4.5	1.85	36.8	15.2	13.7
F_2 R05462	A_2 B963627-2/QL12	*Xtxp65*/Sb05QGM279	1,108,089	2,390,326	9.3	1.95	80.9	35.1	35.3
BC_1F_1 R07370	A_1 QL33//QL33/QL12	Sb05QGM273/Sb05QGM280	573,915	2,673,137	122.6	1.51	96.2	81.56	–

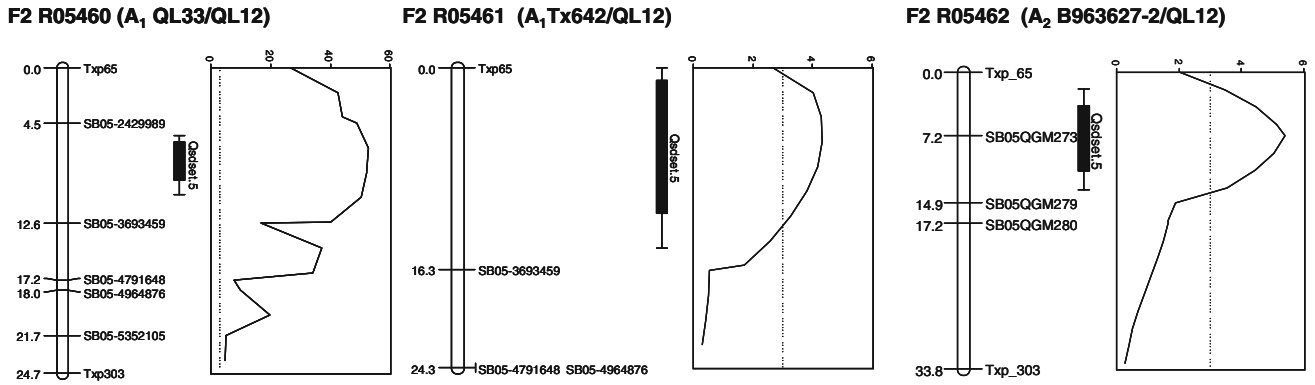

Fig. 4 QTL analysis of seed set on chromosome SBI-05 for three F_2 populations segregating for a major fertility restoration allele

PPR genes cluster near sorghum fertility restoration loci

Further examination of the region spanning known *Rf* genes in sorghum reveals evidence of regional gene duplication. In the 2.6 Mbp region spanning the Rf_5 locus, a total of 13 PPR gene family members were identified with homology to rice Rf_1 through BLAST analysis. To understand the possible relationships between the individual members of the PPR gene cluster flanking the Rf_5 locus, a phylogenetic analysis was performed with Geneious, v4.7 software using deduced amino acid sequences while using rice Rf_1 as a reference sequence. The multi-sequence alignment and unrooted phylogenetic tree (Fig. 6, ESM

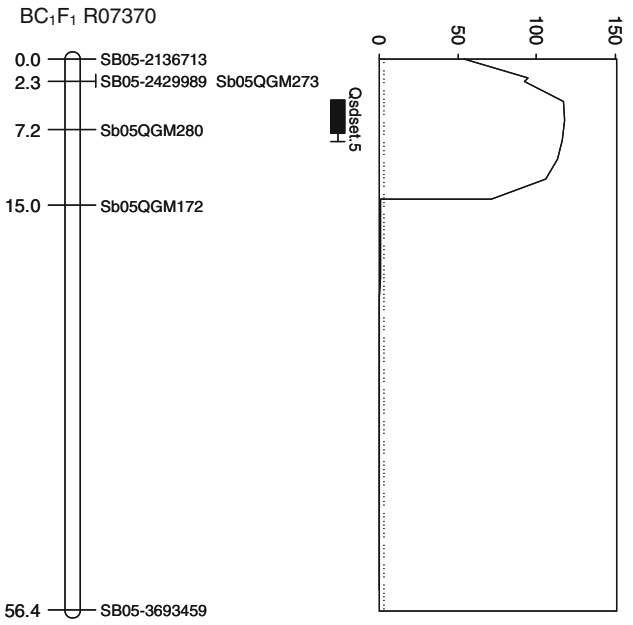

Fig. 5 QTL analysis of seed set on chromosome SBI-05 of BC_1F_1 R07370 population segregating for a major fertility restoration allele in A1 cytoplasm

Fig. S1) shows the relationship between the individual members of the PPR gene cluster around the sorghum *Rf*$_5$ locus and rice *Rf*$_1$. The amino acid sequences were fairly highly conserved between the PPR gene cluster members, with the exception of Sb05g002210 and Sb05g003220. Phylogenetic analyses were also performed for the PPR genes clustering near the *Rf*$_2$ and *Rf*$_1$ loci in sorghum (ESM Figs. S2, S3 and S4). In general, the PPR genes showed a much higher degree of similarity within each *Rf* locus (*Rf*$_1$, *Rf*$_2$ or *Rf*$_5$) than between *Rf* loci (ESM Fig. S4). The PPR gene clusters around *Rf*$_2$ were closely related to the cluster flanking *Rf*$_5$ both of which show a high degree of similarity to rice *Rf*$_1$; by contrast, the PPR gene cluster around sorghum *Rf*$_1$ is quite distinct from family members clustering near *Rf*$_2$ and *Rf*$_5$.

Discussion

In this paper, we identify a previously undescribed major fertility restoration locus located on sorghum chromosome SBI-05. Unique to this fertility restoration gene is its ability to restore fertility in both A_1 and A_2 cytoplasm. Four other major fertility restoration genes, *Rf*$_1$–*Rf*$_4$, have been detailed in sorghum. The other two major genes known to restore fertility to A_1 cytoplasm, *Rf*$_1$ and *Rf*$_2$, are located on SBI-08 and SBI-02, respectively (Klein et al. 2005; Jordan et al. 2010). *Rf*$_3$ and *Rf*$_4$ are associated with restoration of A_3 cytoplasm (Tang et al. 1998), but A_3 cytoplasm appears to have negative effects on grain yield (Moran and Rooney 2003) and is therefore unlikely to be suitable for commercial hybrid seed production. In line with the previous sorghum restoration gene nomenclature, we have designated the major fertility restoration locus described herein as *Rf*$_5$.

Previous research indicates that fertility restoration in sorghum is under relatively complex genetic control involving a mixture of major and minor gene action (Maunder and Pickett 1959; Miller and Pickett 1964). The observation of a continuum of variation in seed set in all of

Table 6 List of genes identified showing homology to PPR proteins between markers Sb05-2429989 and Sb05QGM280 and their homology to rice *Rf*$_1$ (Os10g35240)

Gene ID	Start	End	Gene length (bp)	Protein length	JGI annotation	BLASTN		BLASTP	
						e Value	Score	*e* Value	Score
Sb05g002310[a]	2,446,361	2,448,760	2,399	799aa	"similar to protein *Rf1*, mitochondrial precursor"	2.7e−218	4933	3.7e−231	2223
Sb05g002320	2,454,310	2,455,685	1,376	456aa	"similar to PPR protein"	2e−171	2290	1.9e−206	2091
Sb05g002330[a]	2,459,297	2,461,692	2,395	734aa	"similar to protein *Rf1*, mitochondrial precursor"	1.7e−176	3084	6.3e−156	1513
Sb05g002360	2,525,727	2,528,404	2,677	833aa	"similar to protein *Rf1*, mitochondrial precursor"	3.5e−195	2294	8.1e−140	1361
Sb05g002370[a]	2,541,958	2,544,231	2,273	757aa	"similar to *Rf1* protein, mitochondrial, putative, expressed"	7.4e−173	3927	4.4e−187	1807
Sb05g002606	2,770,011	2,771,460	1,449	406aa	"weakly similar to putative chloroplast RNA processing protein"	4.6e−65	1051	3.4e−47	492
Sb05g002620[a]	2,774,110	2,776,884	2,774	924aa	"similar to PPR protein"	1.1e−215	3838	3.5e−187	1808

[a] Annotated as *bona fide* protein coding genes with high confidence support (http://www.phytozome.net/sorghum)

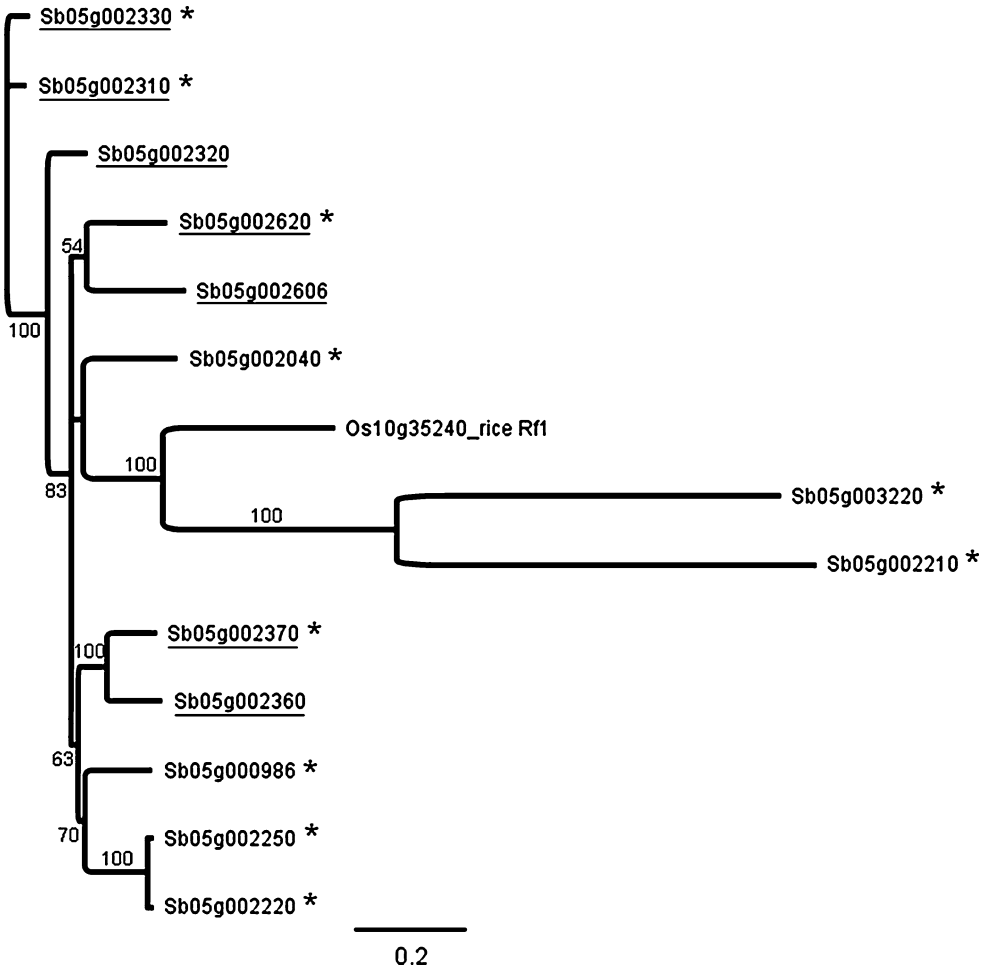

Fig. 6 Unrooted phylogenetic tree of rice Rf_1 and PPR gene cluster in the Rf_5, locus. Branch labels indicate bootstrap support values; branch length represents the number of substitutions per site. *Underlined* genes represent the seven genes within the QTL defined region; *asterisked* genes represent *bona fide* protein coding genes with high confidence support

the populations observed in this study is consistent with the presence of modifiers or partial fertility genes contributing to full restoration as indicated by classical genetic studies (Miller and Pickett 1964). The minor QTL for fertility restoration identified on SBI-04 generates variation in seed set percentages similar to those attributed to the Pf_1 or Pf_2 loci identified by Miller and Pickett (1964). It should be noted that depending on the parental lines, the range of fertility restoration phenotypes differed considerably, with some populations exhibiting a distinct bimodal distribution while other populations showed a broad continuum of restoration phenotypes. This variation in fertility restoration likely reflects the segregation of a series of partial fertility restoration genes, with the number of minor genes being dependent on the specific genetic background.

Within A_1 cytoplasm of sorghum, there are no obvious differences in the restoration phenotype of Rf_5 compared with the previously identified Rf_1 and Rf_2 genes. This similarity in phenotype suggests that the three major *Rf* genes have a common or related cellular function. All *Rf* genes identified to date, except Rf_2 of maize (Cui et al. 1996), are members of the PPR family. Sequence analysis of the Rf_1 locus in sorghum identified a PPR protein that cosegregates with the fertility restoration phenotype (Klein et al. 2005). A PPR gene was also identified as the likely candidate for Rf_2 in sorghum (Jordan et al. 2010). Given the strong association between the PPR gene family and fertility restoration in a range of crops species, and the fact that Rf_5 produces an identical phenotype to Rf_1 and Rf_2 in restoring A_1 cytoplasm in sorghum, it is likely that Rf_5 is also a PPR gene family member.

In this study, linkage analyses delimited the Rf_5 locus to a ~584 kb region of the sorghum genome on SBI-05 that is predicted to encode 70 genes. Of these 70 genes, seven were identified as PPR genes with six PPR genes exhibiting strong sequence homology to Rf_1 of rice (BLASTP *e* values ranging from 3.7e−231 to 8.1e−140). The large number of homologous PPR genes clustering near the Rf_5 locus suggests recent gene duplication in this region of the genome. Clustering of PPR genes has been observed within *Rf* loci in many other species including rice (Kato et al. 2007), maize (Xu et al. 2009), Petunia (Bentolila et al. 2002), Brassica (Brown et al. 2003) and Arabidopsis (Lurin et al. 2004). Similarly in sorghum, the Rf_1 locus contains five tightly

linked, closely related genes within 20 cM of the candidate gene (Sb08g015260, Sb08g015300, Sb08g019750, Sb08g019910, and Sb08g020970) and the *Rf*$_2$ locus contains five tightly linked (within 1 cM of each other), closely related genes (Sb02g004520, Sb02g004530, Sb02g004626, Sb02g004810, Sb02g005000). 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. These authors also observed that nucleotide sequence conservation was not limited to the duplicated genes but was also present in their flanking non-coding regions, suggesting that the complex structure of the *Rf*$_1$ locus in maize may have been generated by homologous recombination. We also observed nucleotide sequence conservation in the 5′ flanking regions of the duplicated PPR genes in the *Rf*$_5$ region, which could also support this theory. Thus, placed in context of the growing body of evidence for fertility restoration genes being PPR gene family members, it can be postulated that the *Rf*$_5$ restoration phenotype is conditioned by one or multiple PPR genes residing within the locus. In spite of this strong evidence, further experimental observations are necessary to confirm the identity of the gene or genes associated with the *Rf*$_5$ restoration phenotype. Even if the one or more of these genes controls fertility restoration, the tight clustering of these genes means that recombination between them will be very rare and conventional fine mapping approaches are unlikely to provide sufficient resolution. As a result, functional analyses will likely be required to determine the causal gene or genes. Given that this locus restores fertility to both A_1 and A_2 cytoplasm, there is a distinct possibility that two or more of the linked PPR genes present in this locus are functional *Rf* genes.

The existence of restorer genes that enable the production of fertile F_1 hybrids using cytoplasmic genetic male sterility is essential to the cost-effective production of hybrid sorghum seed. Currently commercial hybrid sorghum seed production relies almost exclusively on the A_1 cytoplasm. Such reliance on a single cytoplasm potentially leaves the industry vulnerable to an epidemic of a disease as occurred with the southern corn leaf blight (*Bipolaris maydis*). Unlike the maize seed industry, the sorghum industry could not rapidly change to an alternative hybrid production technology to deal with such a threat. Of the alternatives to the A_1 cytoplasm, to date only the A_2 cytoplasm has been demonstrated to have similar agronomic performance to A_1 cytoplasm and appears to be suitable for commercial hybrid seed production (Moran and Rooney 2003; Reddy et al. 2007). In fact in the 1990s in Australia a successful commercial hybrid (DK48M) was marketed using the A_2 cytoplasm (Jordan DR, unpublished data). While it is relatively straightforward to convert female parents to A_2 cytoplasm via backcrossing, the low frequency of male parents that can restore fertility in A_2 cytoplasm has limited its utility in breeding programs. The results of this study indicate that the *Rf*$_5$ gene is able to restore fertility in both A_1 and A_2 cytoplasm at a level consistent with that observed in commercial sorghum hybrids. The availability of the linked SSR markers described in this paper (or SNPs discovered through resequencing technology of target genotypes) would make the process of introgressing *Rf*$_5$ into different genetic backgrounds relatively straightforward. Deploying this gene in commercial breeding programs would provide a degree of insurance for the genetic vulnerability associated with the exclusive use of the A_1 cytoplasm, and would permit commercial companies to rapidly switch to A_2 cytoplasm if required.

In the present study, a partial fertility gene was mapped to SBI-04, which is consistent with the presence of modifiers or partial fertility genes as indicated by the classical genetic studies contributing to full pollen fertility restoration (Miller and Pickett 1964). These modifiers or partial fertility genes have an important impact on sorghum breeding. In applied hybrid breeding programs, *Pf* genes must be excluded from female parent lines. In practice this can be difficult because *Pf* genes are often only unmasked when new female parents are backcrossed into sterile cytoplasm. In addition, in some cases this partial fertility is only expressed under particular environmental conditions and may pass unnoticed for generations (Jordan et al. 2010). Partial pollen fertility in female parents in seed production blocks can result in considerable commercial losses, and commercial breeders are very cautious about making crosses between non-restorers and full or partial restorers when the objective is the development of new female parents. As a result, it is rare for crosses between B and R (or exotic) lines to be made. This constraint has markedly restricted genetic diversity observed in the A/B line germplasm pool as observed in a number of diversity studies (Ahnert et al. 1996; Menz et al. 2004; Mace et al. 2008). In contrast there is strong selection in the case of male parent lines for combinations of restorer genes that give complete restoration of A_1 cytoplasm in the heterozygous state under a range of environmental conditions. Again failure to restore fertility can be commercially expensive and crosses between non-restorers and restorers aimed at developing new male parents are rarely made. In both types of parents, but particularly female parents, these constraints restrict access to genetic variation limiting potential genetic gain for quantitative traits. The use of markers to track the location of genes associated with CMS male fertility restoration and partial fertility will allow breeders to use a more diverse range of parents in crosses with a greater degree of confidence in being able to efficiently recover acceptable male or female parents as required.

Conclusion

The identification of the major restorer genes *Rf*$_1$, *Rf*$_2$, and *Rf*$_5$ and one of the genetic regions involved in partial fertility provides critical tools for developing a detailed understanding of the genetic, cellular, and physiological architecture of this important trait. The present research also extends our knowledge of those genes conditioning fertility restoration, and further expands our molecular tools needed to effectively manipulate this commercially important trait. Knowledge of *Rf* gene location within the genome will permit more effective strategies for developing male and female parents by allowing breeders to identify the most effective alleles, and to ensure that these alleles are present (or absent) in their elite material by marker-assisted selection. Our on-going effort to map and clone a complete complement of major and minor restorer genes in sorghum will provide breeders with an effective arsenal with which to manipulate fertility restoration in their breeding programs.

Acknowledgments The authors wish to thank Julie McCollum for her expert technical assistance. We thank the Australian Grains Research and Development Corporation (GRDC; http://www.grdc.com.au) for part of the financial support for this research, and the USDA-Agricultural Research Service for financial support to RRK.

References

Ahnert D, Lee M, Austin DF, Livini C, Woodman WL, Openshaw SJ, Smith JSC, Porter K, Dalton G (1996) Genetic diversity among elite sorghum inbred lines assessed with DNA markers and pedigree information. Crop Sci 36:1385–1392

Bentolila S, Alfonso AA, Hanson MR (2002) A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. Proc Natl Acad Sci USA 99:10887–10892. doi:101073/pnas102301599

Brown GG, Formanova N, Jin H, Wargachuk R, Dendy C, Patil P, Laforest M, Zhang JF, Cheung WY, Landry BS (2003) The radish *Rfo* restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J 35:262–272. doi:101046/j1365-313X200301799x

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Cui XQ, Wise RP, Schnable PS (1996) The *rf2* nuclear restorer gene of male-sterile T-cytoplasm maize. Science 272:1334–1336

Dahlberg JA, Madera-Torres P (1997) Restorer reaction in A_1 (ATx623), A_2 (A2Tx632), and A_3 (A3SC103) cytoplasms to selected accessions from the Sudan sorghum collection. Int Sorghum Millet Newsl 38:43–58

Drummond A, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2009) Geneious v4.7. http://www.geneious.com/

Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL (2008) NCBI BLAST: a better web interface. Nucl Acids Res 36:W5–W9. doi:101093/nar/gkn201

Jordan DR, Mace ES, Henzell RG, Klein PE, Klein RR (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(7):1279–1287

Kato H, Tezuka K, Feng YY, 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. doi:101038/sjhdy6801026

Klein RR, Klein PE, Mullet J, Minx P, Rooney WL, Schertz KF (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. Theor Appl Genet 111:994–1012

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits by using RFLP linkage maps. Genetics 136:1447–1455

Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A, Le Ret M, Matrin-Magnietee M-L, Mireau H, Peeters N, Renou J-P, Szurek B, Taconnat L, Small I (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16:2089–2103

Mace ES, Jordan DR (2010) Location of major effect genes in sorghum [*Sorghum bicolor* (L.) Moench]. Theor Appl Genet 121:1339–1356

Mace ES, Xia L, Jordan DR, Halloran K, Parh DK, Huttner E, Wenzl P, Kilian A (2008) DArT markers: diversity analyses and mapping in *Sorghum bicolor*. BMC Genomics 9:26. doi:26101186/1471-2164-9-26

Mace ES, Rami JF, Bouchet S, Klein PE, Klein RR, Kilian A, Wenzl P, Xia L, Halloran K, Jordan DR (2009) A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. BMC Plant Biol 9:13

Maunder AB, Pickett RC (1959) The genetic inheritance of cytoplasmic-genetic male sterility in grain sorghum. Agron J 51:47–49

Menz MA, Unruh NC, Klein PE, Mullet JE, Rooney WL, Klein RR (2004) Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. Crop Sci 44:1236–1244

Miller DA, Pickett RC (1964) Inheritance of partial male-fertility in *Sorghum vulgare* Pers. Crop Sci 4:1–4

Moran JL, Rooney WL (2003) Effect of cytoplasm on the agronomic performance of grain sorghum hybrids. Crop Sci 43:777–781

Parh DK, Jordan DR, Aitken EAB, Mace ES, Jun-ai P, McIntyre CL, Godwin ID (2008) QTL analysis of ergot resistance in sorghum. Theor Appl Genet 117:369–382

Reddy BVS, Ramesh S, Reddy PS, Ramaiah B (2007) Combining ability and heterosis as influenced by male-sterility inducing cytoplasms in sorghum [*Sorghum bicolor* (L.) Moench]. Euphytica 154:153–164. doi:101007/s10681-006-9281-6

Schertz KF (1983) Potentials with new cytoplasmic male sterility systems in sorghum. Proc Genet Soc USA 38:1–10

Schmitz-Linneweber C, Small I (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci 13:663–670

Stephens JC, Holland RF (1954) Cytoplasmic male-sterility for hybrid sorghum seed production. Agron J 46:20–23

Tang HV, Chang RY, Pring DR (1998) Cosegregation of single genes associated with fertility restoration and transcript processing of sorghum mitochondrial *orf107* and *urf209*. Genetics 150:383–391

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11:1441–1452

Ullstrup AJ (1972) The impacts of southern corn leaf blight epidemics of 1970–1971. Annu Rev Phytopathol 10:37–50

Wang S, Basten C, Zeng Z-B (2010) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh. http://statgenncsuedu/qtlcart/WQTLCarthtm

Xu XB, Liu ZX, Zhang DF, Liu Y, Song WB, Li JS, Dai JR (2009) Isolation and analysis of rice *Rf1*-orthologus PPR genes co-segregating with *Rf3* in maize. Plant Mol Biol Rep 27:511–517. doi:101007/s11105-009-0105-4