Theor Appl Genet (2017) 130:539–547
DOI 10.1007/s00122-016-2833-9

ORIGINAL ARTICLE

Allelism analysis of *BrRfp* locus in different restorer lines and map-based cloning of a fertility restorer gene, *BrRfp1*, for *pol* CMS in Chinese cabbage (*Brassica rapa* L.)

Huamin Zhang[1] · Junqing Wu[1] · Zihui Dai[1] · Meiling Qin[1] · Lingyu Hao[1] · Yanjing Ren[1] · Qingfei Li[1] · Lugang Zhang[1]

Received: 4 September 2016 / Accepted: 18 November 2016 / Published online: 28 November 2016

Abstract

Key message **In Chinese cabbage, there are two *Rf* loci for *pol* CMS and one of them was mapped to a 12.6-kb region containing a potential candidate gene encoding PPR protein.**

Abstract In Chinese cabbage (*Brassica rapa*), polima cytoplasmic male sterility (*pol* CMS) is an important CMS type and is widely used for hybrid breeding. By extensive test crossing in Chinese cabbage, four restorer lines (92s105, 01s325, 00s109, and 88s148) for *pol* CMS were screened. By analyzing the allelism of the four restorer lines, it was found that 92s105, 01s325, and 00s109 had the same "restorers of fertility" (*Rf*) locus (designated as *BrRfp1*), but 88s148 had a different *Rf* locus (designated as *BrRfp2*). For fine mapping the *BrRfp1* locus of 92s105, a BC_1F_1 population with 487 individuals and a BC_1F_2 population with 2485 individuals were successively constructed. Using simple sequence repeat (SSR) markers developed from *Brassica rapa* reference genome and InDel markers derived from whole-genome resequencing data of 94c9 and 92s105, *BrRfp1* was mapped to a 12.6-kb region containing a potential candidate gene encoding pentatricopeptide repeat-containing protein. Based on the nucleotide polymorphisms of the candidate gene sequence between the restoring and nonrestoring alleles, a co-segregating marker SC718 was developed, which would be helpful for hybrid breeding by marker-assisted screening and for detecting new restorer lines.

Communicated by C. F. Quiros.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-016-2833-9) contains supplementary material, which is available to authorized users.

✉ Lugang Zhang
lugangzh@163.com

[1] State Key Laboratory of Crop Stress Biology for Arid Areas, College of Horticulture, Northwest A&F University, Yangling 712100, Shaanxi, China

Introduction

Cytoplasmic male sterility (CMS), a maternally inherited trait that prevents plants from producing functional pollen, is a widespread phenomenon in higher plants. The previous studies have led to the belief that unusual and often chimeric open reading frames (ORFs) arising from mitochondrial genome rearrangement are responsible for CMS (Hanson and Bentolila 2004; Chase 2007; Fujii and Toriyam 2008; Chen and Liu 2014). In many cases, specific dominant nuclear genes termed "restorer of fertility" (*Rf*) have been found, which can suppress the male sterile phenotype at different molecular levels, such as genomic, post-transcriptional, translational or post-translational, and metabolic levels, and restore the fertility in plants carrying the CMS mitochondrial genome (Chen and Liu 2014). CMS/*Rf* systems have not only facilitated hybrid seed production, but are also an ideal model for studying cytoplasmic inheritance and the crosstalk between the mitochondrial and nuclear genomes.

The restoration of one CMS-associated gene by several different *Rf* genes is a common phenomenon in many plants. In maize, *Rf2*, in combination with one of three other *Rf* genes (*Rf1*, *Rf8*, and *Rf**) can fully or partially restore the fertility of T-CMS maize (Wise et al. 1999). In radish, Ogura CMS can be restored by two *Rf* genes, *Rfo* and *RsRf3* (Uyttewaal et al. 2008; Wang et al. 2013). Similarly, A1-CMS of sorghum and BT-CMS, HL-CMS, and

WA-CMS of rice also have two *Rf* genes (Liu et al. 2004; Klein et al. 2005; Wang et al. 2006; Jordan et al. 2010; Luo et al. 2013). To date, more than ten *Rf* genes have been isolated from different plant species. These include *Rf2* for maize T-CMS (Cui et al. 1996), *PPR592* for petunia *pcf*-CMS (Bentolila et al. 2002), *Rfo* and *Rfk* for radish Ogura-CMS and Kosena-CMS (Brown et al. 2003; Desloire et al. 2003; Koizuka et al. 2003), *PPR13* for sorghum A1-CMS (Klein et al. 2005), *Rf1a* and *Rf1b* for rice BT-CMS (Wang et al. 2006), *Rf17* for rice CW-CMS (Fujii and Toriyam 2009), *Rf2* for rice LD-CMS (Itabashi et al. 2011), *Rf5* for rice HL-CMS (Hu et al. 2012), *Rf4* for rice WA-CMS (Tang et al. 2014), *bvORF20* for sugar beet Owen CMS (Kitazaki et al. 2015), and *Rfp* for rapeseed *pol* CMS (Liu et al. 2012, 2016). All the cloned *Rf* genes encode a pentatricopeptide repeat-containing (PPR) protein, except *Rf2* of maize, *bvORF20* of sugar beet, and *Rf17* and *Rf2* of rice.

Polima CMS (*pol* CMS), discovered by Fu (1981), was the first practically valuable CMS used in hybrid breeding of rapeseed. Since the mapping of *Rfp*-to-*B. napus* linkage group 18 by Jean et al. (1997), several researchers have identified tightly linked molecular markers at the *Rfp* locus. Recently, Liu et al. (2012) mapped the *Rfp* locus to a 29.2-kb region on chromosome A09 of *B. napus* and predicted that a PPR gene was the most likely candidate for *Rfp*. Using transgenic rapeseed plants, they confirmed that the PPR gene, indeed, functions as *Rfp* (Liu et al. 2016).

Pol cytoplasm has been transferred into Chinese cabbage (*Brassica rapa* L. AA) by interspecific cross and has become an important CMS type for hybrid breeding in Chinese cabbage (Ke and Song 1989; Ke et al. 1992). By extensive test crossing, several restorer lines that can fully or partially restore *pol* CMS have been screened in Chinese cabbage (Ke and Zhang 1993; Zhang and Ke 1994). However, it remains unclear whether the *Rf* genes of different restorer lines were the same. In the restorer line 01s325, Xu et al. (2014) mapped *BrRfp* locus on chromosome A09 and identified two markers, SSR03 and SSR2528, which co-segregated with the *BrRfp* locus. However, the restorer genes have not been cloned and the molecular mechanism of fertility restoration by the *pol* CMS/*Rf* system needs to be investigated.

In this study, two *Rf* loci were confirmed from four restorer lines of *pol* CMS in Chinese cabbage by allelism analysis. Moreover, using simple sequence repeat (SSR) markers developed from *B. rapa* reference genome and InDel markers derived from whole-genome resequencing of two parental lines, *BrRfp1* of restorer line 92s105 was mapped to a 12.6-kb region that contained a potential candidate gene encoding PPR. Based on the candidate gene sequences of the restoring and nonrestoring alleles, a molecular marker co-segregating with *BrRfp1* locus was developed, which would be useful for breeding via marker-assisted selection (MAS) and for the screening of new restorer lines.

Materials and methods

Plant materials

Four restorer lines, 92s105, 01s325, 00s109, and 88s148, with very large phenotypic differences were screened from Chinese cabbage germplasms (Ke and Zhang 1993; Zhang and Ke 1994). The lines 92s105, 01s325, and 00s109 could fully restore the fertility of *pol* CMS (Supplementary Fig. S1c), whereas 88s148 could only partially restore it (Supplementary Fig. S1b). The line 94c9 is a *pol* CMS line. For allelism analysis of the restorer loci in the four restorer lines, five crosses, 92s105 × 01s325, 92s105 × 00s109, 00s109 × 01s325, 01s325 × 88s148, and 92s105 × 88s148, were performed to produce the F_1 generation. These five F_1 plants were subsequently test-crossed with 94c9. For fine mapping of the restorer gene, the BC_1F_1 population was produced from the backcross between 94c9 (the acceptor parent) and 92s105 (the donor parent), and one male fertile (heterozygous in *BrRfp1* locus) individual from the BC_1F_1 population was self-pollinated to produce BC_1F_2 population.

All the seeds used in this study were obtained from the Chinese cabbage research team at the College of Horticulture, Northwest A&F University, Yangling, China. All the plants were grown from seedlings in a greenhouse in the early January and were transplanted to the experimental field at the Northwest A&F University, in late February of 2015 and 2016. The plants were vernalized naturally. The male fertility of each individual was assessed by observing at least three flowers during the flowering period twice, at 7-day interval.

Development of SSR and InDel markers

To develop SSR markers, sequences containing five SSR markers identified from the restorer line 01s325 by Xu et al. (2014) were downloaded from the *B. rapa* reference genome (http://brassicadb.org/brad/). The SSR sites were searched using the software SSRHunter (Li and Wan 2005). To develop InDel markers, genomes of the two parental lines, 94c9 and 92s105, were sequenced with HiSeq X Ten (Macrogen, Shenzhen, China) at 99- and 91-fold sequencing depths. After removing duplication and finding the variant with Isaac Variant Caller (IVC), the information for each variant was classified by chromosomes or scaffolds based on the *B. rapa* reference genome available at NCBI Website. Only the insertions or deletions >4 bp were chosen to design the InDel primers. The designing of these primers was based on the *B. rapa* reference genome region, and

Table 1 Results of *BrRfp* allelism analysis of four restorer lines

Cross combination	Number of fertile plants	Number of partially fertile plants	Number of Sterile plants	Mendelian expectations	χ^2 ($P > 0.05$)
94c9 × (92s105 × 01s325)	302				
94c9 × (00s109 × 01s325)	320				
94c9 × (92s105 × 00s109)	248				
94c9 × (01s325 × 88s148)	175	90	91	2:1:1	0.107
94c9 × (92s105 × 88s148)	141	70	69	2:1:1	0.021

150-bp upstream and downstream of the InDel sites, respectively. The primers for the SSR and InDel markers were designed by Primer Premier 5.0 (http://www.premierbiosoft.com/primerdesign/). All the primers used in this study were synthesized by Sangon Biotech Co., Ltd (Shanghai, China).

Genomic DNA extraction, PCR, and gel electrophoresis

Genomic DNA was extracted from young leaves by the cetyl trimethylammonium bromide (CTAB) method as described by Porebski et al. (1997). PCR was performed in 10-μL reaction mixtures containing 1-μL DNA template (50 ng/μL), 0.3-μL forward primer (10 μM), 0.3-μL reverse primer (10 μM), and 5-μL 2 × Taq Master Mix for PAGE (Vazyme, Nanjing, China). The PCR protocol was as follows: 94 °C for 4 min followed by 28–30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s, and a final extension at 72 °C for 5 min. The PCR products were separated on 9% nondenaturing polyacrylamide gel.

Construction of genetic and physical maps and identification of candidate gene

The segregation data of the polymorphic molecular markers and the phenotypic data for each individual in the population were used to construct the genetic linkage map by the JoinMap 4.0 software with a logarithm of the odds (LOD) threshold score of 6.0. The physical map of *BrRfp1* locus was constructed using the physical position of each marker on the chromosome in the NCBI *B. rapa* reference genome. Gene annotation of the delimited region between the two closest markers was performed according to the annotation results of *B. rapa* reference genome in NCBI (http://www.ncbi.nlm.nih.gov/genome/?term=Brassica+rapa) and BRAD (*Brassica* database, http://brassicadb.org/brad).

Cloning and sequence analysis of the candidate gene

The specific primers for the candidate gene were designed according to the reference genome of *B. rapa*. The PCR products were purified using a DNA Gel Extraction Kit (Tiangen, Beijing, China) and used for TA cloning. The purified PCR products were introduced into the pMD18-T Simple Vector (Takara, Dalian, China) and transformed into *E. coli* strain DH5α. The recombinant plasmids were sequenced by Sangon Biotech Co., Ltd (Shanghai, China). The complete coding sequences of the eight Chinese cabbage lines were submitted to GenBank under the following accession numbers: 92s105 (KX671967), 01s325 (KX671968), 00s109 (KX671969), 12CF29 (KX671970), xiakang2 (KX671971), 94c9 (KX671972), 12CF9 (KX671973), and 88s148 (KX671974). Sequence alignment was performed with MegAlign of the DNA Star Lasergene package (http://www.dnastar.com). The PPR motif prediction of the candidate gene was carried out with PPRFinder Online (http://www.plantppr.com).

Results

Allelism analysis of the restorer locus in the four restorer lines

To analyze the allelism of the restorer locus, five populations were generated (Table 1). All the testcross progenies generated from 92s105 × 01s325, 92s105 × 00s109, and 00s109 × 01s325 were fully fertile. However, the testcross progenies generated from 01s325 × 88s148 and 92s105 × 88s148 were fertility segregates and the segregation ratio of fully fertile to partially fertile and male sterile plants best fitted a 2:1:1 ratio (Table 1). These results indicated that the restorer genes of 92s105, 01s325, and 00s109 were at the same locus (designated as *BrRfp1*), but the restorer gene of 88s148 was at a different locus (designated as *BrRfp2*). The results of allelism analysis were consistent with the phenotypic observation that 92s105, 01s325, and 00s109 restore the fertility of *pol* CMS completely, whereas 88s148 restores the fertility of *pol* CMS partially. In addition, *BrRfp1* is epistatic to *BrRfp2* in the restoration of fertility for *pol* CMS.

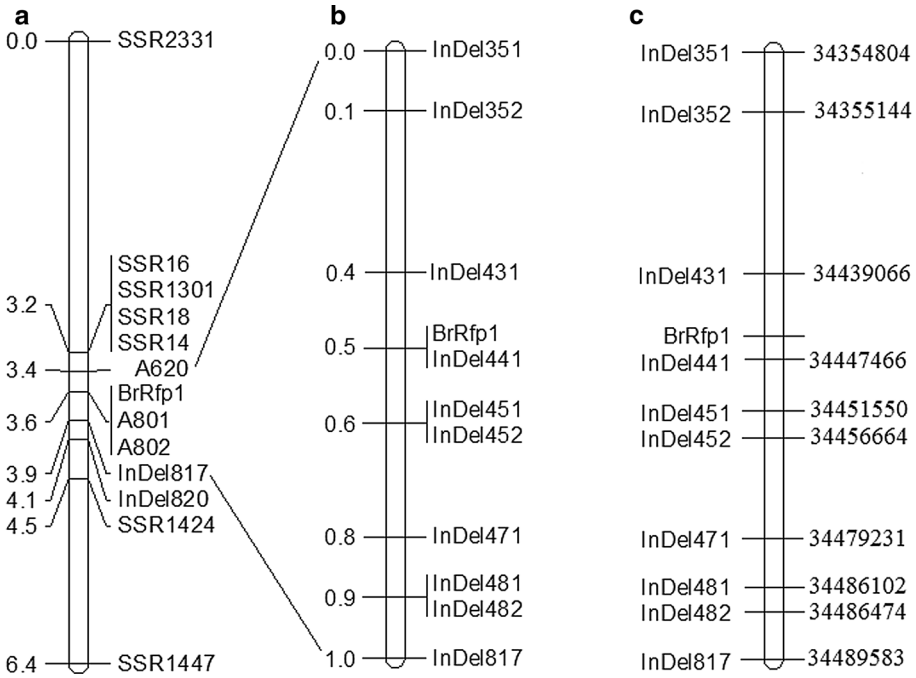

Fig. 1 Genetic and physical maps of the *BrRfp1* locus. **a** Genetic map of *BrRfp1* locus based on 487 BC_1F_1 population. **b** High-resolution linkage map of *BrRfp1* locus based on 2485 BC_1F_2 population. **c** Physical map of molecular markers in **b** on the chromosome A09 (the physical position on the *right side* is based on the *B. rapa* reference genome in NCBI)

Genetic analysis of *BrRfp1* and preliminary mapping

A BC_1F_1 population containing 487 individuals was generated from a backcross between the male sterile line 94c9 and the restorer line 92s105. Phenotypic investigation showed the segregation of 247 fertile individuals and 240 sterile individuals. The χ^2 test revealed that the segregation ratio was consistent with the expected ratio of 1:1 ($\chi^2 = 0.101$, $P > 0.05$), indicating that the male fertility was controlled by a single dominant gene, the same as the restorer line 01s325 (Xu et al. 2014).

For preliminary mapping of *BrRfp1*, five SSR markers on chromosome A09, identified by Xu et al. (2014) from the restorer line 01s325, were first screened in the two parental lines of BC_1F_1 population; however, no polymorphism was observed except for SSR2331. Based on the *B. rapa* reference genome sequence containing these five SSR markers, 121 more SSR markers were developed. Six markers (SSR14, SSR16, SSR18, SSR1301, SSR1424, and SSR1447) (Supplementary Table S1) exhibited polymorphism in the two parental lines. Thereafter, these seven SSR markers were used to assay 487 individuals of the BC_1F_1 population and the *BrRfp1* locus was flanked by SSR1301 and SSR1424 with a genetic distance of 0.4 and 0.8 cM, respectively (Fig. 1a). To further narrow the delimited region of *BrRfp1* locus, 100 primer pairs that amplified about 1.0-kb products were designed. Three primer pairs (A620, A801, and A802) only amplified products from the restorer line 92s105; these three primer pairs were used as dominant markers. When the primer pairs could amplify product from both the parental lines, the product was sequenced to locate the InDel site. Two such sites were found after sequencing the amplified product and then developed into InDel markers (InDel817 and InDel820). Using these five markers (Supplementary Table S1), the previously selected recombinants were tested. The *BrRfp1* locus, co-segregating with A801 and A802, was localized to a region between A620 and InDel 817 (Fig. 1a). The genetic distances of A620 and InDel817 were both 0.2 cM, and the physical distance between A620 and InDel 817 was about 132.5 kb.

Fine mapping of *BrRfp1* locus using whole-genome resequencing data

To fine map the *BrRfp1* locus, a BC_1F_2 population with 2485 individuals generated from one heterozygous male fertile individual in the BC_1F_1 population was produced. To develop more InDel markers in the delimited region, the genome sequences of the two parental lines, 94c9 and 92s105, were obtained by whole-genome resequencing. Based on the whole-genome resequencing data, 24 InDel markers were developed and 9 of them showed polymorphism in the parental lines (Supplementary Table S1). A620, a dominant marker, which is not suitable for detecting recombination events in BC_1F_2 population, was replaced by a co-dominant marker, Indel351, distanced about 1.5 kb from A620. InDel351 and InDel817 were used to screen the recombination events among BC_1F_2 populations. A total of 41 recombinants were detected, including 24 recombination events for InDel351 and 17 recombination events for InDel817. These recombinants were further

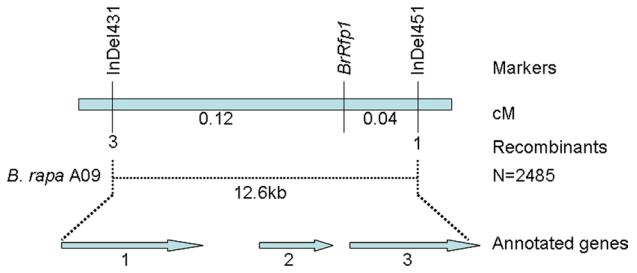

Fig. 2 Identification of *BrRfp1* as a candidate gene. *BrRfp1* is restricted to the region between the markers InDel 431 and InDel 451. This region is about 12.6 kb on the chromosome A09 based on the reference genome of *B. rapa* and there are three annotated genes in this region

Table 2 Gene annotation in mapped region of *B. rapa* reference genome

Gene name	Gene annotation[a]
LOC103843088	Pentatricopeptide repeat-containing protein At1g12620-like
LOC103843201	Pollen-specific protein SF3-like
LOC103843089	Keratin, type II cytoskeletal 2 epidermal

[a] Gene annotation based on NCBI *Brassica rapa* annotation release 100

detected by eight other InDel markers. The segregation data of these ten markers and the male fertility survey data of BC_1F_2 population were applied to construct a high-resolution linkage map (Fig. 1b). All these markers were mapped at a 1.0-cM region around the *BrRfp1* locus. Three of these markers were located on one side of the *BrRfp1* gene, six were on the other side, and one (InDel441) co-segregated with the *BrRfp1* gene. The order of these markers on the linkage map was consistent with their physical position on chromosome A09 of *B. rapa* reference genome (Fig. 1c). Finally, the *BrRfp1* gene of Chinese cabbage was mapped to a region, about 12.6 kb, between InDel431 and InDel451, with a genetic distance of 0.12 cM and 0.04 cM, respectively (Fig. 2).

BrRfp1 candidate gene identification and sequence analysis

According to the gene annotations of this 12.6-kb region in BRAD and NCBI, there are three genes in this region (Table 2), but only one (LOC103842088) encodes for the PPR protein. Therefore, LOC103842088 was considered as the candidate gene of *BrRfp1* in Chinese cabbage.

Using the primers Rfp088 designed against *B. rapa* genome sequences present in NCBI, the full-length sequences of LOC103842088 from the two parental lines, 92s105 and 94c9, were cloned and sequenced. The sequences analysis revealed that LOC103842088 from 92s105 and 94c9, both had no introns and encoded a PPR protein, belonging to P-class, with 650 amino acids comprising 16 PPR motifs. There were 47 single-nucleotide polymorphisms (SNPs) in the coding region between 92s105 and 94c9 (Supplementary Fig. S2). These SNPs had no effect on the deduced protein length but resulted in mutations of 32 amino-acid residues (Supplementary Fig. S3). Liu et al. (2016) verified that *Rfp*, the restorer gene for *pol* CMS in rapeseed, was a PPR gene on chromosome A09. The deduced protein sequences of the *BrRfp1* candidate gene and *Rfp* were compared. There was a difference in five amino-acid residues; however, none of them were located on the eight key code positions indicated by Liu et al. (2016) (Supplementary Fig. S3).

To identify the functional variations between the restoring *BrRfp1* and nonrestoring *Brrfp1* alleles in Chinese cabbage, the genes from other three restorer lines (01s325, 00s109, and 88s148) and three male sterile lines (xiakang2, 12CF9, and 12CF29) were sequenced. The deduced protein sequences from 01s325 and 00s109 were identical to each other, and had differences in three amino-acid residues compared to 92s105. In 88s148, the sequence of the first six PPR motifs was the same as that of 92s105, but the sequence after ten PPR motifs was the same as that of 94c9 (Supplementary Fig. S3). These results are consistent with the allelism analysis and phenotypic observation in the four restorer lines. In addition to the two rapeseed sequences, *Rfp* and *rfp* (Liu et al. 2016), ten deduced protein sequences (including four restoring genes and six nonrestoring genes) were aligned. Among these deduced protein sequences, differences in three amino-acid residues located in the 11th PPR motif of all the six nonrestoring genes were found (Fig. 3).

Development of a co-segregation molecular marker of *BrRfp1*

Based on the nucleotide polymorphisms of the candidate gene, *BrRfp1*, between the restorer lines and sterile lines, a dominant marker SC718 (Supplementary Table S1) was developed (Supplementary Fig. S2). Individuals (2485 in number) of BC_1F_2 population were detected using SC718 and the results were entirely consistent with the phenotype (Fig. 4). To detect the transferability of SC718 in the different populations, the F_2 population, $14F_2Q_{21}$, derived from the cross 94c6 × tkxf was tested using SC718 and another co-segregating marker, InDel441. SC718 also co-segregated with the phenotype, but InDel441 showed no polymorphism (data not shown). Therefore, SC718 could be used as a reliable molecular marker for MAS breeding and screening of new restorer germplasms containing the *BrRfp1* locus in Chinese cabbage.

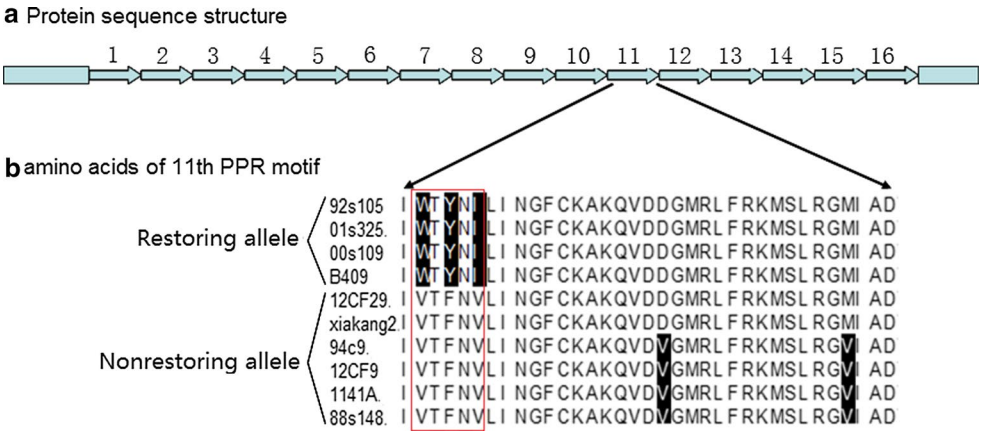

Fig. 3 Structure and homology analyses of the *BrRfp1* sequence. **a** Deduced protein sequence structure of *BrRfp1*. **b** Sequence alignment of the eleventh PPR motif in four restoring and six nonrestoring alleles. The mutant amino acids are indicated in *black*. The *red rectangle* indicates the mutations at all the nonrestoring alleles (color figure online)

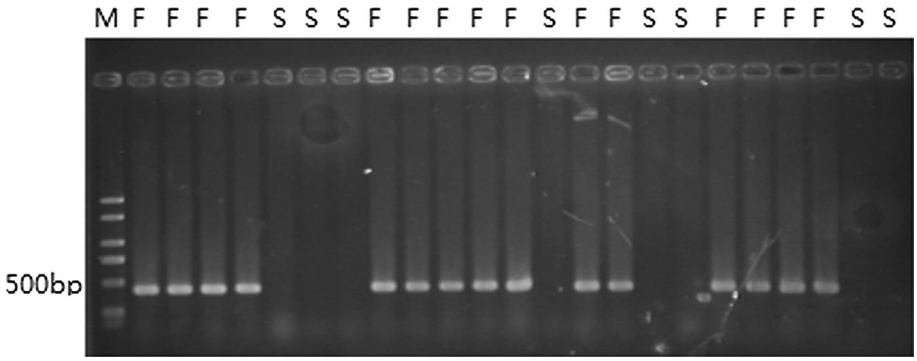

Fig. 4 Co-segregation analysis in the partial BC_1F_2 population using SC718 marker. *M* DL2000 marker, *F* male fertile plant, *S* male sterile plant

Discussion

Among the four restorer lines of Chinese cabbage used in this study, 92s105, 01s325, and 00s109 could fully restore *pol* CMS, but 88s148 could only partially restore *pol* CMS. Both allelism and candidate gene sequence analysis indicated that the restorer locus of 88s148 was a new locus different from other three restorer lines. The *pol* cytoplasm of Chinese cabbage was introduced from rapeseed by an interspecific cross (Ke and Song 1989, Ke et al. 1992). In both the Chinese cabbage and rapeseed, restorer lines having different restorer capabilities were determined (Fu et al. 1989; Ke and Zhang 1993). The phenomenon of the partial restoration of infertility was also found in maize. In maize, except for the full restorer gene *Rf1*, two partial restorer genes, *Rf8* and *Rf**, for T-CMS were found, which had a distinct way of processing the T-*urf13* mitochondrial transcripts (Dill et al. 1997; Wise et al. 1999). In Chinese cabbage and rapeseed, it had been found that some *pol* CMS lines were sensitive to temperature (Fu et al. 1989; Zhang and Hao 2001). These germplasms could be applied to "two lines" hybrid breeding. The previous studies were all focused on mapping and cloning of the restorer genes with full restorer capability, and rarely considered the partial restorer and temperature sensitive restorer. The cloning of partial restorer and temperature sensitive restorer genes would provide more insights into the molecular mechanisms of nuclear–mitochondrial interactions in the *pol* CMS/*Rf* system.

For the development of both the SSR and InDel markers, a great deal of genomic sequence information is required. The release of Chinese cabbage whole-genome sequence (Wang et al. 2011) was a great convenience for the development of the SSR markers; however, this strategy is inefficient and time consuming. In this study, only six (5%) SSR markers showed polymorphism between 94c9 and 92s105 among the 121 SSR markers derived from *B. rapa* reference genome sequences. Furthermore, 100 primer pairs, which were used to develop the InDel markers after sequencing and comparison of the amplification product, were designed according to the sequences of the delimited region downloaded from the *B. rapa* reference genome. Using this method, only two (2%) InDel markers were developed. For developing more InDel markers, the whole genomes of 94c9 and 92s105 were sequenced at about 100-fold sequencing depths. By comparing the sequenced genomes to the *B. rapa* reference genome, a large number of SNPs and InDels were found. Only in the preliminary mapped region (about 132.5-kb), there were 1520 SNPs and 280 InDels. Twenty-four InDel primers were designed

and synthesized; nine (37.5%) of them showed amplification polymorphism. These InDels could greatly accelerate the process of fine mapping the target gene. Finally, *BrRfp1* from the restorer line 92s105 was successfully mapped to a 12.6-kb region. Therefore, the whole-genome resequencing is a very efficient method to develop InDel markers.

When *BrRfp1* locus was mapped to a 132.5-kb region using the BC_1F_1 population, there were two genes, Bra026927 (BRAD) and LOC103843088 (NCBI), distanced about 50.3 kb, which encoded the PPR proteins. The phenomenon of the distribution of PPR genes in clusters is also found in other mapped restorer loci. For example, in the petunia *Rf* locus, there are two PPR genes adjacent to each other, sharing 93% similarity in the predicted protein sequences (Bentolila et al. 2002). The *Rfo* locus of radish was also found to contain three clustered genes encoding similar PPR proteins (Desloire et al. 2003; Brown et al. 2003). In rice, the *Rf* loci of BT-CMS, HL-CMS, and WA-CMS were all mapped to a region about 310 kb on the chromosome 10 containing nine genes encoding the PPR proteins (Komori et al. 2004; Wang et al. 2006; Hu et al. 2012; Tang et al. 2014). In *Arabidopsis thaliana*, most *Rf-PPR-like* (*RFL*) genes are mainly clustered in two regions on chromosome 1 (Fujii et al. 2011). *Rf* and *RFL* genes are clustered in the chromosomal regions possibly because of their rapid and dynamic evolution through gene duplication and functional divergence (Fujii et al. 2011).

Because the restorer lines for *pol* CMS were found from *B. napus*, *B. rapa*, and *B. juncea* (AABB), the restorer gene for *pol* CMS was believed to be present on the A genome of the Brassica genus (Fu et al. 1989; Yang and Fu 1991). Liu et al. (2016) verified a PPR gene on the chromosome A09 of rapeseed, which functioned as an *Rfp* in transgenic plants. In rapeseed, the changes in eight amino-acid residues located at positions 2, 5, and 35 of the PPR motifs were considered as the cause of functional divergence. In this study, *BrRfp1* of the restorer line 92s105 was also mapped to a PPR gene on the chromosome A09, which shared high similarity to *Rfp* in *B. napus*, with only five different amino-acid residues present in the predicted protein sequences. In the restorer lines 01s325 and 00s109, the deduced protein sequences of *BrRfp1* were highly similar to *Rfp*, having only two different amino-acid residues (Supplementary Fig. S3). All the different amino-acid residues were not in the key position pointed out by Liu et al. (2016). Therefore, the restorer gene, *BrRfp1*, for *pol* CMS in Chinese cabbage is probably identical to the restorer gene, *Rfp*, in *B. napus*. Chinese cabbage and rapeseed, which originated independently in China and Europe, are two species of the Brassica genus. Although Chinese cabbage and rapeseed have independently evolved more than 1000 years (Nagaharu 1935) ago, they may have the same restorer gene for *pol* CMS. The probable reason for this could be that the cytoplasm and restorer gene of *pol* CMS originated from ancient *B. rapa* and then integrated into *B. napus* genome with the hybridization between ancient *B. rapa* and ancient *B. oleracea*. Whether the restorer gene for *pol* CMS in *B. juncea* is identical to that in *B. rapa* and *B. napus* needs further investigation.

In angiosperms, the PPR gene family is one of the largest gene families in nearly all the genomes sequenced so far (Cheng et al. 2016). In *B. rapa*, there are 592 PPR members; 347 of them are P-class PPR proteins (http://www.plantppr.com). Almost all the PPR proteins are targeted to plastids and mitochondria, where they play important roles in post-transcriptional processes, including RNA stabilization, RNA cleavage, intron splicing, RNA editing, and in the initiation of translation (Barkan and Small 2014). One PPR motif has the ability to recognize a single nucleotide, and a specific RNA sequence is recognized by an array of PPR domains (Yin et al. 2013; Barkan and Small 2014). A few studies reported that the amino-acid residues at the positions 2, 5, and 35 of a PPR motif were the molecular determinants for RNA-binding specificity (Fujii et al. 2011; Yin et al. 2013). For example, there are two different amino-acid substitutions at the position 5 in two different PPR motifs of a nonrestoring allele of *Rfk* in radish CMS (Koizuka et al. 2003). An amino-acid mutation at the position 35 of the third PPR motif results in RFL9 losing its function of processing the *rps3* and *orf240* transcript in the mitochondria of *Arabidopsis thaliana* (Arnal et al. 2014). The comparison of the four restoring and six nonrestoring alleles for *pol* CMS in Chinese cabbage and rapeseed revealed that the three amino-acid residue variations at the positions 2, 4, and 6 of the 11th PPR motif were present in all the nonrestoring alleles (Fig. 3). The alteration of Trp^{406} to Val, caused by the change of TG^{1216} to GT at the position 2, could probably account for the loss of the restoration function (Supplementary Fig. S3).

Author contribution statement HZ constructed the mapping populations, performed marker development and mapping analysis, and wrote the paper. JW assisted with phenotypic survey and DNA extraction, ZD, MQ, and LH assisted to extract DNA. YR and QL provided valuable research ideas. LZ designed and supervised the study.

Acknowledgements This work was supported by grants from The National Key Research and Development Program of China (2016YFD0101701) and the National Science and Technology Support Program of China (2014BAD01B0802).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards The authors declare that this study complies with the current laws of the countries in which the experiments were performed.

References

Arnal N, Quadrado M, Simon M et al (2014) A restorer-of-fertility like pentatricopeptide repeat gene directs ribonucleolytic processing within the coding sequence of rps3-rpl16 and orf240a mitochondrial transcripts in *Arabidipsis thaliana*. Plant J 78:134–145

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

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

Brown GG, Formanová N, Jin H et al (2003) The radish *Rfo* restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J 35:262–272

Chase CD (2007) Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends Genet 23:81–90

Chen L, Liu Y (2014) Male sterility and fertility restoration in crops. Annu Rev Plant Biol 65:579–606

Cheng S, Gutmann B, Zhong X et al (2016) Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. Plant J 85:532–547

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

Desloire S, Gherbi H, Laloui W et al (2003) Identification of the fertility restoration locus, *Rfo*, in radish, as a member of the pentatricopeptide-repeat protein family. EMBO Rep 4:588–594

Dill CL, Wise RP, Schnable PS (1997) *Rf8* and *Rf** mediate unique T-urf13-transcript accumulation, revealing a conserved motif associated with RNA processing and restoration of pollen fertility in T-cytoplasm maize. Genetics 147:1367–1379

Fu TD (1981) Production and research of rapeseed in the People's Republic of China. Eucarpia Cruciferae Newsl 6:6–7

Fu T, Yang X, Yang G (1989) Development and studies on polima cytoplasmic male sterile 'three lines' in *Brassica napus* L. J Huazhong Agric Univ 8:201–207

Fujii S, Toriyam K (2008) Genome barriers between nuclei and mitochondria exemplified by cytoplasmic male sterility. Plant Cell Physiol 49:1484–1494

Fujii S, Toriyam K (2009) Suppressed expression of retrograde-regulated male sterility restores pollen fertility in cytoplasmic male sterile rice plant. Proc Natl Acad Sci USA 106:9513–9518

Fujii S, Bond CS, Small ID (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

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

Hu J, Wang K, Huang W et al (2012) The rice pentatricopeptide repeat protein RF5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. Plant Cell 24:109–122

Itabashi E, Iwato N, Fujii S et al (2011) The fertility restorer gene, *Rf2*, for Lead Rice-type cytoplasmic male sterility of rice encodes a mitochondrial glycine-rich protein. Plant J 65:359–367

Jean M, Brown GG, Landry BS (1997) Genetic mapping of nuclear fertility in canola (*Brassica napus* L.) using DNA markers. Ther Appl Genet 95:321–328

Jordan DR, Mace ES, Henzell RG et al (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

Ke G, Song Y (1989) Breeding of alloplasmic male sterile line in Chinese cabbage (*Brassica campestris* L. ssp. pekinensis (Lour) olsson) and its application. Shaanxi J Agric Sci 3:9–11

Ke G, Zhang L (1993) Study on the relartionship of restoration and maintenance of alloplasmic male sterile in Chinese cabbage. Acta Agricuturae Boreali-ccidentalis Sinica 2:15–20

Ke G, Zhao Z, Song Y et al (1992) Breeding of alloplasmic male sterile line CMS3411-7 in Chinese cabbage (*Brassica campestris* L. ssp. pekinensis (Lour) olsson) and its application. Acta Horticulturae Sinica 19:333–340

Kitazaki K, Arakawa T, Matsunaga M et al (2015) Post-translational mechanisms are associated with fertility restoration of cytoplasmic male sterility in sugar beet (*Beta vulgaris*). Plant J 83:290–299

Klein RR, Klein PE, Mullet JE et al (2005) Fertility restorer locus Rf1 of sorghum (Sorghum bicolor L.) encodes a pentatricopeptide repeat protein not present in the collinear region of rice chromosome 12. Theor Appl Genet 111:994–1012

Koizuka N, Imai R, Fujimoto H et al (2003) Genetic characterization of a pentatricopeptide repeat protein gene, *orf687*, that restores fertility in the cytoplasmic male-sterile Kosena radish. Plant J 34:407–415

Komori T, Ohta S, Murai N et al (2004) Map-based cloning of a fertility restorer gene, *Rf-1*, in rice (*Oryza sativa* L.). Plant J 37:312–325

Li Q, Wan JM (2005) SSRHunter: development of a local searching software for SSR sites. Hereditas 27:808–810

Liu XQ, Xu X, Tan YP et al (2004) Inheritance and molecular mapping of two fertility-restoring loci for Honglian gametophytic cytoplasmic male sterility in rice (*Oryza sativa* L.) Mol Gen. Genomics 271:586–594

Liu Z, Liu P, Long F et al (2012) Fine mapping and candidate gene analysis of the nuclear restorer gene *Rfp* for *pol* CMS in rapeseed (*Brassica napus* L.). Theor Appl Genet 125:773–779

Liu Z, Yang Z, Wang X et al (2016) A mitochondria-targeted PPR protein restores *pol* cytoplasmic male sterility by reducing o*rf224* transcript levels in oilseed rape. Mol Plant 9:1082–1084

Luo D, Xu H, Liu Z et al (2013) A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. Nat Genet 45:573–577

Nagaharu U (1935) Genome analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Jpn J Bot 7:389–452

Porebski S, Bailey LG, Baum BR (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Report 15:8–15

Tang H, Luo D, Zhou D et al (2014) The rice restorer *Rf4* for wild-abortive cytoplasmic male sterility encodes a mitochondrial-localized PPR protein that functions in reduction of *WA352* transcripts. Mol Plant 7:1497–1500

Uyttewaal M, Arnal N, Quadrado M et al (2008) Characterization of *Raphanus sativus* pentatricopeptide repeat proteins encoded by the fertility restorer locus for Ogura cytoplasmic male sterility. Plant Cell 20:3331–3345

Wang Z, Zou Y, Li X et al (2006) Cytoplasmic male sterility of rice with Boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. Plant Cell 18:676–687

Wang X, Wang H, Wang J et al (2011) The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet 43:1035–1039

Wang ZW, Wang C, Gao L et al (2013) Heterozygous alleles restore male fertility to cytoplasmic male-sterile radish (*Raphanus sativus* L.): a case of overdominance. J Exp Bot 64:2041–2048

Wise RP, Gobelman-Werner K, Pei D et al (1999) Mitochondrial transcript processing and restoration of male fertility in T-cytoplasm maize. J Hered 90:380–385

Xu X, Sun X, Zhang Y et al (2014) Identification of AFLP and SSR markers linked with the male fertility restorer gene of CMS 06J45 in heading Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). Plant Breed 133:615–619

Yang G, Fu T (1991) A preliminary study on the restoring-maintaining relationship in rapeseed (*Brassica napus* and *Brassica campestris*). Acta Agronomica Sinica 17:151–156

Yin P, Li Q, Yan C et al (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. Nature 504:168–171

Zhang L, Hao D (2001) Investigation on the sterility changeover of male sterility line CMS7311 in heading Chinese cabbage. Acta Botanica Sinica 43:1123–1128

Zhang L, Ke G (1994) The genetic law and restoration of alloplasmic male sterile in Chinese cabbage (*Brassica campestris* L. ssp. pekinensis (Lour) olsson). Acta Agricuturae Boreali-ccidentalis Sinica 3:45–50