Theor Appl Genet (2012) 125:467–477
DOI 10.1007/s00122-012-1846-2

ORIGINAL PAPER

Construction of a high-resolution linkage map of *Rfd1*, a restorer-of-fertility locus for cytoplasmic male sterility conferred by DCGMS cytoplasm in radish (*Raphanus sativus* L.) using synteny between radish and *Arabidopsis* genomes

Youngcho Cho · Young-Pyo Lee · Beom-Seok Park · Tae-Ho Han · Sunggil Kim

Received: 29 December 2011 / Accepted: 5 March 2012 / Published online: 21 March 2012

Abstract Cytoplasmic male sterility caused by Dongbu cytoplasmic and genic male-sterility (DCGMS) cytoplasm and its nuclear restorer-of-fertility locus (*Rfd1*) with a linked molecular marker (A137) have been reported in radish (*Raphanus sativus* L.). To construct a linkage map of the *Rfd1* locus, linked amplified fragment length polymorphism (AFLP) markers were screened using bulked segregant analysis. A 220-bp linked AFLP fragment sequence from radish showed homology with an *Arabidopsis* coding sequence. Using this *Arabidopsis* gene sequence, a simple PCR marker (A220) was developed. The A137 and A220 markers flanked the *Rfd1* locus. Two homologous *Arabidopsis* genes with both marker sequences were positioned on *Arabidopsis* chromosome-3 with an interval of 2.4 Mb. To integrate the *Rfd1* locus into a previously reported expressed sequence tag (EST)-simple sequence repeat (SSR) linkage map, the radish EST sequences located in three syntenic blocks within the 2.4-Mb interval were used to develop single nucleotide polymorphism (SNP) markers for tagging each block. The SNP marker in linkage group-2 co-segregated with male fertility in an F_2 population. Using radish ESTs positioned in linkage group-2, five intron length polymorphism (ILP) markers and one cleaved amplified polymorphic sequence (CAPS) marker were developed and used to construct a linkage map of the *Rfd1* locus. Two closely linked markers delimited the *Rfd1* locus within a 985-kb interval of *Arabidopsis* chromosome-3. Synteny between the radish and *Arabidopsis* genomes in the 985-kb interval were used to develop three ILP and three CAPS markers. Two ILP markers further delimited the *Rfd1* locus to a 220-kb interval of *Arabidopsis* chromosome-3.

Communicated by I. Paran.

Y. Cho · T.-H. Han · S. Kim (✉)
Department of Plant Biotechnology,
Biotechnology Research Institute,
Chonnam National University,
Gwangju 500-757, Korea
e-mail: dronion@jnu.ac.kr

Y.-P. Lee
Biotech Research Center,
Dongbu Advanced Research Institute,
Dongbu Hannong Co., Ltd.,
Daejeon 305-708, Korea

B.-S. Park
The Agricultural Genome Center,
National Academy of Agricultural Science,
Suwon 441-707, Korea

Introduction

Cytoplasmic male-sterility (CMS), the inability to produce viable pollen grains, has been observed in more than 150 plant species, including many economically important crops (Hanson 1991). CMS has been utilized as an economical emasculation method in F_1 hybrid seed production and has been a model system for studying communication between nuclear and mitochondrial genomes (Budar et al. 2003; Hanson and Bentolila 2004; Knoop 2004; Kubo and Newton 2008). All cases of CMS reported to date are caused by aberrant mitochondrial genes, often chimeric genes consisting of partial sequences of known mitochondrial genes and unknown sequences (Budar et al. 2003; Hanson and Bentolila 2004). Dynamic mitochondrial DNA rearrangement may cause chimeric genes. Such rearrangement is assumed to be caused by recombination that is mediated by repeat sequences less than 1 kb (Small et al. 1989; Kmiec et al. 2006).

Configuration of plant mitochondrial genomes is complex, and multipartite structures of subgenomes exist in a

single cell (Backert et al. 1997; Oldenburg and Bendich 2001; Allen et al. 2007). Even the stoichiometry of subgenomes varies at an intraspecific level (Palmer 1988; Allen et al. 2007). Specific stoichiometry of subgenomes is maintained through generations, but can be changed by triggers such as tissue culture and nuclear genes (Sakai and Imamura 1993; Bellaoui et al. 1998; Janska et al. 1998; Kim et al. 2007). Nuclear genes controlling stoichiometry of subgenomes such as *Msh1* (Abdelnoor et al. 2006), *RecA* (Shedge et al. 2007), and *OSB1* (Zaegel et al. 2006) have been reported.

Fertility of CMS plants is sometimes restored by nuclear restorer-of-fertility (Rf) genes. Except for the maize *Rf2* gene encoding aldehyde dehydrogenase (Cui et al. 1996), the rice *Rf17* encoding protein with unknown function (Fujii and Toriyama 2009), and the rice *Rf2* coding for a glycine-rich protein (Itabashi et al. 2011), all Rf genes isolated to date code for pentatricopeptide repeat (PPR) proteins (Bentolila et al. 2002; Brown et al. 2003; Desloire et al. 2003; Koizuka et al. 2003; Komori et al. 2004; Klein et al. 2005). Interactions between CMS-inducing genes and genes encoding PPR proteins has recently been investigated to elucidate the role of PPR proteins in the suppression of CMS-inducing mitochondrial genes (Wang et al. 2006; Gillman et al. 2007; Fujii et al. 2011).

Since CMS in radish (*Raphanus sativus* L.) was first reported by Ogura (1968), the genes have been introduced into radish cultivars used in commercial seed production. Furthermore, Ogura CMS was introduced into *Brassica* species by backcrossing (Bannerot et al. 1974; Dickson 1985) or protoplast fusion (Pelletier et al. 1983; Menczel et al. 1987; Jarl et al. 1989) to implement stable F_1 hybrid seed production in *Brassica*. In addition, the mitochondrial gene *orf138* responsible for Ogura CMS induction (Bonhomme et al. 1991; Grelon et al. 1994) and the nuclear Rf gene, *Rfo*, that restorers Ogura CMS, were isolated. The latter was isolated by map-based cloning and revealed to encode a PPR protein (Brown et al. 2003; Desloire et al. 2003; Koizuka et al. 2003).

In the past decade, a new CMS, designated Dongbu cytoplasmic and genic male-sterility (DCGMS), was discovered in radish germplasm introduced from Western Europe (Lee et al. 2008). Cytological examination of pollen development of DCGMS plants showed that microspore abortion occurred at later stages than in Ogura CMS plants. DCGMS plants are particularly useful in radish breeding programs since the frequency of DCGMS maintainer lines in radish breeding germplasm in Korea was much higher than Ogura CMS maintainer lines (Lee et al. 2008). Molecular markers for distinguishing the DCGMS cytotype from three other radish cytotypes were developed based on the mitochondrial (Lee et al. 2009) and chloroplast variation (Kim et al. 2009).

Breeding lines containing an Rf gene for CMS that was conferred by the DCGMS cytotype were also identified. Fertility restoration was conditioned by a single locus, *Rfd1*, and a linked molecular marker was developed (Kim et al. 2010). However, construction of a high-resolution linkage map was not developed due to a lack of radish genomic information. Since the first linkage map for the radish genome was constructed using restriction fragment length polymorphism (RFLP) markers by Bett and Lydiate (2003), a couple of rough linkage maps were used in quantitative trait locus (QTL) analysis of root shape and red pigmentation (Tsuro et al. 2008), resistance to beet cyst nematodes (Budahn et al. 2009), and resistance to clubroot disease (Kamei et al. 2010). The most saturated linkage map for radish, which was mainly constructed using simple sequence repeat (SSR) markers derived from expressed sequence tags (ESTs), was recently reported (Shirasawa et al. 2011). Since most SSR markers were designed based on radish coding genes, syntenic blocks of the radish genome to the *Arabidopsis* genome were identified and showed that radish genomes have been duplicated or triplicated since divergence from a common ancestor (Shirasawa et al. 2011).

In this study, a high-resolution linkage map of the *Rfd1* locus was constructed using the EST sequences anchored in the EST-SSR linkage map and synteny between radish and *Arabidopsis* genomes.

Materials and methods

Plant materials

An F_2 population was produced from the F_1 hybrid originating from the cross between the male-sterile line (MS19) containing DCGS cytoplasm and a male-fertile breeding line (R121). The F_1 hybrid was produced in a previous study (Lee et al. 2008). F_2 seeds were germinated in a greenhouse in 128-cell plug trays measuring 54 × 28 × 8 cm^3. The F_2 plants were vernalized through a winter season and transplanted in the field. Multiple visual examinations were used to evaluate the male-fertility phenotypes of 401 F_2 plants. Total genomic DNA was extracted from fresh leaf tissue from seedlings at the three-leaf stage using a cetyl trimethylammonium bromide (CTAB) method (Doyle and Doyle 1987).

Amplified fragment length polymorphism (AFLP) and bulked segregant analysis (BSA)

The BSA approach (Michelmore et al. 1991) was used to identify additional molecular markers linked to the *Rfd1*

locus. Ten F_2 individuals consisting of male-sterile and male-fertile bulks were selected based on both the male-fertility phenotype and genotypes of the A137 marker, which has been reported to be linked to the *Rfd1* locus at a distance of 5 cM (Kim et al. 2010). The male-fertile F_2 plants that were heterozygous for the A137 marker were excluded when making DNA bulks.

AFLP analysis was performed as described by Vos et al. (1995) with slight modifications. Total DNA was digested with *Eco*RI and *Mse*I restriction enzymes. The adaptor-ligated DNA was pre-amplified using primers with a single selective nucleotide (*Eco*RI + N/*Mse*I + N). The primer combinations of *Eco*RI + NNN and *Mse*I + NNN were used for selective amplifications. The amplified products were separated on 6 % polyacrylamide gels in 1 × TBE buffer. Gels were stained with a staining reagent (SILVER SEQUENCE™, Promega, Madison, WI) according to the manufacturer's instructions. Sequencing of the AFLP fragment linked to the *Rfd1* locus was performed as described by Kim et al. (2010).

Amplification by polymerase chain reaction (PCR) for molecular-marker development and sequencing of PCR products

PCR was performed in 10-μL reaction mixtures containing 0.05 μg template, 1 μL 10 × PCR buffer, 0.2 μL forward primer (10 μM), 0.2 μL reverse primer (10 μM), 0.2 μL dNTPs (10 mM each), and 0.1 μL polymerase mix (Advantage 2 Polymerase Mix, Clontech, Palo Alto, CA). For analyses of a large number of samples, 0.25 U *Taq* polymerase (SP-Taq polymerase, Cosmo Genotech, Seoul, Korea) were used with the same conditions. PCR amplification was performed with an initial denaturation at 95 °C for 4 min, 10 cycles of 95 °C for 30 s, 67 °C (0.8 °C decrements in each cycle) for 30 s, and 72 °C for 2 min, 30 cycles of 95 °C for 30 s, 59 °C for 30 s, and 72 °C for 2 min, and a final 7-min extension at 72 °C. PCR products were visualized on 1.5 % agarose gels or 9 % polyacrylamide gels after ethidium bromide staining. For genotyping cleaved amplified polymorphic sequence (CAPS) markers, PCR products were digested with their respective restriction enzymes for 3 h at 37 °C. The primer sequences for molecular markers developed in this study are presented in Table 1.

PCR products were purified using a QIAquick PCR Purification kit (QIAGEN, Valencia, CA) and sequenced directly. Sequencing reactions were carried out using Big Dye (Applied Biosystems, Foster City, CA) according to the manufacturer's protocol and the sequences were obtained using an ABI PRISM 3730XL Analyzer (Applied Biosystems).

Pyrosequencing and assembly of radish cDNA

Total RNA was extracted using an RNA extraction kit and following the manufacturer's instructions (RNeasy Plant Mini Kit, QIAGEN). Whole plants at the four-leaf stage grown in aseptic conditions were used for this extraction. A male-fertile breeding line (701DB) was used for RNA extraction. The cDNAs produced were fragmented by nebulization into 500–800 bp fragments. Pyrosequencing was carried out using a Genome Sequencer FLX System (454 Life Sciences, Branford, CT). The read sequences were assembled into contigs using Newbler Assembler software (454 Life Sciences). A total of 28,479 contigs from radish cDNAs were formed. Transcriptome sequence data were deposited in GenBank with the accession numbers from JT784548 to JT813026.

Linkage analysis

The linkage relationship between the molecular markers and the *Rfd1* locus was calculated using JoinMap version 4.0 (Stam 1993). Map distances (cM) were calculated from recombinant frequencies using the Kosambi function (Kosambi 1944).

Results

Development of a molecular marker linked to the *Rfd1* locus

A novel locus, *Rfd1,* that restores fertility to male-sterile DCGMS radish plants and a linked marker, A137, were reported in our previous study (Kim et al. 2010). To develop additional linked markers for the construction of a linkage map of the *Rfd1* locus, the BSA approach was used in combination with AFLP analysis. A total of 64 primer combinations were used in AFLP analysis. A polymorphic AFLP band of approximately 220 bp was detected between two bulks, and the AFLP marker co-segregated with male-fertility phenotypes in a F_2 population (Fig. 1a).

To convert the AFLP marker into a simple PCR marker, sequences of the AFLP fragment were obtained after cloning the AFLP product. The AFLP fragment sequence showed 81 % homology to the 5′ coding sequences of an *Arabidopsis* gene encoding an ethylene-responsive transcription factor RAP2-2 (Fig. 1b). A series of four reverse primers were designed based on the exon 2 sequence of the *Arabidopsis* homologous gene. PCR amplification was performed using a forward primer designed from the radish AFLP sequence and four serial reverse primers. Single clear PCR products were obtained from the male-fertile F_2 bulks with the first two primer combinations. However, no

Table 1 Molecular markers developed in this study

Marker	Primer sequences (5′–3′) Forward/reverse	Type of marker	Indels/restriction enzymes[a]	Size of PCR products (bp)[b]
A220	TTTGAGTTAGACAAGAAGACAACCA/ CATTTTCCCCAAGGACGCTGCCTAATC	Dominant simple PCR		1,021 (JQ425087)
ILP1501	CTTTGCCTAAGGGGTTCGAT/ CAGTCAGTGGCTGAGAATCAA	ILP	26 bp	289 (JQ425088)/315 (JQ425089)
ILP2859	TTGAGATCCGCCACTCTCTT/ CTTCCTCAGCTGCATCCTCT	ILP	16 bp	251 (JQ425090)/267 (JQ425091)
ILP2760	GGCCTGTCATCTTCTTCGTC/ TTCTTTAACACGGTGACTGGTG	ILP	140 bp	360 (JQ425092)/220 (JQ425093)
ILP0662	TGGGAGTTTTCGAGATGGAG/ GCAATGGCCTGCATAAAAA	ILP	22 bp	191 (JQ425094)/213 (JQ425095)
ILP2138	CCATGAGGAAGAAGCTCGAC/ CATCCTCGTCAGCTTGCATA	ILP	49 bp	397 (JQ425096)/348 (JQ425097)
CAPS2505	ATGATGGTCATGGTGGTCCT/ GGTCTAGACACCCGCCTAAA	CAPS	*Bfa*I	200 (JQ425098)/200 (JQ425099)
CAPS3521	GCCTTGGCTATCTCTGTTGC/ AAACAAAAATAATTGGATCAGATGG	CAPS	*Taq*I	497 (JQ425100)/497 (JQ425101)
CAPS3100	ACAAGCTGAGGATTGTCACG/ TCGGTTTCTGAACTCAACCA	CAPS	*Eco*RV	810 (JQ425102)/808 (JQ425103)
ILP3290	TACGCGATGAAGATCATGGA/ GAACCGAGCTTCATCCTCTG	ILP	63 bp	590 (JQ425104)/527 (JQ425105)
ILP3510	AAGATCGCTCTTCATCTTCAGG/ TCTGAAAGGCTCTGATGAAACA	ILP	11 bp	115 (JQ425106)/126 (JQ425107)
ILP3585	AGAGAGCGGTTGCATGAGAT/ TCTTGAAGCTCTTCGCCATC	ILP	6 bp	88 (JQ425108)/94 (JQ425109)
CAPS3700	TCGAAGCTCCATGGGAATAC/ ATCTGCGGACGTAGGAAATG	CAPS	*Ssp*I	653 (JQ425110)/653 (JQ425111)

[a] Lengths of indels on the intron sequences used in ILP marker development and names of restriction enzymes used in CAPS marker development are indicated

[b] Sizes of PCR products of dominant allele-linked sequence (GenBank accession)/recessive allele-linked sequence (GenBank accession) are shown

positive PCR products were observed in any male-sterile F_2 bulks (Fig. 1c). After sequencing the positive PCR products, five additional forward and reverse primers were designed based on radish sequences. However, no PCR products were amplified in any male-sterile F_2 individual plants (data not shown). This indicates that the gene including the AFLP fragment might be deleted in the male-sterile parental line used in this study. Therefore, the primer combination of Rs-F1 and At-R1 was used to develop a dominant simple PCR marker designated 'A220.'

Integration of the *Rfd1* locus into the radish EST-SSR linkage map

Genotyping of the A137 and A220 markers in the F_2 population showed that the A220 marker was positioned on the opposite side of *Rfd1* compared to the A137 marker (data not shown). In addition, the two *Arabidopsis* genes homologous to the radish genes containing the A137 and A220 marker sequences were both positioned on *Arabidopsis* chromosome-3. The A137 and A220 homologs were located on the 2,295 and 4,738 K positions of *Arabidopsis* chromosome-3, respectively (Fig. 2).

To integrate the *Rfd1* locus into the recently reported radish linkage map (Shirasawa et al. 2011), the synteny between the *Arabidopsis* and radish genomes was utilized. In addition to constructing a radish-linkage map based mainly on EST-SSRs, Shirasawa et al. (2011) identified syntenic blocks between *Arabidopsis* and radish genomes. Three regions of the radish-linkage map had syntenic relationships with *Arabidopsis* chromosome-3 containing the A137 and A220 homologs (Fig. 2). To determine the region containing the *Rfd1* locus, three EST sequences from which three SSR markers (RSS2505, RSS1153, and RSS1674) had been developed were used to detect single nucleotide polymorphisms (SNP) between male-sterile and male-fertile F_2 bulks.

Primers were designed based on the putative exons flanking one or two introns. PCR amplification was carried out using male-sterile and male-fertile F_2 bulk DNAs. The

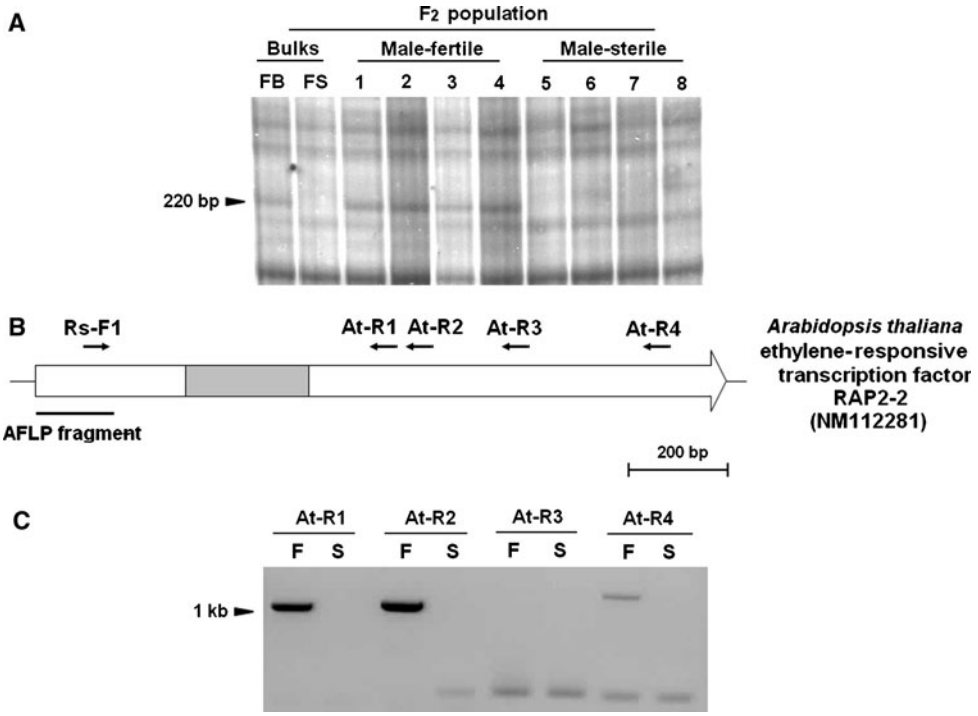

Fig. 1 Development of the molecular marker A220 linked to the *Rfd1* locus. **a** The AFLP profile was amplified with a primer combination of *E*-ATC and *M*-CAC. The polymorphic band is indicated with an *arrowhead*. *FB* Male-fertile F_2 bulk, *FS* male-sterile F_2 bulk. *1–4* Male-fertile F_2 individuals, *5–8* male-sterile F_2 individuals. **b** Organization of the *Arabidopsis* genes showing homology with the AFLP fragment sequence linked to the *Rfd1* locus. The position of AFLP fragment is indicated with a *horizontal line*. *Arrow-shaped boxes* indicate the 5′–3′ direction. The *empty* and *filled boxes* indicate exons and introns, respectively. *Horizontal arrows* indicate primer-binding sites. **c** PCR products from male-fertile and male-sterile F_2 bulks amplified using the primer combinations in **b**. A common forward primer, Rs-F1, was used in all four reactions. Only reverse primers are shown. *F* Male-fertile F_2 bulks, *S* male-sterile F_2 bulks

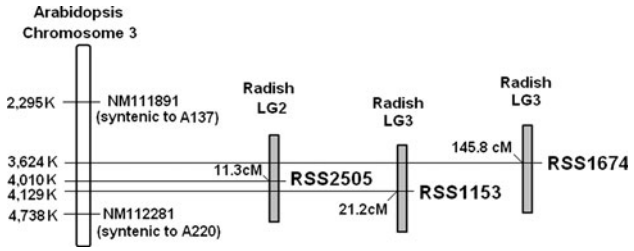

Fig. 2 Syntenic relationship between *Arabidopsis* chromosome-3 and radish SSR markers located at three different positions in the EST-SSR linkage map (Shirasawa et al. 2011). The syntenic regions are connected with horizontal lines. The coordinates of the syntenic regions of *Arabidopsis* chromosome-3 are shown in kilobase pairs (Kb) and positions of radish SSR markers in the radish EST-SSR linkage map are shown in cM. A137 and A220 are molecular markers flanking the *Rfd1* locus

sequence of the RSS1153-derived PCR product was monomorphic between the two bulks and the RSS1674-derived PCR product sequences showed more than six SNPs and indels. However, all SNP nucleotide peaks in the chromatograph of capillary sequencing overlapped in both bulks, indicating that this region did not co-segregate with the *Rfd1* alleles. Meanwhile, the RSS2505-derived sequences showed six SNPs and no SNP nucleotide peaks overlapped between the two bulks. Thus, the region containing the RSS2505 marker co-segregated with the *Rfd1* alleles and the *Rfd1* locus is positioned in the upper part of the linkage group-2 of the radish EST-SSR linkage map.

Development of molecular markers linked to the *Rfd1* locus using EST information anchored in the EST-SSR linkage map

Because the *Rfd1* locus was located on the linkage group-2 of the radish EST-SSR linkage map, eight SSR markers linked to the RSS2505 marker were tested for the presence of polymorphism in our population (Fig. 3). Two SSR markers (RSS0818 and RSS2497) had polymorphisms between the male-sterile and male-fertile F_2 bulks. However, six other SSR markers were classified as monomorphic or difficult to score marker genotypes due to multiple non-specific band patterns. Therefore, EST sequences from which SSR markers originated were used to develop PCR markers based on intron sequence variations. The EST sequences were obtained from Shirasawa et al. (2011).

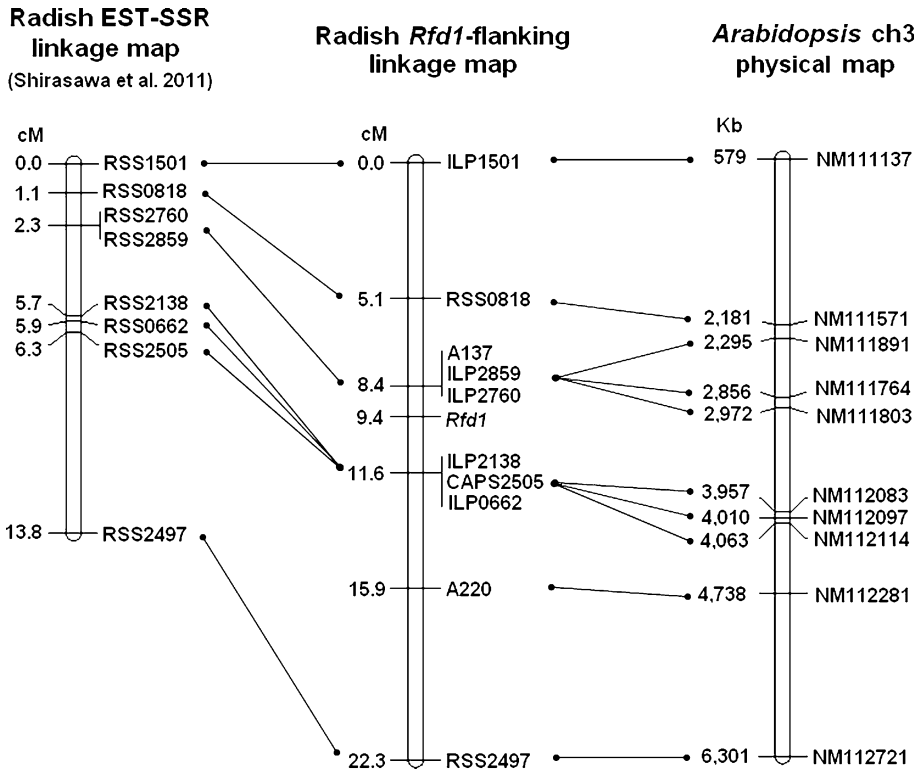

Fig. 3 Linkage maps of the molecular markers flanking the *Rfd1* locus. The linkage map in the *center* was constructed using the F_2 population. The map on the *left* is the radish EST-SSR linkage map constructed by Shirasawa et al. (2011). The physical map of *Arabidopsis* chromosome-3 is shown on the *right*. The coordinates of physical distances (Kb) are shown to the *left* of the map and syntenic regions are connected with *horizontal lines*

Primers were designed from exon sequences of which positions were predicted based on *Arabidopsis* orthologous genes for detecting polymorphisms within the intron sequences. Five intron length polymorphisms (ILPs) and one CAPS marker were developed based on polymorphic intron sequences (Table 1; Fig. 3). The F_2 population consisting of 95 individuals was analyzed using A137, two SSRs, and seven developed markers together with male-fertility phenotypes to construct a linkage map of the *Rfd1* locus (Fig. 3). The order of EST-based markers was consistent with that of the EST-SSR linkage map, though the genetic distances varied due to the differences in mapping populations. The constructed linkage map delimited the *Rfd1* locus between the ILP2760 and ILP2138 markers. In addition, the marker order and the positions of homologous *Arabidopsis* genes on the physical map were also collinear (Fig. 3). The recombinant-based interval between the ILP2760 and ILP2138 markers corresponded to a physical interval of 985 kb on *Arabidopsis* chromosome-3.

Construction of a high-resolution linkage map of the *Rfd1* locus using synteny between the *Arabidopsis* and radish genomes

Identification of a syntenic relationship between the radish *Rfd1*-flanking regions and *Arabidopsis* chromosome-3 permitted the use of sequence information of *Arabidopsis* genes within the 985-kb interval for developing additional linked markers and for narrowing the 985-kb interval. Radish EST sequences showing high homology with *Arabidopsis* genes were identified from 28,479 cDNA contigs produced in this study. Multiple EST sequences were frequently retrieved after a local BLAST search due to the duplicated or triplicated nature of radish genomes compared to the *Arabidopsis* genome.

To establish the criteria for selecting the most plausible ESTs linked to the *Rfd1* locus, the relationship between radish ESTs linked to the *Rfd1* and *Arabidopsis* genes was analyzed. The positive ESTs generally showed high bit-scores and high coverage rates for homologous sequences between radish and *Arabidopsis* ESTs (Table 2). However, it was difficult to select the most plausible EST when two or three ESTs with similar high bit-scores were retrieved. *Arabidopsis* genes that retrieved only one homologous EST were excluded since there was a high possibility that undetected homologous radish genes in our EST library or non-expressed homologous pseudogenes would interfere with reliable amplification of single PCR products. Therefore, only *Arabidopsis* genes showing BLAST search outputs of two or three ESTs with high bit-scores and low *E* values were selected. After aligning two or three homologous radish EST sequences, primers for unique sequences of each EST were designed from the predicted exons flanking two or three introns.

Three ILPs and three CAPS markers were developed using synteny between the radish and *Arabidopsis* genomes (Table 1; Fig. 4). Additional F_2 individuals consisting of 401 plants were analyzed with ILP2760 and ILP2138 to

Table 2 Comparison of sequence homology between radish EST sequences and *Arabidopsis* genes in the 2.4-Mb interval of chromosome-3 that are syntenic to the radish *Rfd1*-flanking regions

Arabidopsis genes[a]	Homologous radish EST sequences[b]			Radish markers
NM111571/AT3G06910 (1,793 bp)[c]	*R16358*	R23912		RSS0818
	88 %, 0.47[d]	93 %, 0.15		
NM111891/AT3G10550 (2,523 bp)	*R21914*	R02031	R24948	A137
	90 %, 0.95	85 %, 0.68	86 %, 0.53	
NM111764/AT3G09300 (1,377 bp)	R20520	R13538	*R18581*	ILP2859
	91 %, 0.86	91 %, 0.92	*91 %, 0.89*	
NM111803/AT3G09690 (2,051 bp)	*R25012*	R23641		ILP2760
	85 %, 0.75	81 %, 0.44		
NM112083/AT3G12480 (882 bp)	R17193	*R26114*	R09331	ILP2138
	90 %, 0.74	*90 %, 0.58*	87 %, 0.26	
NM112097/AT3G12620 (1,158 bp)	*R12016*	R08588	R11201	CAPS2505
	88 %, 0.87	87 %, 0.93	81 %, 0.17	
NM112114/AT3G12780 (1,446 bp)	*R26501*	R19076	R22981	ILP0662
	91 %, 0.80	89 %, 0.84	90 %, 0.87	
NM180251/AT3G14230 (1,637)	*R03968*	R01422	R13836	A220
	87 %, 0.46	87 %, 0.58	90 %, 0.29	

The EST sequences linked to the *Rfd1* locus are shown in italics

[a] GenBank accession numbers/locus tags are shown in order of position on *Arabidopsis* chromosome-3

[b] EST sequences are shown in descending order of alignment bit-scores. If bit-scores were identical, homology levels were used for sorting

[c] Length of full-length cDNA sequences

[d] Percent identity of radish ESTs with *Arabidopsis* gene sequences and the proportion of homologous sequences within the total length of EST sequences

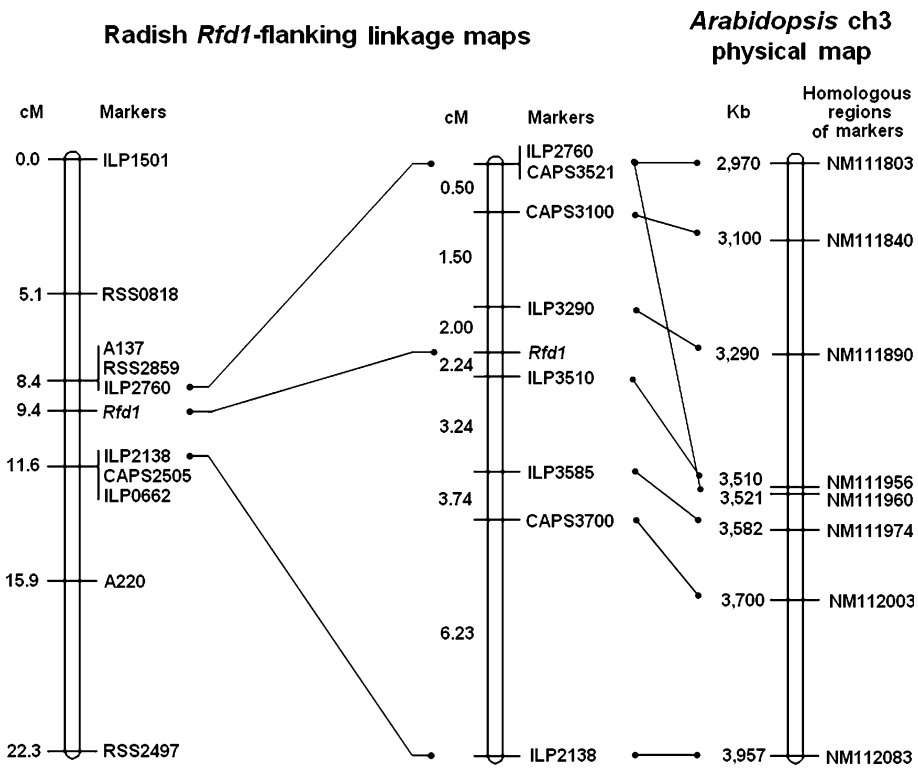

Fig. 4 A high-resolution linkage map of the *Rfd1* locus. A high-resolution linkage map of the *Rfd1* locus between the ILP2760 and ILP2138 markers is shown in the *center* and the linkage map in Fig. 3 is shown at the *left* for reference. The physical map of *Arabidopsis* chromosome-3 is positioned on the *right* and the syntenic regions are connected with *horizontal lines*

identify recombinants between the two markers. The identified 25 recombinants were analyzed with six closely linked markers to construct a light-resolution linkage map (Fig. 4). The ILP3290 and ILP3510 markers flanking the *Rfd1* locus had two and one recombinants between the *Rfd1* locus and the markers, respectively. Five markers were collinear to their corresponding *Arabidopsis* genes, but the position of the CAPS3521 marker was not collinear, indicating an occurrence of gene translocation. In conclusion, the markers ILP3290 and ILP3510 delimited the *Rfd1* locus within a 220-kb interval on the *Arabidopsis* chromosome-3. This information will be useful for isolating the *Rfd1* gene.

Discussion

Unlike *Arabidopsis thaliana* and some *Brassica* species, such as *B. rapa* and *B. napus*, genomic information for radish is limited. This may be partially due to its relatively low economic importance worldwide, even though it is a major vegetable crop in eastern Asian countries. However, genomic information on *Arabidopsis* and *Brassica* species can be utilized in genomic and genetic studies of radish because radish is a close relative of both species. In this study, a linkage map of the *Rfd1*-flanking regions was successfully constructed using synteny between the radish and *Arabidopsis* genomes.

The Brassicaceae family is estimated to consist of 338 genera and 3,700 species (Bailey et al. 2006). Previous phylogenetic studies of the Brassicaceae family showed that radish was a relative of both *Arabidopsis* and *Brassica* species (Yang et al. 1999; Kim et al. 2009). In particular, radish was reported to be more closely related to *Brassica rapa* than other *Brassica* species when variations within the chloroplast genomes were used to construct phylogenetic trees (Warwick and Black 1997; Yang et al. 2002; Kim et al. 2009). However, when the phylogeny were analyzed using RFLP and random amplified polymorphic DNA (RAPD) techniques (Thormann et al. 1994) or nuclear internal transcribed spacer (ITS) sequences (Yang et al. 1999), radish was shown to be more closely related to *B. nigra* than *B. rapa*. For these reasons, Song et al. (1990) suggested that radish originated from interspecific hybridization between *B. rapa* and *B. nigra* in which *B. rapa* was the female parent. Furthermore, Yang et al. (2002), using noncoding chloroplast DNA sequences, estimated that a hybridization event occurred in approximately 60 % of the divergence events between the *rapaloleracea* lineage that includes *B. rapa* and the *nigra* lineage.

The hypothesis of the hybrid origin of radish suggests that the nuclear genome of *B. nigra* might be most closely related to that of radish, but genomic information on *B. nigra* is also scant. Using comparative mapping, Lagercrantz (1998) suggested that the *B. nigra* genome triplicated after divergence from the *Arabidopsis* and *Brassica* lineages. The recent release of the draft of whole-genome sequences of *B. rapa* also clearly show that the *B. rapa* genome has triplicated compared with the *Arabidopsis* genome (The *Brassica rapa* genome sequencing project consortium 2011). If radish originated from the hybridization between the *rapaloleracea* and *nigra* lineages, the radish genome may also show genome triplication. Indeed, Shirasawa et al. (2011) showed that the radish genome contained duplicated or triplicated syntenic blocks compared with the *Arabidopsis* genome. Comparison of the nuclear genome organization of radish, *B. nigra*, and *B. rapa* will be an interesting issue to determine the interspecific hybrid origin of radish.

The collinearity of syntenic blocks between the *Arabidopsis* and radish genomes was useful in developing molecular markers. However, the duplicated or triplicated nature of the radish genome prevented clear PCR amplification of *Arabidopsis* orthologs in radish when intron polymorphism-based markers, such as ILP markers were developed in this study. The failure to amplify some single PCR products may be due to the high level of homology among radish paralogous genes. The homology level of radish paralogs was greater than 85 % in most cases, though the homology level varied between specific genes (Table 2). The high level of homology among radish paralogs indicates that the triplication event occurred at a relatively recent evolutionary time in the *Brassica* lineage. Indeed, the triplication event in *B. rapa* was estimated to have occurred 5–9 million years ago (The *Brassica rapa* genome sequencing project consortium 2011). Due to the high homology among radish paralogs, for successful PCR amplification, *Arabidopsis* genes within the syntenic block needed to be selected when two or three radish homologous cDNA sequences were retrieved for an *Arabidopsis* gene. In those cases, primers could be designed based on dissimilar sequences among radish paralogs as determined by the sequence alignment. Significantly dissimilar sequences were generally located at the 3′ ends of EST sequences.

However, the local BLAST results for many *Arabidopsis* genes positioned within the 985-kb interval retrieved either no hit or only one homolog. This may be due to the low coverage of the radish transcriptome by the 28,530 cDNA contigs produced in this study. A total of 41,174 genes were identified from a draft whole-genome sequence of *B. rapa* (The *Brassica rapa* genome sequencing project consortium 2011). Since the genome sizes of *B. rapa* and radish were reported to be as similar as 0.54 and 0.59 pg/1C, respectively (Johnston et al. 2005), the radish genome is estimated to contain approximately 40,000 genes. Therefore, RNA extraction from more diverse plant tissues will be necessary to increase transcriptome coverage for

more efficient development of intron-based markers. Substantial gene loss of triplicated genome fragments may be another reason for the low homolog detection from the local BLAST search. In the case of *B. rapa*, the least fractionated sub-genome retains 70 % of the *Arabidopsis* orthologs, but the most fractionated sub-genome retains only 36 % of the *Arabidopsis* orthologs on the syntenic regions (The *Brassica rapa* genome sequencing project consortium 2011). Similarly, it is predicted that the radish genome has also experienced significant gene loss since genome triplication occurred.

We delimited the *Rfd1* locus within a 220-kb syntenic region of *Arabidopsis* chromosome-3 using PCR markers developed based on the synteny between radish and *Arabidopsis* genomes. One or two recombinants were identified between the closest markers and the *Rfd1* locus. To narrow the interval further, a larger segregating population should be produced and more closely linked markers should be developed. Isolation of the *Rfd1* gene will allow for the introduction of DCGMS male-sterility into other *Brassica* species in combination with the restorer-of-fertility gene, *Rfd1*, to establish new F_1 hybrid seed production systems in which stable and natural male-sterility are currently unavailable. Similar strategies were successfully applied in the case of Ogura CMS (Pelletier et al. 1983; Menczel et al. 1987; Jarl et al. 1989). In addition, isolation of the *Rfd1* gene will permit cloning of the mitochondrial gene responsible for male sterility conferred by the DCGMS cytotype and should prove useful in studying the interaction between a mitochondrial male sterility-inducing gene and its nuclear restorer.

Acknowledgments This work was supported by a grant from the Next-Generation BioGreen 21 Program (Plant Molecular Breeding Center No. PJ007992), Rural Development Administration and a grant (108169-05-01-SB010) from the Technology Development Program for Agriculture and Forestry, Ministry of Agriculture and Forestry, Republic of Korea. The authors thank Seung-Hee Kim, Ji-wha Hur, and Jeong-Ahn Yoo for their dedicated technical assistance.

References

Abdelnoor RV, Christensen AC, Mohammed S, Munoz-Castillo B, Moriyama H, Mackenzie SA (2006) Mitochondrial genome dynamics in plants and animals: convergent gene fusions of a *MutS* homologue. J Mol Evol 63:165–173

Allen JO, Fauron CM, Mink P, Roark L, Oddiraju S, Lin GN, Meyer L, Sun H, Kim K, Wang C, Du F, Xu D, Gibson M, Cifrese J, Clifton SW, Newton KJ (2007) Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. Genetics 177:1173–1192

Backert S, Neilsen BL, Börner T (1997) The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. Trend Plant Sci 2:477–483

Bailey CD, Koch MA, Mayer M, Mummenhoff K, O'Kane SL Jr, Warwick SI, Windham MD, Al-Shehbaz IA (2006) Toward a global phylogeny of the Brassicaceae. Mol Biol Evol 23:2142–2160

Bannerot H, Loulidard L, Cauderon MY, Tempe J (1974) Transfer of cytoplasmic male sterility from *Raphanus sativus* to *Brassica oleracea*. Proc Eucarpia Meeting Cruciferae. pp 52–54

Bellaoui M, Martin-Canadell A, Pelletier G, Budar F (1998) Low-copy-number molecules are produced by recombination, actively maintained and can be amplified in the mitochondrial genome of Brassicaceae: relationship to reversion of the male sterile phenotype in some cybrids. Mol Gen Genet 257:177–185

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

Bett KE, Lydiate DJ (2003) Genetic analysis and genome mapping in *Raphanus*. Genome 46:423–430

Bonhomme S, Budar F, Ferault M, Pelletier G (1991) A 2.5 kb *Nco* I fragment of Ogura radish mitochondrial DNA is correlated with cytoplasmic male-sterility in *Brassica* cybrids. Curr Genet 19:121–127

Brown GG, Formanova N, Jin H, Wargachuk R, Dendy C, Patil P, Laforest M, Zhang J, 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

Budahn H, Peterka H, Mousa MAA, Ding Y, Zhang S, Li J (2009) Molecular mapping in oil radish (*Raphanus sativus* L.) and QTL analysis of resistance against beet cyst nematode (*Heterodera schachtii*). Theor Appl Genet 118:775–782

Budar F, Touzet P, de Paepe R (2003) The nucleo-mitochondrial conflict in cytoplasmic male sterilities revised. Genetica 117:3–16

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, Marhadour S, Clouet V, Cattolico L, Falentin C, Giancola S, Renard M, Budar F, Small I, Caboche M, Delourme R, Bendahmane A (2003) Identification of the fertility restoration locus, *Rfo*, in radish, as a member of the pentatricopeptide-repeat protein family. EMBO Rep 4:588–594

Dickson MH (1985) Male sterile persistent white curd cauliflower NY 7642A and its maintainer NY 7642B. HortScience 20:957

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Fujii S, Toriyama K (2009) Suppressed expression of *RETROGRADE-REGULATED MALE STERILITY* restores pollen fertility in cytoplasmic male sterile rice plants. Proc Natl Acad Sci 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 108:1723–1728

Gillman JD, Bentolila S, Hanson M (2007) The petunia restorer of fertility protein is part of a large mitochondrial complex that interacts with transcripts of the CMS-associated locus. Plant J 49:217–227

Grelon M, Budar F, Bonhomme S, Pelletier G (1994) Ogura cytoplasmic male-sterility (CMS)-associated *orf138* is translated into a mitochondrial membrane polypeptide in male-sterile *Brassica* cybrids. Mol Gen Genet 243:540–547

Hanson MR (1991) Plant mitochondrial mutations and male sterility. Annu Rev Genet 25:461–486

Hanson MR, Bentolila S (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell 16:154–169

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

Janska H, Sarria R, Woloszynska M, Arrieta-Montiel M, Mackenzie SA (1998) Stoichiometric shifts in the common bean mitochondrial genome leading to male sterility and spontaneous reversion to fertility. Plant Cell 10:1163–1180

Jarl CI, van Grinsven M, van den Mark F (1989) Correction of chlorophyll-defective male-sterile winter oilseed rape (*Brassica napus*) through organelle exchange: molecular analysis of the cytoplasm of parental lines and corrected progeny. Theor Appl Genet 77:135–141

Johnston JS, Pepper AE, Hall AE, Chen ZJ, Hodnett G, Drabek J, Lopez R, Price HJ (2005) Evolution of genome size in Brassicaceae. Ann Bot 95:229–235

Kamei A, Tsuro M, Kubo N, Hayashi T, Wang N, Fujimura T, Hirai M (2010) QTL mapping of clubroot resistance in radish (*Raphanus sativus* L.). Theor Appl Genet 120:1021–1027

Kim S, Lim H, Park S, Cho K, Sung S, Oh D, Kim K (2007) Identification of a novel mitochondrial genome type and development of molecular makers for cytoplasm classification in radish (*Raphanus sativus* L.). Theor Appl Genet 115:1137–1145

Kim S, Lee Y, Lim H, Ahn Y, Sung S (2009) Identification of highly variable chloroplast sequences and development of cpDNA-based molecular markers that distinguish four cytoplasm types in radish (*Raphanus sativus* L.). Theor Appl Genet 119:189–198

Kim K, Lee Y, Lim H, Han T, Sung S, Kim S (2010) Identification of *Rfd1*, a novel restorer-of-fertility locus for cytoplasmic male-sterility caused by DCGMS cytoplasm and development of simple PCR markers linked to the *Rfd1* locus in radish (*Raphanus sativus* L.). Euphytica 175:79–90

Klein RR, Klein PE, Mullet JE, 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 collinear region of rice chromosome 12. Theor Appl Genet 111:994–1012

Kmiec B, Woloszynska M, Janska H (2006) Heteroplasmy as a common state of mitochondrial genetic information in plants and animals. Curr Genet 50:149–159

Knoop V (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. Curr Genet 46:123–139

Koizuka N, Imai R, Fujimoto H, Hayakawa T, Kimura Y, Kohno-Murase J, Sakai T, Kawasaki S, Imamura J (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, Takakura Y, Kuraya Y, Suzuki S, Hiei Y, Imaseki H, Nitta N (2004) Map-based cloning of a fertility restorer gene, *Rf-1*, in rice (*Oryza sativa* L.). Plant J 37:315–325

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Kubo T, Newton KJ (2008) Angiosperm mitochondrial genomes and mutations. Mitochondrion 8:5–14

Lagercrantz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150:1217–1228

Lee Y, Park S, Lim C, Kim H, Lim H, Ahn Y, Sung S, Yoon M, Kim S (2008) Discovery of a novel cytoplasmic male-sterility and its restorer lines in radish (*Raphanus sativus* L.). Theor Appl Genet 117:905–913

Lee Y, Kim S, Lim H, Ahn Y, Sung S (2009) Identification of mitochondrial genome rearrangements unique to novel cytoplasmic male sterility in radish (*Raphanus sativus* L.). Theor Appl Genet 118:718–728

Menczel L, Morgan A, Brown S, Maliga P (1987) Fusion-mediated combination of Ogura-type cytoplasmic male sterility with *Brassica napus* plastids using X-irradiated CMS protoplasts. Plant Cell Rep 6:98–101

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci 88:9828–9832

Ogura H (1968) Studies on the new male sterility in Japanese radish, with special reference to the utilization of this sterility towards the practical raising of hybrid seeds. Mem Fac Agr Kagoshima Univ 6:39–78

Oldenburg DJ, Bendich AJ (2001) Mitochondrial DNA from the Liverwort *Marchantia polymorpha*: circularly permuted linear molecules, head-to-tail concatemers, and a 5′ protein. J Mol Biol 310:549–562

Palmer JD (1988) Intraspecific variation and multicircularity in *Brassica* mitochondrial DNAs. Genetics 118:341–351

Pelletier G, Primard C, Vedel F, Chetrit P, Remy R, Rousselle P, Renard M (1983) Intergeneric cytoplasmic hybridization in Cruciferae by protoplast fusion. Mol Gen Genet 191:244–250

Sakai T, Imamura J (1993) Evidence for a mitochondrial sub-genome containing radish *AtpA* in a *Brassica napus* cybrid. Plant Sci 90:95–103

Shedge V, Arrieta-Montiel M, Christensen AC, Mackenzie SA (2007) Plant mitochondrial recombination surveillance requires unusual *RecA* and *MutS* homologs. Plant Cell 19:1251–1264

Shirasawa K, Oyama M, Hirakawa H, Sato S, Tabata S, Fujioka T, Kimizuka-Takagi C, Sasamoto S, Watanabe A, Kato M, Kishida Y, Kohara M, Takahashi C, Tsuruoka H, Wada T, Sakai T, Isobe S (2011) An EST-SSR linkage map of *Raphanus sativus* and comparative genomics of the Brassicaceae. DNA Res 18:221–232

Small I, Suffolk R, Leaver CJ (1989) Evolution of plant mitochondrial genomes via substoichiometric intermediates. Cell 58:69–76

Song KM, Osborn TC, Williams PH (1990) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 3. Genome relationships in *Brassica* and related genera and the origin of *B. oleracea* and *B. rapa* (syn. *campestris*). Theor Appl Genet 79:497–506

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3:739–744

The *Brassica rapa* genome sequencing project consortium (2011) The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet 43:1035–1039

Thormann CE, Ferreira ME, Camargo LEA, Tivang JG, Osborn TC (1994) Comaprison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. Theor Appl Genet 88:973–980

Tsuro M, Suwabe K, Kubo N, Matsumoto S, Hirai M (2008) Mapping of QTLs controlling root shape and red pigmentation in radish, *Raphanus sativus* L. Breed Sci 58:55–61

Vos P, Hogers R, Bleeker M, Reijans M, Van Der Lee T, Hornes M, Friiters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wang Z, Zou Y, Li X, Zhang Q, Chen L, Wu H, Su D, Chen Y, Guo J, Luo D, Long Y, Zhong Y, Liu Y (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

Warwick SI, Black LD (1997) Phylogenetic implications of chloroplast DNA restriction site variation in subtribes Raphninae and Cakilinae (Brassicaceae, tribe Brassiceae). Can J Bot 75:960–973

Yang Y, Lai K, Tai P, Ma D, Li W (1999) Molecular phylogenetic studies of Brassica, Rorippa, Arabidopsis and allied genera based

on the internal transcribed spacer region of 18S–25S rDNA. Mol Phylogenet Evol 13:455–462

Yang Y, Tai P, Chen Y, Li W (2002) A study of the phylogeny of *Brassica rapa*, *B. nigra*, *Raphanus sativus*, and their related genera using noncoding regions of chloroplast DNA. Mol Phylogenet Evol 23:268–275

Zaegel V, Guermann B, Le Ret M, Andrés C, Meyer D, Erhardt M, Canaday J, Gualberto JM, Imbault P (2006) The plant-specific ssDNA binding protein OSB1 is involved in the stoichiometric transmission of mitochondrial DNA in *Arabidopsis*. Plant Cell 18:3548–3563