Theor Appl Genet (2006) 113:397–406
DOI 10.1007/s00122-006-0303-5

ORIGINAL PAPER

Analysis of *MS2Bnap* genomic DNA homologous to *MS2* gene from *Arabidopsis thaliana* in two dominant digenic male sterile accessions of oilseed rape (*Brassica napus* L.)

S. W. Hu · Y. F. Fan · H. X. Zhao · X. L. Guo · C. Y. Yu · G. L. Sun · C. H. Dong · S. Y. Liu · H. Z. Wang

Received: 23 July 2005 / Accepted: 28 April 2006 / Published online: 7 June 2006

Abstract PCR technique was employed to isolate gene homologous to the *MS2Bnap* (X99922.1) from two rapeseed (*Brassica napus* L.) dominant digenic male sterile lines, namely 220A (male sterile) and 220B (male fertile), 6A (male sterile) and 6C (male fertile). The isolated 2,581 bp sequences from 220A (named 220A-gDNA, GenBank accession number AY288778), 220B (220B-gDNA, AY257490), 6A (6A-gDNA, DQ060318) and 6C (6C-gDNA, DQ060319) all contained six introns. Forty-one single nucleotide polymorphism (SNP) sites were detected by alignment of these four sequences, seven of them dispersed in the exon regions. Two SNPs (1247, 1656) were detected between 220A-gDNA and 220B-gDNA, and the one at nucleotide 1247 of 220A-gDNA with A replaced by C was a missense mutation, which may be the putative male sterility site in 220A. All eight SNPs identified between 6A-gDNA and 6C-gDNA were located in the third intron, so the proteins encoded by them are the same. The one SNP between 6A-/6C-gDNA and 220A-/220B-gDNA at nucleotide 2474 of 220A-/220B-gDNA with C replaced by G was a missense mutation. Mutation site of BNMS2PROT (CAA68190.1) encoded by *MS2Bnap* in 220A(254) and 6A/6C(584) is different, which indicated dominant digenic male sterile line 220AB and 6CA have some difference in the molecular level. Comparison of structure of *MS2Bnap* in *B. napus* with that of *MS2* in *Arabidopsis thaliana* revealed that the similarity of exons between these two genes is higher than that of introns.

Communicated by H. C. Becker

S. W. Hu · Y. F. Fan · X. L. Guo · C. H. Dong · S. Y. Liu · H. Z. Wang (✉)
Oil Crops Research Institute,
Chinese Academy of Agricultural Sciences,
Wuhan, Hubei 430062, People's Republic of China
e-mail: wanghz@oilcrops.cn

S. W. Hu · Y. F. Fan · C. Y. Yu
College of Agronomy, Northwest A&F University,
Yangling, Shaanxi 712100, People's Republic of China

H. X. Zhao
College of Life Sciences, Northwest A&F University,
Yangling, Shaanxi 712100, People's Republic of China

G. L. Sun
Department of Biology, Saint Mary's University,
B3H 3C3 Halifax, Nova Scotia, Canada

Introduction

Male sterility is an inheritable trait characterized by the inability of a plant to produce functional pollen. Two main types of male sterility in plant were discovered: cytoplasmic male sterility (CMS) and genic male sterility (GMS). In oilseed rape, both type of the male sterility have been widely used in the breeding program for utilizing heterosis.

Remarkable progresses have been made in isolating and cloning CMS-related genes and their restoring genes (Cui et al. 1996; Bentolila et al. 2002; Brown et al. 2003; Desloire et al. 2003; Koizuka et al. 2003; Komori et al. 2004), and understanding molecular mechanisms of sterilization and restoration (Hanson 1991; Bentolila et al. 2002). However, there is a little information available on the molecular mechanism of GMS, although GMS has been a valuable system for hybrid seed production and improvement of plant populations

(Rao et al. 1990). Up to date, several types of GMS have been reported in rapeseed (reviewed by Fu 2000), such as monogenic dominant GMS (Mathias 1985; Wang et al. 2003), digenic dominant GMS (DDGMS) (Li et al. 1985; Dong and Du 1993; Wang et al. 1999; Hu 2003), monogenic recessive GMS (Chowdhury and Das 1967; Das and Pandey 1961; Luo et al. 1992; Takagi 1970; Zuberi and Zuberi 1983), digenic recessive GMS (Hou et al. 1990; Pan and Zhao 1990; Sun et al. 1997) and multigenic recessive GMS (Chen et al. 1998; Sun et al. 2002; Wang et al. 2001). Molecular markers associated with male fertility-related genes in rapeseed GMS have been identified (Tu et al. 1999; Gan et al. 1999; Jiang et al. 2000; Lu et al. 2004; Hu et al. 2003; Ke et al. 2005).

MS2 gene determines a nuclear male sterile mutant (monogenic recessive) phenotype in *Arabidopsis* and was isolated and characterized using the *En/spm-I/dspm* transposon-tagging system (Aarts et al. 1993). *MS2* gene encodes one open reading frame of 616 amino acids, which probably located in microbodies of tapetum cells at the time of exine wall deposition on the young microspores. The possible function of the MS2 protein was proposed as a fatty acyl reductase in the formation of pollen wall substances (Aarts et al. 1997). The homologue of *MS2* in *Brassica napus*, *MS2Bnap*, is isolated by cold plague screening from *B. napus* anther specific cDNA library (Hodge et al. 1992). Li et al. (2002) isolated a fragment homologous to *MS2Bnap* gene from digenic recessive GMS line S45AB of *B. napus* using RT-PCR technique, and found that there existed an amino acid difference between fertile S45B and sterile S45A, which may be the putative male sterility site in S45A. However, the full-length genomic DNA sequence corresponding to *MS2Bnap* has not been reported.

So far, four DDGMS accessions have been reported in *B. napus* (Li et al. 1985; Dong and Du 1993; Wang et al. 1999; Hu 2003). The previous investigations indicated that the maintainers and restorers of DDGMS Qianyou 2AB were different from those of 6CA (Wang et al. 1999), while Shaan-GMS has similar maintainers and restorers with 6CA (Hu 2003), however the molecular markers for dominant male sterility gene *Ms* was different from each other (Hu et al. 2003). In this paper, we used two DDGMS lines, namely 220AB, and a F_1 between 6CA (Li et al. 1985) and a breeding line 220, to clone the genomic DNA sequence corresponding to *MS2Bnap*. The objectives of this study are (i) to compare the genomic DNA sequence corresponding to *MS2Bnap* from 220A, 220B, 6C and 6A (refers to the male fertile plants and male sterile plants in the F_1 between 6CA and the breeding line 220, respectively), with the *MS2Bnap* to reveal their *MS2Bnap* gene structure; (ii) to reveal if there exist any difference between *MS2Bnap* gene in 220A and 220B, and that between 6A and 6C; (iii) to compare *MS2Bnap* genes in *B. napus* with *MS2* gene in *Arabidopsis thaliana* to reveal their relationship in the molecular level.

Materials and methods

Plant materials

Two kinds of rapeseed DDGMS accessions, namely 220AB and F_1 between a DDGMS line 6CA and a breeding line 220 were used in this study. 220AB was derived from a DDGMS Shaan-GMS crossed with the line 220, followed by successive backcrossing to the line 220 (Hu 2003). In the population 220AB, male sterile plants 220A (Msmsrfrf) and male fertility plants 220B (msmsrfrf) are isogenic lines, which differ in the fertility-related locus Ms–ms. In the F_1 population derived from a cross of 6CA and the line 220, male sterility plants (called as 6A) with genotype Msmsrfrf and male fertility plants (called as 6C) with genotype msmsrfrf are also isogenic in the locus Ms–ms.

Isolation of genomic DNA

Five plants were randomly collected from each of 220A, 220B, 6A and 6C at the flowering stage. The total genomic DNA of each plant was extracted according to the protocol of Saghai-Maroof et al. (1984).

PCR amplification

Six primers (Table 1, synthesized by BioAsia Corporation in Shanghai, China) reported by Li et al. (2002) were used to amplify the genomic DNA fragment corresponding to *MS2Bnap* gene in all five plants of each of 220A, 220B, 6A or 6C. The total reaction volume for DNA amplification was 20 μl. Reaction mixtures contained 10 mM Tris–HCl (pH 8.8 at 25°C), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM each of dATP, dGTP, dCTP and dTTP (Promega, Madison, WI, USA), 0.5 μM each of the two primers, 1.0 units EX Taq DNA polymerase (Takara Biotech. Co., Dalian, China) and 50 ng genomic DNA. DNA amplification was performed using MJ Research thermocycler, Model PTC-200 (Watertown, MA, USA) with a program of 3 min at 94.0°C, followed by 35 cycles of 30 s at 94.0°C, 30 s at 60°C and 1 min at 72.0°C, then ended with the final extension for 10 min at 72.0°C. After amplification, PCR products were separated by electrophoresis in 1.5% agarose gel with 1 × TAE buffer, stained with ethidium bromide

Table 1 Primer sequences and their positions in the *MS2Bnap* complete sequence

No. of primer	Sequence of primers	Primers' place in the sequence of *MS2Bnap*-cDNA
P1	5′-AATGGAATGGACAGTTTACTGTC-3′	1–23
P2	5′-GAAGGTTGTATTGGCAGCTGAG-3′	778–799
P3	5′-CAGATTCAGCAGAGGAGATTGC-3′	732–753
P4	5′-GAAGCTCAGCTAAGTCCTCG-3′	1,483–1,502
P5	5′-CGTGTATCAGATCGCTTC-3′	1,441–1,458
P6	5′-TTTGACCTAAGCCCTTCC-3′	1,909–1,927

and photographed by Image Master® VDS (Pharmacia, Sweden) under UV light.

Cloning PCR products and data analysis

The expected bands of PCR products were harvested and purified using DNA harvesting kit (Sangon Corporation, Shanghai, China), and then cloned into pGEM®-T easy vector according to the manufacturer's instruction (Promega, Madison, WI, USA). The plasmid DNA was isolated and sequenced by Sangon Corporation. The sequence data were analyzed using the DNAMAN software Version 3.0 (Lynnon BioSoft, Quebec, Canada). For each sample, five clones were randomly selected for sequencing. The sequences of the genomic DNA fragment corresponding to *MS2Bnap* gene in 220A, 220B, 6A and 6C were finally confirmed by sequencing five plants, each with five clones.

Full-length genomic DNA (+ 48172 to + 51053 of AP002040.1) for *MS2* (X73652.1) gene in *A. thaliana* was obtained by searching DDBJ database with Blast program (http://www.ncbi.nlm.nih.gov/entrez/nucleotide.html). The gene structure of *MS2Bnap* in *B. napus* and *MS2* in *A. thaliana* was compared using the DNAMAN software Version 3.0 too.

Results

PCR amplification

Of the primers tested, three pairs, P1/P2, P3/P4 and P5/P6 amplified DNA fragments from the genomic DNA of 220A, 220B, 6A and 6C (*B. napus*). Three expected fragments (about 1.1 kb, 1.0 kb, 570 bp) were obtained from both 220A and 220B (Fig. 1), and from 6A and 6C. (The figure was not showed.)

Sequence analysis of genomic DNA corresponding to *MS2Bnap*

Three fragments (Fig. 1) amplified from male fertile 220B and male sterile 220A, and from male fertile 6C and male sterile 6A were cloned and sequenced. The sequence analysis indicated that fragments I, II and III were 1,122, 1,017, 570 bp in length, respectively. By overlapping the sequences of these three fragments, we got a full-length of genomic DNA sequence corresponding to *MS2Bnap* from 220A and 220B, and from 6A and 6C, and named them as 220A-gDNA (GenBank accession number AY288778), 220B-gDNA (AY257490), 6A-gDNA (DQ060318) and 6C-gDNA (DQ060319), respectively. These four sequences all are 2,581 bp in length from transcription initiation site to end site. Comparison of these sequences with the reported *MS2Bnap* detected six introns with total length of 654 bp (Fig. 2). These six introns positioned at + 504 to + 738 bp (235 bp), + 843 to + 930 bp (88 bp), + 1,131 to + 1,223 bp (93 bp), + 1,325 to + 1,399 bp (75 bp), + 1,764 to + 1,842 bp (79 bp) and + 2,335 to + 2,418 bp (84 bp), respectively. Each intron had a typical structure of GT/AG dinucleotides at the both ends. The AT content ranged from 66.67 to 76.19% with an average of 70.18%.

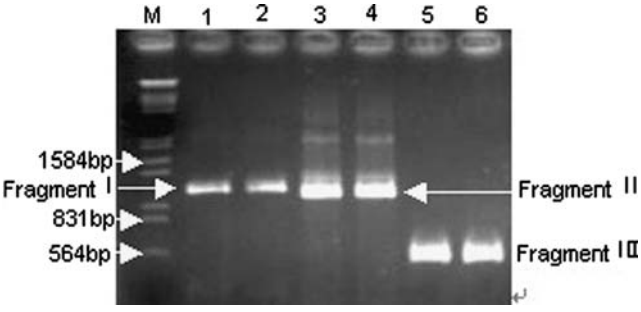

Fig. 1 Agarose gel electrophoresis pattern of three pair primers. *Lanes 1 and 2* the PCR products amplified with primer pair P1/P2 for fragment I, *lanes 3 and 4* the PCR products amplified with primer pair P3/P4 for fragment II, *lanes 5 and 6* the PCR products amplified with primer pair P5/P6 for fragment III, *M* λ DNA marker digested by *EcoR* I and *Hind* III, *lanes 1, 3 and 5* DNA pattern from 220B plants (male fertile), *lanes 2, 4 and 6* DNA pattern from 220A plants (male sterile)

Alignments of the sequences of 220A-gDNA, 220B-gDNA, 6A-gDNA and 6C-gDNA detected 41 single nucleotide polymorphism (SNP) sites. Of them, seven SNPs were located in the exon regions (Table 2). There

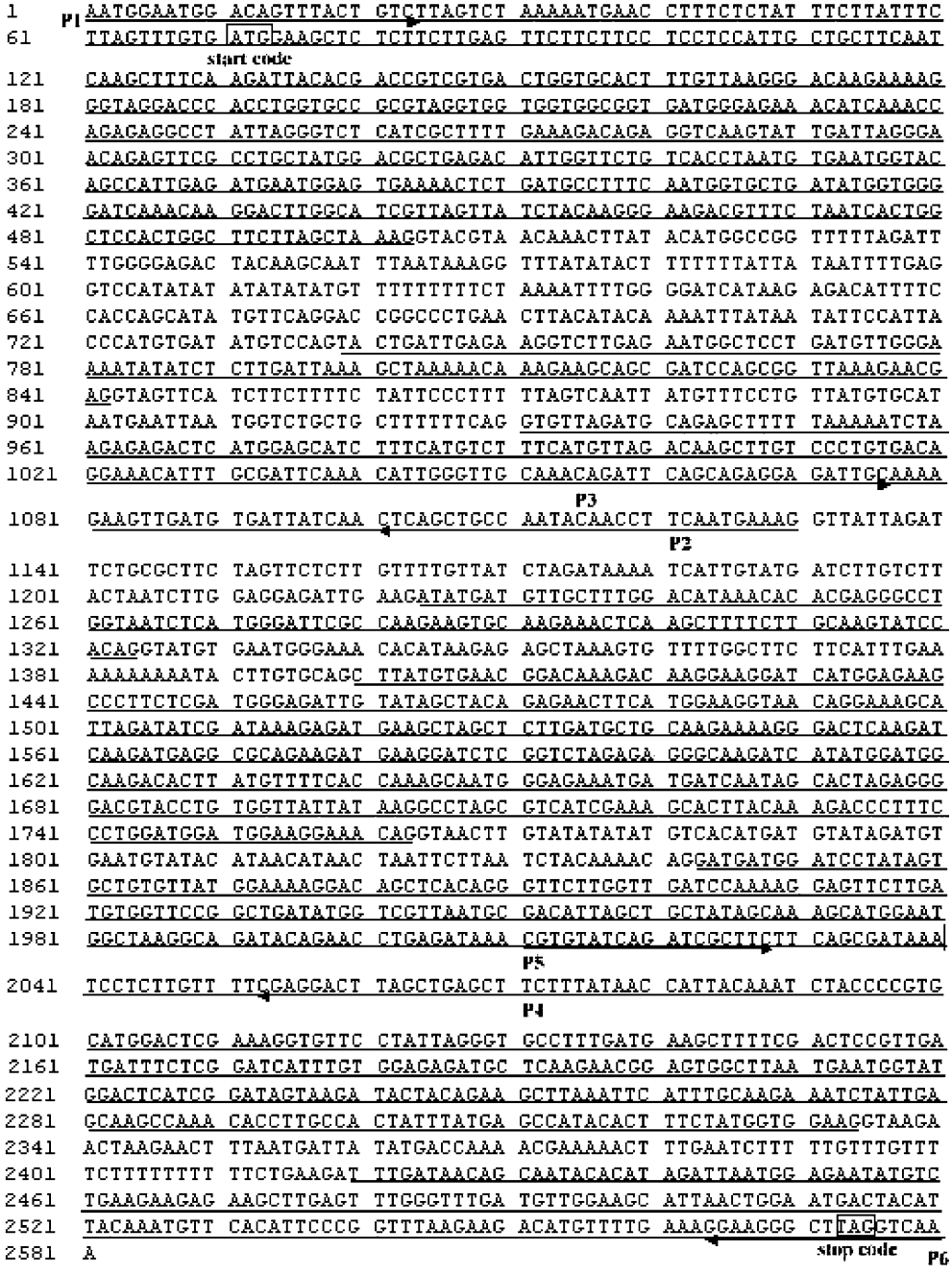

Fig. 2 Structure of *MS2Bnap* gene from rapeseed "220B" (*B. napus*). *Sequence underlined* exons, *sequence without underlined* introns, *sequence under-arrowed* primers, *ATG* start codon, *TAG* stop codon

existed two SNPs (1247, 1656) between 220A-gDNA and 220B-gDNA, and the one at nucleotide 1247 of 220A-gDNA with A replaced by C was a missense mutation. There existed eight SNPs between 6A-gDNA and 6C-gDNA, which all located in the third intron region. The one SNP between 6A-gDNA/6C-gDNA and 220A-gDNA/220B-gDNA at nucleotide 2474 of 220B-gDNA with C replaced by G was a missense mutation.

Each of 220B-gDNA, 220A-gDNA, 6C-gDNA and 6A-gDNA contained an open reading frame of 1,848 bp (CDS:71–503, 740–843, 932–1,131, 1,225–1,325, 1,399–1,762, 1,842–2,333, 2,418–2,580 bp) that encodes a putative polypeptide of 616 amino acids. The protein, 220B-PROT encoded by 220B-gDNA showed 100% identity to the protein, BNMS2PROT (CAA68190.1) encoded by *MS2Bnap*. Alignment of 220B-PROT and the protein, 220A-PROT encoded by

Table 2 Single nucleotide polymorphisms (SNPs) among four *MS2Bnap* genomic DNA fragments between two different DDGMS accessions in *B. napus*

Locations of SNPs	I_3										E_4		I_4		
	1142	1145	1147	1179	1183	1184	1187	1196	1207	1221	1247	1320	1328	1336	1337
220B-gDNA	C	C	C	A	A	A	T	G	C	A	A	C	T	G	G
220A-gDNA	C	C	C	A	A	A	T	G	C	A	C[a]	C	T	G	G
6C-gDNA	C	C	C	A	A	A	T	G	T	C	A	T	C	C	C
6A-gDNA	A	T	T	T	T	T	C	A	T	C	A	T	C	C	C
MS2Bnap											A	T			

Locations of SNPs	I_4														
	1348	1349	1351	1352	1353	1355	1356	1358	1367	1370	1373	1374	1378	1379	1380
220B-gDNA	G	A	A	G	C	A	A	G	C	C	C	A	G	A	A
220A-gDNA	G	A	A	G	C	A	A	G	C	C	C	A	G	A	A
6C-gDNA	A	T	T	T	T	G	T	A	G	T	T	T	A	T	T
6A-gDNA	A	T	T	T	T	G	T	A	G	T	T	T	A	T	T
MS2Bnap															

Locations of SNPs	I_4				E_5	I_5	E_6			I_6	E_7
	1381	1390	1391	1394	1656	1175	2077	2083	2206	2412	2474
220B-gDNA	A	A	C	G	A	A	T	T	C	T	C
220A-gDNA	A	A	C	G	G	A	T	T	C	T	C
6C-gDNA	T	T	A	A	A	T	C	C	C	G	G[a]
6A-gDNA	T	T	A	A	A	T	C	C	C	G	G[a]
MS2Bnap					A		T	T	A		C

I intron, *E* exon

[a] Refers to missense mutants

220A-gDNA detected one amino acid difference located at 254 in BNMS2PROT (Fig. 3), which is corresponding to the location 1247 in the nucleotide sequence for 220A-gDNA (Table 2). The protein, 6C-PROT encoded by 6C-gDNA is the same as the protein, 6A-PROT encoded by 6A-gDNA, however, one amino acid difference between 6C-PROT/6A-PROT and BNMS2PROT was detected at the location of 584 of BNMS2PROT (Fig. 3) which corresponds to the location 2474 in the nucleotide sequence of 6C-gDNA/6A-gDNA (Table 2).

Comparison of *MS2Bnap* genes in *B. napus* with *MS2* gene in *A. thaliana*

The reported *MS2* gene (X73652.1) was 2,126 bp in length, which contained an open reading frame of 1,848 bp (+ 72 to + 1922 in *MS2*) that encodes a polypeptide of 616 amino acids. The full-length genomic DNA sequence corresponding to *MS2* gene locates in the region from + 48172 to + 51053 of the DNA sequence (AP002040.1). Alignment of *MS2* gene with its full-length genomic DNA detected eight introns with total length of 770 bp (Fig. 4). These eight introns positioned at + 66 to + 242 bp (177 bp), + 682 to + 765 bp (84 bp), + 870 to + 946 bp (77 bp), + 1,147 to + 1,237 bp (91 bp), + 1,339 to + 1,407 bp (69 bp), + 1,610 to + 1,683 bp (74 bp), + 1,846 to + 1,951 bp (106 bp) and + 2,444 to + 2,355 bp (912 bp), respectively. Each intron had a typical structure of GT/AG dinucleotides at the both ends. The AT content ranged from 60 to 76% with an average of 68.25%.

Comparison of full-length genomic DNA of *MS2Bnap* in *B. napus* with that of *MS2* gene in *A. thaliana* revealed that the similarity between exons of them ranged from 50.8 to 96.3% with the average of 85.6% and the similarity between introns of them ranged from 48.0 to 76.8% with the average of 60.7% (Figs. 4, 5). The first exon of *B. napus MS2Bnap* gene was corresponded to the second exon of *A. thaliana MS2* gene with the similarity of 83.9% (which is shorted for Ebn-1-Eat-2), Ebn-2-Eat-3 (90.4%), Ebn-3-Eat-4 (90.%), Ebn-4-Eat-5 (92.1%), Ebn-5-Eat-6 + Eat-7 (91.1, 96.3%), Ebn-6-Eat-8 (88.6%) and Ebn-7-Eat-9 (87.7%). And the second intron of *B. napus MS2Bnap* gene was corresponded to the third intron of *A. thaliana MS2* gene with the similarity of 64.9% (which is shorted for Ibn-2-Iat-3), Ibn-3-Iat-4 (76.8%), Ibn-4-Iat-5 (64.2%),

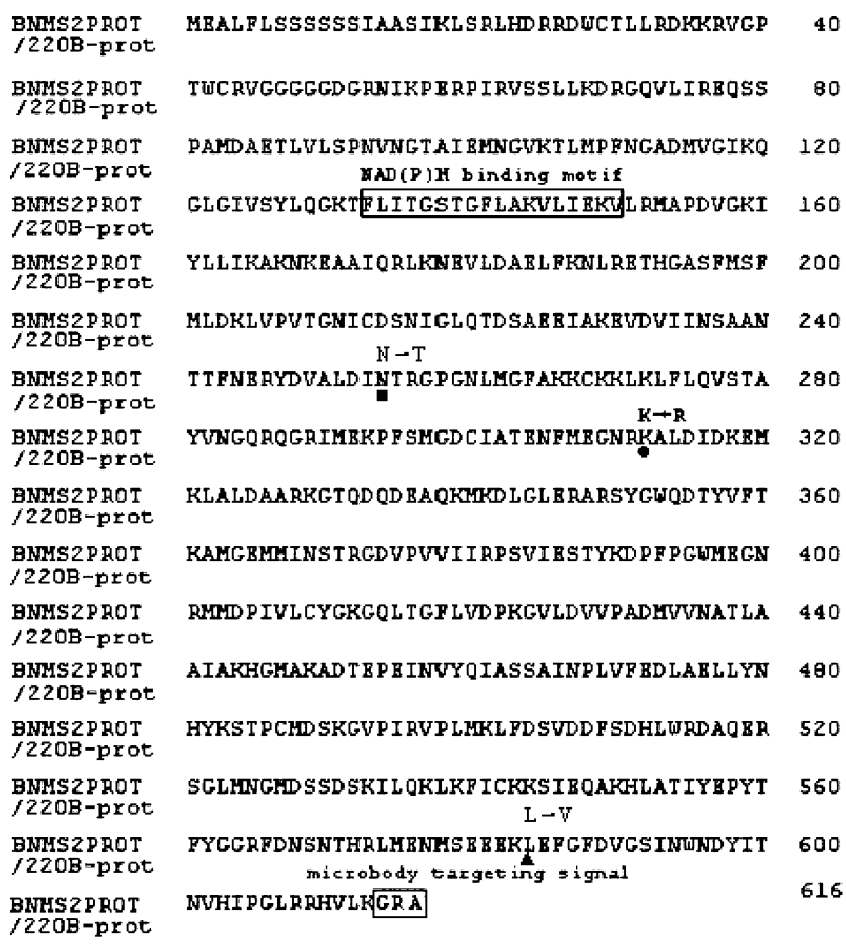

Fig. 3 Comparison of the predicted amino acid sequence of BNMS2PROT (CAA68190.1), 220B-PROT, 220A-PROT and 6A-PROT. *Sequence boxed* related motif, *letter dotted with filled square* a putative male sterility amino acid residue of 220A, *letter dotted with filled upper triangle* a mutant amino acid residue of 6CA, *letter dotted with filled circle* a putative male sterility amino acid residue of S45A (adapted from Li et al. 2002)

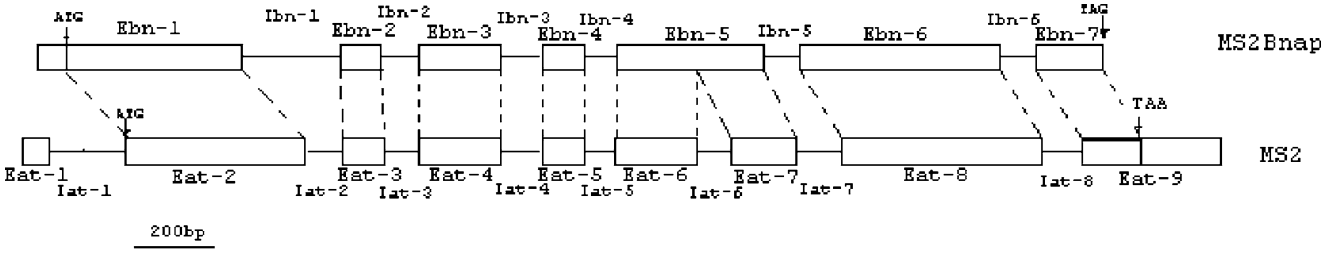

Fig. 4 Comparison of the structure of *MS2Bnap* gene in *B. napus* and *MS2* gene in *A. thaliana Ebn-1*, *Ebn-2*,...,*Ebn-7* refers to exons and *Ibn-1*, *Ibn-2*,...,*Ibn-6* refers to introns of *B. napus MS2Bnap* gene, respectively; *Eat-1*, *Eat-2*,...,*Eat-9* refers to exons and *Iat-1*, *Iat-2*,...,*Iat-8* refers to introns of *A. thaliana MS2* gene, respectively; *ATG* start codon,; *TAG/TAA* stop codon

Ibn-5-Iat7 (63.2%), Ibn-6-Iat-8 (63.2%) and Ibn-1-Iat-1, Iat-2, Iat-6 (48.0, 48.0, 54.1%).

Discussion

Four genomic DNA sequences corresponding to *MS2Bnap* (x99922.1), namely 220B-gDNA and 220A-gDNA in rapeseed DDGMS line 220B and 220A (*B. napus* L), and 6C-gDNA and 6A-gDNA in the F_1 between 6CA crossed with the breeding line 220 were isolated by PCR technique. These four sequences are 2,581 bp in length, and composed of six introns with total of 654 bp in length compared with the sequence of *MS2Bnap*. A typical structure of GT and AG is present at the both ends of each intron. Blasting in NCBI indicated that the sequences of 220B-gDNA (AY257490), 220A-gDNA (AY288778), 6C-gDNA (DQ060319) and 6A-gDNA (DQ060318) were the first cloned genomic sequences homologous to *MS2Bnap* in *B. napus*.

MS2 gene in *A. thaliana* was the first cloned and characterized fertility-related gene using the maize

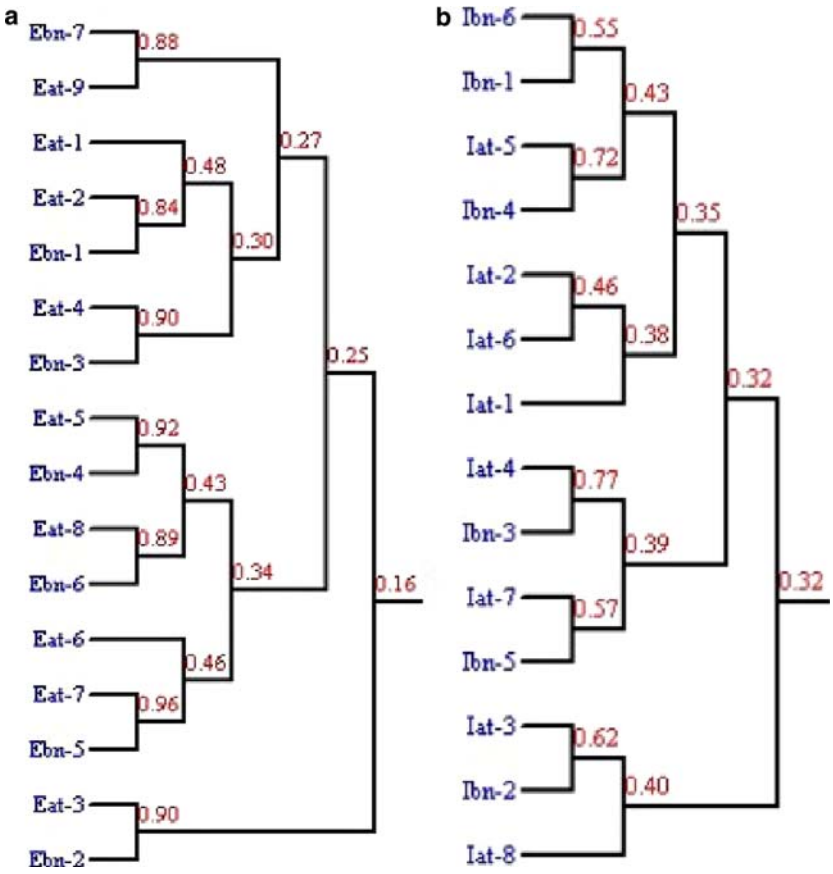

Fig. 5 Comparison of *MS2Bnap* gene in *B. napus* with *MS2* gene in *A. thaliana*. **a** Dendrogram of exons. *Ebn-1*, *Ebn-2*,...,*Ebn-7* refers to exons of *MS2Bnap* gene and *Eat-1*, *Eat-2*,...,*Eat-9* refers to exons of *MS2* gene. **b** Dendrogram of introns. *Ibn-1*, *Ibn-2*,...,*Ibn-6* refers to introns of *MS2Bnap* gene and *Iat-1*, *Iat-2*,...,*Iat-8* refers to introns of *MS2* gene

Enhancer–Inhibitor transposable element system (Aarts et al. 1993). The expression of *MS2* is confined to the tapetum during the period starting at the time of microspore release from tetrads and ending before the first mitotic microspore division (Aarts et al. 1997). *MS2* gene encodes a polypeptide of 616 amino acids, which has two important motifs. One is a conserved motif [I, V, F]-X-[I, L, V]-T-G-F-L-[G, A] at positions 133–150 which has a proposed role in NAD(P)H binding, and another is an addition C-terminal microbody targeting signal. The MS2 products are probably located in microbodies of tapetum cells at the time of exine wall deposition on the young microspores. The major component of exine is sporopollenin. Reduction of fatty acyl groups to fatty alcohol groups may be one of the steps in the formation of sporopollenin. A possible function of the MS2 protein as a fatty acyl reductase in the formation of pollen wall substances was proposed by Aarts et al. (1997). Mutation of *ms2* is caused by frameshift or stop mutations resulting from the excision of the inserted I element in the target region AAA of ACA AAC (position + 1793 to + 1803 of *MS2*, corresponding position + 600 aa to + 601 aa of MS2 product), which resulted in the products of these mutants missing the last three amino acids (Gly-Arg-Ala, + 614 aa to + 616 aa), a C-terminal microbody targeting signal, and unable to locate in microbodies of tapetum cell. Thus the functional pollen cannot be formed in these mutants. *MS2Bnap*, a homologue of *MS2* in *B. napus*, coded a very similar protein BNMS2PROT (CAA68190.1) involved in the formation of functional pollen. Li et al. (2002) reported that there were four diverged nucleotides dispersed in the coding region of the fragment homologous to *MS2Bnap* in rapeseed digenic recessive GMS line S45AB, three of them were synonymous mutations, one was missense mutation, which maybe the male sterile locus of S45A (Fig. 3). Our results showed that sequence of 220B-gDNA contained an open reading frame of 1,848 bp and encoded a polypeptide of 616 amino acids, which showed 100% identity to BNMS2PROT. However, owning to the mutation occurred at + 1,274 bp in the sequence of 220A-gDNA, 220A-gDNA encoded a polypeptide (220A-PROT), which has one amino acid difference compared to BNMS2PROT (Fig. 3). In the population of 220AB, 220A and 220B are isogenic lines with only difference in the fertility-related locus Ms–ms. It could be

inferred that the one amino acid difference of 220A-PROT in 220A plants may be one of the reasons that 220A plants express male sterility. However, this hypothesis needs to be confirmed by further evidences such as gene transformation. The protein encoded by 6C-gDNA and 6A-gDNA is the same. While, there was one amino acid difference between 6C-PROT/6A-PROT and BNMS2PROT, which indicated that this amino acid difference may have no relationship with the male sterile of 6A plants. Furthermore, mutation site of BNMS2PROT encoded by *MS2Bnap* in 220A(254) and 6A/6C(584) is different, which indicated DDGMS 220AB and 6CA have some difference in the molecular level. This was also consistent with the result of Hu et al. (2003), who found that the molecular markers for dominant male sterility gene *Ms* in Shaan-GMS was different from that in 6A.

Arabidopsis thaliana (the model dicotyledonous plant) is closely related to *Brassica* crop species. Comparative genomics indicated that there exist genome-wide collinearity between these two species (Lagercrantz et al. 1996; Cavell et al. 1998; Bancroft and O'Neill 2000). Furthermore, the similarity between expression sequences in *B. napus* and their counterparts in *Arabidopsis* reach to 85%, and a large amount of genes in *A. thaliana* have similar functions with their counterparts in *B. napus* (Cavell et al. 1998; Sillito et al. 2000). Li et al. (2003) reported a global gene for gene alignment of the genomes of *Brassica oleracea* and *A. thaliana* by construction of a transcriptome map based on *B. oleracea* cDNAs obtained from leaf tissue. The map consisted of 247 cDNA markers obtained by the sequence-related amplified polymorphism technique. After sequencing 190 of the polymorphic cDNA bands, FASTA detected 169 sequences with similarity to genes reported in *Arabidopsis*. There was extensive collinearity between the two genomes for chromosomal segments rather than for whole chromosomes, often showing inversions and deletions/insertions.

Alignment of four DNA sequences (220B-gDNA, 220A-gDNA, 6C-gDNA and 6A-gDNA) detected 41 SNP sites, however only seven of them are dispersed in the exon region. Comparison of full-length genomic DNA of *MS2Bnap* in *B. napus* with that of *MS2* gene in *A. thaliana* revealed that the similarity between these two sequences in exon region [85.64%, this figure was corresponded to the results of Cavell et al. (1998) and Sillito et al. (2000), who found the similarity between expression sequences in *B. napus* and its counterparts in *Arabidopsis* reach to 85%] is greater than that in intron region (60.7%). These results are consistent with the findings of Fourmann et al. (2002), who used 32 amplified consensus gene markers to amplify homologous gene sequences from one line of *A. thaliana* and three *Brassica* species (four lines of *B. napus*, one line of *B. oleracea* and one line of *Brassica rapa*), and found that a very high level of conservation was observed between coding sequences in the three *Brassica* species and *A. thaliana.* Furthermore, by sequencing 58 of the 102 genes for more than one rapeseed genotype and comparison the exons and introns of 22–35 sequences for each available pair line, they found that, as expected, there were many more (about four-times more) changes in intron than in exon sequences. Molhoj et al. (2001) also found that a membrane-anchored endo-1,4-β-glucanase from *B. napus* and its orthologous *KOR* from *A. thaliana* have identical exon lengths intercepted by the introns located at the identical position in the coding region.

So far, four DDGMS accessions have been reported in *B. napus* (Li et al. 1985; Dong and Du 1993; Wang et al. 1999; Hu 2003). Genetic investigations have revealed that male fertility of these DDGMS was controlled by two pair of nuclear genes (Li et al. 1985; Hu 2003). If male sterility gene was designated as Ms, its allele recessive gene as ms, dominant inhibition gene as Rf, which can inhibit the expression of the Ms and result in the restoration of F_1, and its allele recessive gene as rf, then genotypes MS_rfrf expressed male sterility, and other seven genotypes expressed male fertility. In the present study, we have employed two pairs of isogenic lines different in the fertility-related locus Ms–ms in oilseed rape (*B. napus*) to analyze *MS2Bnap* genomic DNA homologous to *MS2* gene from *A. thaliana* and reveal if there exist any difference in *MS2Bnap* gene between two pairs of these lines, and to relate these differences to male sterility. As concerned with the role of the dominant inhibition gene Rf in DDGMS, according to the classic genetic explanation, the product encoded by Rf gene can inhibit the expression of Ms gene, resulting in the restoration of F_1. However, what the product of Rf gene is and how it reacts in response to Ms gene in the molecular level is not clear. The investigation is underway by using the isogenic lines which differ in the fertility-related locus Rf–rf with microarray technique.

In summary, it could be concluded that the isolated *MS2Bnap* genomic DNA homologous to *MS2* gene in *B. napus* was 2,581 bp in length and contained six introns. The one SNP between 220A-gDNA and 220B-gDNA at nucleotide 1247 of 220A-gDNA with A replaced by C was a missense mutation, which may be the putative male sterility site in 220A. Mutation site of BNMS2PROT (CAA68190.1) encoded by *MS2Bnap* gene in 220A(254) and 6A/6C(584) is different. The similarity of exons between *MS2Bnap* in *B. napus* and

MS2 in *A. thaliana* is higher than that of introns between these two genes.

Acknowledgements This work was partly supported by China Postdoctoral Science Foundation (2004036075), Chinese Chunhui Project (Z2004-1-61007), Scientific Research Project of Key laboratory of Oil Crop Genetics and Improvement of Chinese Ministry of Agriculture (Wuhan, China), The Project for the Excellent Young Scientists of Northwest A&F University (04ZR006) and The Program for Excellent Young Talents in Northwest A&F University. Rapeseed DDGMS line 6CA was kindly provided by Mr Shulin Li at Shanghai Academy of Agricultural Sciences, Shanghai, China.

References

Aarts MGM, Dirkse WG, Stiekema WJ, Pereira A (1993) Transposon tagging of a male sterility gene in *Arabidopsis*. Nature 363(24):715–717

Aarts MGM, Hodge R, Kalantidis K, Florack D, Wilson ZA, Mulligan BJ, Stiekema WJ, Scott R, Pereira A (1997) The *Arabidopsis* MALE STERILITY 2 protein shares similarity with reductase in elongation/condensation complexes. Plant J 12(3):615–623

Bancroft I, O'Neill C (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. alboglabra that are homoeologous to sequenced regions of chromosome 4 and 5 of *Arabidopsis thaliana*. In: Plant and animal genome 7th conference, Town & Country Hotel, San Diego, 9–12 January

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, Richard W, Charles D, 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(2):262–272

Cavell AC, Lydiate DJ, Parkin IAP, Dean C, Trick M (1998) Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. Genome 41(1):62–69

Chen FX, Hu BC, Li C, Li QS, Chen WS, Zhang ML (1998) Genetic studies on GMS in *Brassica napus* L. I. Inheritance of recessive GMS line 9012A. Acta Agron Sin 24(4):431–438

Chowdhury JB, Das K (1967) Male sterility in Brown sarson. Indian J Genet Plant Breed 27:284–288

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

Das K, Pandey PD (1961) Male sterility in brown sarson. Indian J Genet Plant Breed 21:185–190

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(6):588–594

Dong YL, Du H (1993) Study on a dominant genic male sterility discovered in *Brassica napus* L. Southwest China J Agric Sci 6(4):6–10

Fourmann M, Barret P, Froger N, Baron C, Charlot F, Delourme, Brunel D (2002) From *Arabidopsis thaliana* to *Brassica napus*: development of amplified consensus genetic markers (ACGM) for construction of a gene map. Theor Appl Genet 105:1196–1206

Fu TD (2000) Breeding and utilization of rapeseed hybrid. Hubei Science and Technology, Wuhan, China

Gan Y, Zeng FY, Zhao Y, Zhang YZ (1999) A preliminary study on RAPD markers linked to sterile gene locus of GMS line 79.7 in *Brassica napus*. Southwest China J Agric Sci 12S(Issue on AHTS):111–115

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

Hodge R, Paul W, Draper J, Scott R (1992) Cold-plaque screening: a simple technique for the isolation of low abundance, differentially expressed transcripts from conventional cDNA libraries. Plant J 2:257–260

Hou GZ, Wang H, Zhang RM (1990) Genetic study on genic male sterility (GMS) material 117A in *Brassica napus*. Oil Crops China 12(2):7–10

Hu SW (2003) Studies on the inheritance of the newly discovered genic male sterility accession "Shaan-GMS" and molecular mechanism of its genic male sterility in *Brassica napus*. Dissertation for Doctor's Degree, Northwest A&F University, Yangling

Hu SW, Liu SY, Yu CY, Guo XL, Zhao HX, Hu XJ, Liu YY (2003) Identification of RAPD markers linked to dominant genic male sterile gene in 220AB derived from Shaan-GMS (*Brassica napus* L.). Chin J Oil Crop Sci 25(3):5–7

Jiang LC, Pu XB, Wang R, Zhang QX, Cai PZ (2000) A primary study of genetic male sterile gene RAPD markers in *Brassica napus* L. Chin J Oil Crop Sci 22(2):1–4

Ke LP, Sun YQ, Hong DF, Liu PW, Yang GS (2005) Identification of AFLP markers linked to one recessive genic male sterility gene in oilseed rape, *Brassica napus*. Plant Breed 124(4):367–370

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(4):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(3):315–325

Lagercrantz U, Putterill J, Coupland G, Lydiate D (1996) Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. Plant J 9(1):13–20

Li SL, Qian YX, Wu ZH (1985) Inheritance of genic male sterility in *Brassica napus* and its utilization to commercial production. Acta Agric Shanghai 1(2):1–12

Li DM, Hou L, Luo XY, Pei Y, Yang GW (2002) Cloning and sequence analysis of fragment homologous to *MS2Bnap* gene in rapeseed recessive genic male sterile line S45AB (*Brassica napus* L). Acta Agron Sin 28(1):1–5

Li G, Gao M, Yang B, Quiros CF (2003) Gene for gene alignment between the *Brassica* and *Arabidopsis* genomes by direct transcriptome mapping. Theor Appl Genet 107:168–181

Lu GY, Yang GS, Fu TD (2004) Molecular mapping of a dominant genic male sterility gene Ms in rapeseed (*Brassica napus*). Plant Breed 123(3):262–265

Luo HY, He CP, Huang ZS, Wei ZF, Huang W (1992) Studies on breeding of hybrid rape I. Selection of NMS double-use line of S445. J Guizhou Agric Sci 20(5):1–4

Mathias R (1985) A new dominant gene for male sterility in rapeseed *Brassica napus* L. Z Pflanzenzuchtg 94(2):170–173

Molhoj M, Johansen B, Ulvskov P, Borkhardt B (2001) Expression of a membrane-anchored endo-1,4-β-glucanase from

Brassica napus, orthologous to *KOR* from *Arabidopsis thaliana*, is inversely correlated to elongation in light-grown plants. Plant Mol Biol 45:93–105

Pan T, Zhao Y (1990) A study of the breeding of double-low GMS line in rapeseed (*B. napus* L.). Crop Res 4(3):47–49

Rao MK, Devi KU, Arundhati A (1990) Application of genic male sterility in plant breeding. Plant Breed 105:1–25

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphism in barley. Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Sillito D, Parkin IAP, Mayerhofer R, Lydiate DJ, Good AG (2000) *Arabidopsis thaliana*: a source of candidate disease-resistance genes for *Brassica napus*. Genome 43(3):452–460

Sun CC, Fang GH, Zhao H, Wang WR, Qian XF, Li YL (1997) Analysis of genotype of a recessive genic male sterile two-type line 22118AB in *Brassica napus* L. and its utilization. Acta Agric Shanghai 13(1):11–15

Sun CC, Zhao H, Wang WR, Li YL, Qian XF, Fang GH (2002) Inheritance and utilization of recessive genic male sterile two-type line 20118A in *Brassica napus* L. Chin J Oil Crop Sci 24(4):1–4

Takagi Y (1970) Monogenic recessive male sterility in rapeseed (*Brassica napus* L.) induced by gamma irradiation. Z Pflanzenzuchtg 64:242–247

Tu JX, Fu TD, Zheng YL, Yang GS, Ma CZ, Yang YN (1999) Studies on the recessive genic male sterility (GMS) and its genetic markers in rapeseed (*Brassica napus* L.). DNA molecular confirmation for P6-9 purple stem gene linked to male fertility gene of GMS. Acta Agron Sin 25(6):669–673

Wang TQ, Tian ZP, Huang ZS, Wei ZF, Shao MB (1999) Selection and breeding of dominant male sterile line with low erucic and glucosinolate content (Qianyou 2AB) in *B. napus*. Guizhou Agric Sci 27(5):14–18

Wang H, Tang XH, Zhao JX (2001) Genetic study on ecotype genetic male sterile of H90S in *Brassica napus* L. Chin J Oil Crop Sci 23(4):11–15

Wang DJ, Tian JH, Wang H, Hu XP, Li DR, Guo AG (2003) RAPD markers linked to mono-dominant GMS in rapeseed. Acta Bot Boreal-Occident Sin 23(9):1556–1560

Zuberi MI, Zuberi S (1983) Inheritance of male sterility in *Brassica campestris* var. toria. Indian J Genet Plant Breed 43:438–440