Collinearity between AFLP markers flanking dominant genic male sterility gene in *Brassica napus* and homologous loci in *Arabidopsis thaliana*

SONG Laiqiang[1,2], FU Tingdong[2], MA Caozhi[2], ZOU Xiaofen[1]

[1]*Jiangxi Academy of Agricultural Sciences, Nanchang 330200, Jiangxi, China;*
[2]*National Key Laboratory of Crop Genetic Improvement, Wuhan 430070, Hubei, China Email: songlaiqiang@yahoo.com.cn*

Abstract

Two patterns for restoration of dominant genic male sterility (DGMS) in *Brassica napus* have been proposed: an unlinked dominant gene or a dominant allele at the very locus of DGMS gene (Ms) restores the male fertility. The later cannot even exclude the possibility of a tightly linked gene or pseudoallele. Therefore, there will not a final explanation until the cloning and function analysis of relevant genes are achieved. *Arabidopsis thaliana* is a model plant with a complete sequenced genome. *B. napus* is closely related to it, and both belong to Cruciferae. Moreover, comparative genome analysis has revealed wide collinear relationship between them and conservativity of gene organization during evolution. The present study was focused on collinearity analysis between flanking markers tightly-linked to the Ms and correspondent homologous loci in *A. thaliana* so as to facilitate fine mapping and cloning. Seven AFLP marker fragments were cloned and sequenced, and the flanking regions of 3 small size fragments were obtained by PCR walking. BLAST search results showed that 6 of the 7 marker sequences had their corresponding homologous loci in *A. thaliana*, and that all these loci were located in a 2Mb region between 1g04950 and 1g10400. Collinearity was evident between the Ms markers and the corresponding homologous loci. Combined with our studies on inheritance and the previous studies on comparative mapping between Brassica A genome and *A. thalian* genome, the Ms was inferred to be likely on the short arm of Chromosome 1, the long arm of Chromosome 5, or the long arm of Chromosome 8 in Brassica A genome.

Key words: *Brassica napus*, dominant genic male sterility, AFLP, *Arabidopsis thaliana*, comparative genome

Introduction

Dominant genic male sterility (DGMS) in *Brassica napus* has been demonstrated to be controlled by two unlinked dominant genes (Li et al.,1985). However, recent researches on two sterile lines 609AB and Rs1046AB indicate that DGMS fertility restorer gene (Mf) is another dominant allele at the DGMS gene (simplified as Ms) locus (Song et al., 2005; 2006a;2006b). Only with the cloning and functional analysis of the relevant gene can the mechanism of fertility restoration be finally elucidated, for a tightly-linked gene or pseudoallele confering the fertility restoration can not be exluded even in the case of multiple allele inheritance. As a first step toward cloning of the Ms, Lu et al (2004a; 2004b) primarily mapped the Ms and identified 5 amplified fragment length polymorphism (AFLP) markers tightly linked to the Ms, of which the two nearest markers on either side of the Ms are 3.7cM and 5.9cM away, respectively. Song et al (2006b) identifies 8 AFLP markers closely linked to the Ms, of which the two nearest markers are at distances of 0.3 cM and 1.6 cM on either side. Based on the different sequence information between two parents, an Ms marker has been successfully converted into a restorer marker. But all these efforts have failed to assign the Ms on a certain chromosome or an accepted genetic linkage group. *Arabidopsis thaliana* serving as a model plant for comparative genome investigation with Brassica species possesses many advantages including availability of complete genome sequence, small genome size and close relation to Brassica. Furthermore, comparative genome analysis has disclosed wide collinear relationship between them and conservativity in gene structure and content during evolution. The present study was an attempt to provide valuable information for further fine positioning and cloning of the Ms by analyzing collinearity between flanking markers tightly-linked to the Ms in *B. napus a*nd correspondent homologous loci in *A. thaliana.*

Materials and methods

Materials

Plant materials used in this study included a dominant genic male sterility line 609AB, a temporary maintainer GS2467, and two doubled haploid (DH) lines: S148, a sterile line derived from 609AB, and S467 derived from GS2467.

Population construction: a backcross population designated as Popu1 was constructed by crossing 609A (sterile plants in 609AB) with GS2467 and further backcrossings were finished with GS2467 as recurrent parent. This population was used to develop AFLP markers. Another backcross population designated as Popu2 was produced in a similar way with DH parents instead. Plants of Popu2 were first grown in plastic culture plates with safe soil medium and were then transplanted to a field when plants had 4-5 leaf age. The fertility was determined during flowing (Song et al., 2006b).

Methods

DNA preparing: the individuals of the parents and the population were numbered and tagged, and the young leaves of each individual were collected and stored at -20°C. A modified CTAB method (Doyle and Doyle, 1990) was used to extract DNA from each individual. The measurement and adjustment of DNA concentration was performed as described by Liu et al (2005).

Popu1 was employed to develop Ms markers with AFLP technique combined with bulked segregant analysis. The developed markers were confirmed in Popu2 and mapped around Ms (Song et al., 2006).The recovering, cloning and sequencing of AFLP fragments was carried out following the procedure described by Ke et al (2005) and Song et al (2006b). PCR walking was conducted as described by Sibert et al (1995) and Song et al (2006b) with the related primers and enzymes listed in Table 1.

Table 1 Gene-specifc primers and enzymes for PCR walking

Marker	Primer	Sequence(5′—3′)	Enzyme	Length stretched
SA12MG14	W6F1-3	ATAAACACCGGCTCAAGATCATCGAA	*Pvu*II	651
	W6F2-3	GTTGAATATGATACCTTGACTCTGACT		
SA13MG06	W4F1	TTGAGTTTGCCACTGAATAGAGAAGAG	*Dra*I	830
	W4F2	ACCGTCTTCATCAACTCATCTCTCTGA		
	W4R1	TCAGAGAGATGAGTTGATGAAGACGGT	*Dra*I	427
	W4R2	TCTATTCAGTGGCAAACTCAACATCCT		

Note: F: Forward direction; R: Reverse direction

Sequences of AFLP fragments or extended AFLP fragments with PCR walking were blasted against the National Center for Biotechnology Information (NCBI) database. Based on the 2 or 3 highest homologous sequences in *A. thaliana* obtained, a homologous locus for each AFLP marker was determined. The homologous loci for all the markers constituted a marker homeologous region. Comparison of this region with the marker linkage group was made to analyze the collinearity between flanking region of the Ms in *Brassica napus* and corresponding homologous loci in *A. thaliana.*

Results

Cloning and PCR walking of AFLP fragments

Eight AFLP markers were first developed from Popu1 with AFLP procedure combined with bulked segregant analysis method. These markers were then subjected to recovery and cloning, of which seven AFLP markers EA07MG12, EA09MC11, SA13MG06, P03MC14, P04MC11, SA12MG14 and P05MG15 were successfully recovered, cloned and sequenced. Sequence lengths of these markers were 81bp, 241bp, 74bp, 407bp, 210bp, 182bp and 78bp, respectively. The PCR walking of markers SA12MG14 and P05MG15 had accomplished and reported before. The marker SA12MG14 was extended on one side by 362bp, while the marker P05MG15 was extended on either side by 797bp and 998bp. In this study, besides the other side of the marker SA12MG14 was extended by 651bp, the marker SA13MG06, which is only 74bp in length but near to the Ms, was subjected PCR walking with the gene-specific primers in table 1. The flanking regions of 830bp and 427bp on either side were obtained.

Blast search for homologous sequences for markers

The sequences of AFLP fragments or extended AFLP fragments obtained by PCR walking were presented to the blast search for homologous sequences in *A. thaliana* against the National Center for Biotechnology Information (NCBI) database. The blast of the EA07MG12 did not result in any hits in *A. thaliana*. The reason for this might be that the fragment was too small to reflect the entire sequence information, or that the fragment was a specific one in *B. napus*. The BLAST searches of the markers EA09MC11, SA13MC06, P03MC14, P04MC11, SA12MG14 and P05MG14 found homologous sequences in *A. thaliana* or in *B. rapa*. One of the hit sequences of the extended SA13MC06 fragment was a 217bp region in Clone KBrH138P04 located on cytogenetical Chromosome 1 in *Brassica rapa (subsp. Pekinensis),*with score of 176, indentity of 85% and expectation value of 1e-40. One of the hit sequences of the extended fragment of marker P05MG14 was a 1777bp region in Clone KBrB037O12 with score 1558, identities of more than 98% and e-value of 0. This was consistent with the testcross result that there exist fertility restorers in *B. rapa*.

Collinearity analysis between Ms marker loci and homologous loci in A. thaliana

All the homologous loci in *A. thaliana* (Table 2) corresponding to the 6 Ms markers resided on Chromosome 1 within a region of 2Mbp between loci At1g04950 (1403.6kbp) and At1g10400 (3415kbp). Figure 1 displayed the collinearity between a 10.7cM flanking region of the Ms and the homeologous region in *A. thaliana* Chromosome 1. This collinear relation was not absolute. The homologous loci corresponding to the two co-segregating markers EA09MC11 and P04MC11 were at a distance of 321kbp, and the homologous loci corresponding to the two co-segregating markers SA13MC06 and SA12MG14 were 1127kbp away. The markers P05MG15 and SA12MG14 separated at a distance of 1.9cM, but the two homologous loci were only 220kbp away from each other, much shorter than the expected value of 357.2kbp calculated based on an average of 188kbp/cM. These phenomena were likely caused by the events of chromosome deletion, insertion or inversion during evulution besides the resolving power of the genetic linkage map.

Table2 Homologous loci in *Arabidopsis thaliana* corresponding to Ms markers

AFLP markers	Homologous loci	AFLP markers	Homologous loci
EA09MC11	At1g06560	P03MC14	At1g04950
SA13MC06	At1g07450	P04MC11	At1g05640
SA12MG14	At1g10400	P05MG15	At1g09645

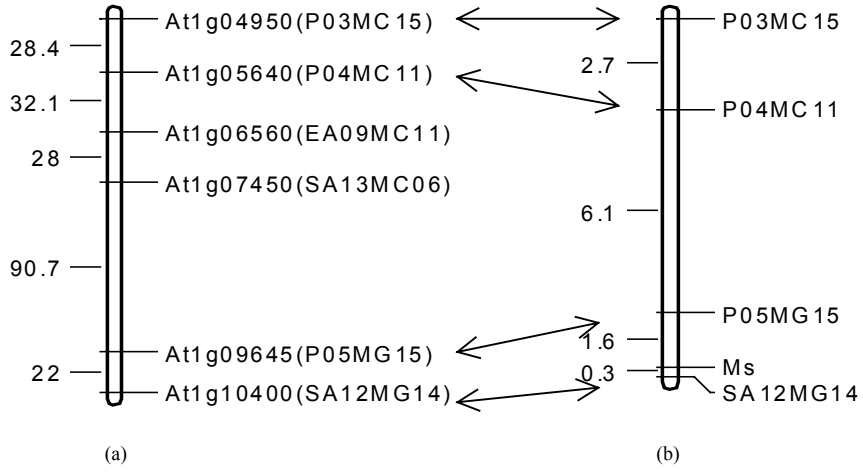

Fig.1 Collinearity analysis between the Ms marker loci and the corresponding homologous loci on chromosome 1 in *Arabidopsis thaliana*. (a) Physical map displays the related loci on Chromosome 1 in *A. thaliana* homologous to the Ms (×10kbp); (b) Linkage map shows the markers linked to Ms(cM)

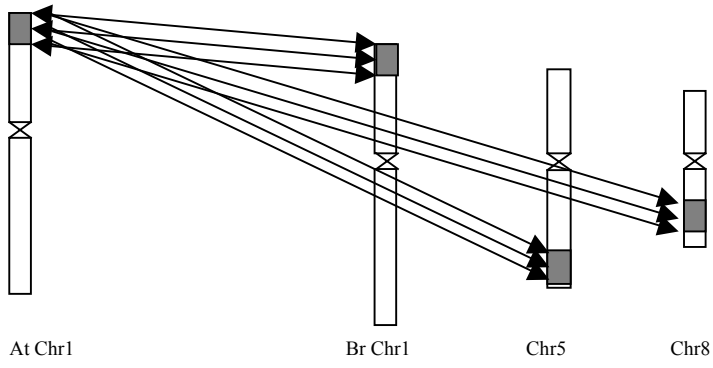

Fig. 2 Corresponding relations between the marker homologous region in *A. thaliana* and those on Chromosome 1, 5 and 8 of *B. rapa* (After the chart from website www.brassica-rapa.org, Yang Tae-Jin)

Discussion

Studies show that there are still large conservative segments in Brassica genomes during evolution after their progenitor differentiated from *A. thaliana*, and that these segments remain high collinearity among related species, which is the foundation of genome comparing and gene cloning based on the collinear relationships. However, the cloning of restorer gene for oug cytoplasmic male sterility (cms) reveals that there is no PPR (Pentatricopeptide Repeat) motif in the collinear region in *A. thaliana* (Brown et al.,2003). A similar PPR domain protein gene to the map-based cloned restorer gene Rf1 for cms in *Sorghum bicolor* is not found in the collinear region on Chromosome 12 in *Oryza sativa*(Klein et al.,2005). Therefore, collinear relationships are relative. Still, analysis of collinearity is helpful for marker conversion, new marker development and gene cloning.

Mapping with common probes or markers among different species is a principal approach in comparative genomic researches, especially in the case that the complete sequence of the target genome is not available. But under the condition of available complete genome sequences, with marker sequences we can apply Internet to undertake sequence-based comparison of genome organization and gene arrangement among species. In the practice of Multinational *Brassica rapa* Sequencing Project, South Korea researchers acquire much important information about the genome through comparing Bacteria BAC end sequences (BES) with *A. thaliana* genome sequence(www.Brassica-rapa.org). In present study, the marker sequences were utilized to blast for homologous sequnces in *A. thaliana* or in *B. rapa*. Comparing marker loci with the homologous loci resulted in distinct collinearity between a 10.7cM region flanking Ms and a 2Mb region on Chromosome 1 of *A. Thaliana*, which was likely disrupted by chromosome inversion, insertion or deletion. Moreover, the homologous loci corresponding to the Ms markers developed in Popu1 close to the homologous loci corresponding to an Ms marker and an restorer marker

developed in Rs1046AB by Lu (2004), together taking up a 3.5Mb region on Chromosome 1 of *A. thaliana* (unpublished data). This corresponding relationship was consistent to the genetical analysis results that the two Ms genes from 609AB and Rs1046AB are allelic to each other, and that the Ms is allelic to the Mf (Song et al., 2005;2006a)

Based only on the corresponding relationship and two homologous sequences in *B. rapa* obtained in the BLAST search, we cannot ascertain that the Ms gene locates on Chromosome 1 of Brassica A genome. According to the comparative analysis of Chromosome 1 of *A. thaliana* and Brassica A genome, the region on Chromosome 1 of *A. thaliana* homologous to the Ms markers is corresponding to the short arm on Chromosome 1, the long arm on Chromosome 5, or the long arm on Chromosome 8. These three chromosomes have been assigned to the linkage groups N9, N5 and N8 by Parkin et al (1995), and to the Chinese cabbage linkage groups LG6, LG4 and LG7 by Lim Y P (www. brassica-rapa.org). These corresponding relationships facilitate to assignment of the Ms on a certain chromosome and to further fine mapping.

Conclusion

There exists distinct collinearity between the Ms flanking region of 10.7 cM and a 2Mb region of homeologous counterpart on *A. thaliana* Chromosome 1. Two homologous sequences are found in *B. rapa* genome, one on the genetical Chromosome 1, the other on unknown location but with high identity. The position of the Ms is further limited.

References

Brown G G, Formanova N, Jin H, Wargachuk R, Dendy C, Patil P, Laforest M, Zhang J, Cheung W Y, Landry B S(2003) The radish Rfo restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J, 35(2): 262-72.

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus12: 13–15

Klein R R, Klein P E, Mullet J E, Minx P, Rooney W L, Schertz K F (2005) Fertility restorer locus Rf1of sorghum (Sorghum bicolor L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. Theor Appl Genet, 111: 994–1012

Ke LP, Sun YQ, Liu PW, Yang GS (2004) Identification of AFLP fragments linked to one recessive genic male sterility (RGMS) in rapeseed (*Brassica napus* L.) and conversion to SCAR markers for marker-aided selection. Euphytica 138: 163–168.

Li SL,Qian YX, Wu ZH (1985) Inheritance and utilization of genetic male sterility in rapeseed(*Brassica napus* L.). Acta Agriculturae Shanghai 1(2): 1-12

Liu ZW, Fu TD, Tu JX, Chen BY (2005) Inheritance of seed colour and identification of RAPD and AFLP markers linked to the seed colour gene in rapeseed (*Brassica napus* L.). Theor Appl Genet 110: 303-310

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

Lu GY, Yang GS, Fu TD (2004b) Linkage map construction and mapping of a dominant genic male sterility gene (Ms) in *Brassica napus*. Acta Genetica Sinica 31 (11) : 1309～1315

Parkin I A P, Sharpe A G, Keith D J, Lydiate D J (1995). Identification of the A and C genomes of amphidiploid Brassica napus (oilseed rape). Genome, 38: 1122-1131

Siebert PD, Chenchick A, Kellogg DE, Lukyanov KA, Lukyanov SA (1995) An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res 23: 1087–1088

Song LQ, Fu TD, Yang GS, Tu JX, Ma CZ (2005) Genetic Verification of Multiple Allelic Gene for Dominant Genic Male Sterility in 609AB (*Brassica napus L.*). Acta Agronomica Sinica 31(7): 869-875

Song LQ, Fu TD, Yang GS, Tu JX, Ma CZ (2006a) Allelism analysis of dominant genic male sterility gene and its restorer gene in Brassica napus. Scientia Agricultura Sinica, 39(3): 456-462

Song LQ, Fu TD, Tu JX, Ma CZ, Yang GS (2006b) Molecular validation of multiple allele inheritance for dominant genic male sterility gene in Brassica napus L.Theor Appl Genet, 113(1), 55- 62