Theor Appl Genet (2010) 120:843–851
DOI 10.1007/s00122-009-1215-y

ORIGINAL PAPER

High-resolution mapping of the *Brassica napus Rfp* restorer locus using *Arabidopsis*-derived molecular markers

Nataša Formanová · Rachel Stollar · Rachel Geddy · Laetitia Mahé · Martin Laforest · Benoit S. Landry · Gregory G. Brown

Received: 11 September 2009 / Accepted: 30 October 2009 / Published online: 18 November 2009
© Springer-Verlag 2009

Abstract The two forms of cytoplasmic male sterility (CMS) native to the oilseed rape or canola species *Brassica napus, nap* and *pol,* have novel features that may provide insight into the molecular mechanisms through which CMS/nuclear restorer systems evolve. One such feature is the finding that the distinct nuclear restorer genes for the two systems represent different alleles or haplotypes of the same nuclear locus. Improved understanding of how these systems have evolved will require molecular cloning and characterization of this novel locus. We have employed an approach that exploits the regional co-linearity between the *Arabidopsis* and *Brassica* genomes to construct a high-resolution genetic map of the nuclear restorer for the *pol* system, *Rfp*. Specifically, *Arabidopsis*-derived sequences have been used as a set of ordered RFLP probes to localize *Rfp* to a region of the *B. napus* genome equivalent to a 115 kb interval on *Arabidopsis* chromosome 1. Based on the known relationship of physical distances between orthologous segments of *Arabidopsis* and *Brassica* chromosomes, it is anticipated that the *B. napus* restorer locus is now mapped to sufficient resolution to permit its isolation and characterization.

Communicated by C. Quiros.

N. Formanová · R. Stollar · R. Geddy · L. Mahé · G. G. Brown (✉)
Department of Biology, McGill University,
Montreal, QC, Canada
e-mail: greg.brown@mcgill.ca

M. Laforest · B. S. Landry · G. G. Brown
DNA LandMarks, St-Jean-sur-Richelieu, QC, Canada

Present Address:
N. Formanová
Bayer BioScience N.V., Technologipark 38,
9052 Ghent, Belgium

Present Address:
R. Geddy
Environment and Energy Division,
Department of Foreign Affairs and International Trade Canada,
111 Sussex Drive, Ottawa, ON K1N 1J1, Canada

Present Address:
L. Mahé
Agropolis Fondation, 34394 Montpellier Cedex 5, France

Present Address:
B. S. Landry
Percival Consultants, 162 des tournesols,
St-Jean-sur-Richelieu, QC J2W 1H9, Canada

Introduction

Cytoplasmic male sterility (CMS) is a widespread trait among higher plants that results from the expression of novel open reading frames (ORFs) in the mitochondrial genome (Hanson and Bentolila 2004; Pelletier and Budar 2007). For many examples of CMS, nuclear genes termed restorers-of-fertility (*Rf*) have been identified that can suppress the CMS phenotype. In general, such nuclear restorer genes are specific for a particular form of CMS and function by down-regulating the expression of the novel, CMS-associated ORF (Chase 2007; Schmitz-Linneweber and Small 2008). CMS thus provides a model for studying interactions between the nuclear and mitochondrial genomes, in addition to being a practical tool for facilitated hybrid seed production. Moreover, it has attracted interest from evolutionary biologists because CMS/restorer gene systems are thought to represent a type of "intra-genomic conflict". Maternally inherited CMS-conferring mtDNA

spreads in a population due to the selective advantage conferred by not allocating resources into pollen production; as CMS individuals become more frequent, however, pollen becomes limiting, creating the selective pressure for the evolution of new restorer genes (Budar et al. 2003; Touzet and Budar 2004).

CMS is widely used for hybrid seed production in the oilseed rape species *Brassica napus*. The two forms of CMS that are native to *B. napus*, *pol* and *nap*, have novel properties that promise to provide insight into the molecular mechanisms through which CMS/nuclear restorer systems evolve. The *pol* and *nap* CMS-associated ORFs, *orf224* and *orf222*, respectively, possess a high degree of sequence similarity, suggesting that they evolved relatively recently from a common ancestor (L'Homme et al. 1997; Brown 1999). In addition, the nuclear restorer genes for the two systems, *Rfp* (for *pol*) and *Rfn,* map to the same nuclear locus, again suggesting a recent evolutionary origin (Li et al. 1998). Both *Rfp* and *Rfn* act by specifying RNA processing events that decrease levels of their respective cognate transcripts, but differ in their specificity: *Rfn* does not act on *orf224* transcripts and *Rfp* does not act on *orf222* transcripts (Li et al. 1998; Menassa et al. 1999; Elina 2007). Thus, *Rfp* and *Rfn* represent different alleles or haplotypes of a single nuclear locus that specify distinct mtRNA processing events related to the down-regulation of their respective CMS-associated genes.

Additional insight into the evolution of CMS in *B. napus* will require the identification and characterization of the *Rfn/Rfp* locus. Nuclear restorer genes that have been identified for other systems and species have, for the most part, been found to encode P-type pentatricopeptide repeat (PPR) proteins (Chase 2007; Schmitz-Linneweber and Small 2008). Such proteins are characterized by tandem copies of a canonical 35-amino acid repeat and are thought to specify the sites of transcript modification events such as nuclease processing and editing by binding to specific sequences on RNA molecules (Schmitz-Linneweber and Small 2008). It thus seems likely that *Rfn* and *Rfp* will be found to encode related PPR proteins that act to down-regulate expression of their corresponding CMS-associated ORFs by binding to and mediating processing events on different sets of mitochondrial transcripts.

Rfn and *Rfp* have previously been mapped to *B. napus* linkage group 18 (Landry et al. 1991) using RFLP markers (Jean et al. 1997; Li et al. 1998). One such marker, cRF1, was found to be completely linked to *Rfp* in three populations, together comprising 195 individuals. This marker was used to recover cosmid clones from a *B. rapa* line into which *Rfp* had been introgressed. Several of these were assembled into a contig that spanned the cRF1-linked polymorphism (Formanová et al. 2006). The sequences in this contig were found to correspond to a region near the 4.6 Mb coordinate of *Arabidopsis* chromosome 1 (Formanová et al. 2006).

We here describe a strategy that has allowed us to more precisely map the *Rfp* locus by exploiting regional co-linearity between the *B. napus* and *Arabidopsis* genomes (Parkin et al. 2005). *Rfp* has been localized to a *B. napus* chromosomal region that corresponds to an *Arabidopsis* region 115 kb in length. Given that physical distances in *Brassica* genomes are on average, <1.5 times those of the *Arabidopsis* genome (O'Neill and Bancroft 2000), it is likely that *Rfp* has now been mapped to sufficient resolution to permit its cloning and characterization.

Materials and methods

Plant material and fertility scoring

BC1 mapping populations segregating at the *Rfp* locus derived from crosses between *pol* CMS 'Karat' *(rfp/rfp;pol)* or *pol* CMS Westar *(rfp/rfp;pol)* and the *pol* restorer 'Westar-Rf' *(Rfp/Rfp;pol)* (the KW and WW mapping populations, respectively), have been described (Jean et al. 1997). To generate a larger mapping population, a single F1 individual (*rfp/Rfp*; *pol*; Karat × Westar-Rf) was crossed to three sterile *pol* CMS Karat plants (*rfp/rfp; pol*). Plants were grown to maturity in growth chambers under standard conditions (16-h photoperiod, 22/16°C, day/night temperatures). One thousand seeds recovered from these plants were planted and raised to seedling stage under the same conditions. Genomic DNA was isolated according to Zhang and Stewart (2000) from single leaf discs taken at the two to four leaf stage, and scored for markers 14o and 4.4BB, as described below. Plants in which recombination between the markers had taken place were grown to maturity and scored for fertility. Male fertility was assessed by the careful observation of five flowers per plant at least two times during the flowering period. Flower morphology was used as an additional criterion to aid in the assessment of fertility of plants segregating for the restoration of the *pol* CMS, since a small amount of pollen can be produced by plants possessing only the maintainer allele *rfp* (Fan et al. 1986; Jean et al. 1997). Flowers from male-fertility restored plants look identical to those of a fertile maintainer plant; flowers from a *pol* CMS plant have shrunken petals, and the style of the pistil is longer and often bent. CMS anthers also have shorter filaments, a poorly developed or absent pollen sac, and no or a reduced amount of released pollen. The morphological contrast between CMS and normal flowers was sufficient to allow plants carrying a restorer allele to be unambiguously distinguished from those with only maintainer alleles.

RFLP analysis

Plant genomic DNA used for RFLP analysis was isolated as described by Landry et al. (1991) with the following modification: purification by CsCl gradient was replaced with two successive chloroform/isoamylalcohol (24:1, v/v) extractions. The cRf1 marker used in the present study had been selected by targeted mapping as being linked to the *Rfp* locus (Jean et al. 1997). Individual DNA fragments were purified from digests of *Arabidopsis* BACs by gel extraction (QIAEX II, Qiagen). These probes were evaluated for their capacity to detect *Rfp*-linked polymorphisms by probing gel blots of EcoRI, BamHI and EcoRV digests of the parental lines and pools of fertile and sterile BC1 progeny of the KW and WW populations. Specifically, digested DNA samples (5µg/lane) were electrophoresed in 0.8% agarose gels and transferred onto Hybond N^+ membranes (Amersham) by the alkaline transfer method recommended by the manufacturer. Probes were radioactively labeled with α-^{32}P-dCTP using RediprimeII (Amersham Pharmacia Biotech Inc) and purified from unincorporated nucleotides by passage through Bio-Gel P-60 (Bio-Rad Laboratories). Conditions for hybridization were as described previously (Landry et al. 1991) with the following modification: the final high-stringency wash was performed at 55 instead of 62°C when *Arabidopsis* probes were used.

AFLP analysis

DNA from the lines 'Karat' (*rfprfp;pol)* and 'Westar-Rf' *(Rfp/Rfp; pol)* as well as bulked DNAs from sterile and fertile BC1 plants of the KW population (15–20 individuals per bulk) were tested using AFLP Analysis System I, AFLP Start Primer Kit (Life Technologies, Gaithersburg, MD), as described by Vos et al. (1995). Briefly, 250 ng of genomic DNA was completely digested with EcoRI/MseI followed by adapter ligation, pre-amplification and amplification according to the manufacturer's recommendations for ^{33}P labeling. Amplified products were resolved on a 6% polyacrylamide denaturing gel, transferred to Whatman paper, dried, and exposed to X-ray film (Kodak X-OMAT, AR) for 24 h. In total, 64 primer combinations were tested.

For isolation of AFLP fragments, dried gels and films were aligned and a selected area was excised from the gel using a razor blade. The gel slice was placed in a 1.5-ml tube and homogenized; 100 µl of extraction buffer (0.5 M ammonium acetate, 5 mM EDTA) was then added, and the tube was left at 4°C for 24 h. The sample was then centrifuged at 14,000 rpm for 5 min, the supernatant transferred to a fresh tube and the DNA was precipitated by ethanol. DNA pellets were dissolved in 10 µl of sterile water, and 1–4 µl was used for a second round of AFLP amplification. Products of the second AFLP reaction were run on a polyacrylamide gel, the band of interest was cut from the gel and the DNA was extracted again as described above. The *Rfp*-linked polymorphic band detected as AFLP 14o was cloned using the TA-cloning kit from Invitrogen.

SCAR genotyping

A PCR reaction was performed on each BC1 plant from the mapping population. Two pairs of primers were used for amplifying both the 14o and 4.4BB markers in the same reaction.

14o marker:
14o1: CTTCTTCGACAGAGAAACGTG
14o4: CAGAAGCTGATTCCTGATGATG.
4.4BB marker:
4R: AAGCTTCCACAGGTCCAGAGTAC
3F: GTTCTTACAGTGGCGTAAAGATGAG.

For each reaction, 25 ng of DNA (or 1/50 of the DNA sample from the micro-extraction of a single 5-mm leaf disc) was used in a 25 µl of a reaction mixture (SuperMix, Gibco). Amplification was started with one cycle of 5 min at 94°C, 30 s at 57°C and 1.5 min at 72°C, followed by 35 cycles of 30 s at 94°C, 30 s at 57°C, 1.5 min at 72°C and 10 min at 72°C. Twenty microliters of each PCR sample was analyzed on a 2% agarose gel.

SNP discovery and analysis

A series of primers corresponding to *Brassica* orthologs of *Arabidopsis* genes located between the 4.25 and 4.45 Mb coordinates of chromosome 1, were used to generate amplicons from sterile and fertile bulked DNAs of the KW population of Jean et al. (1997). Sequence analysis of the products generated using one such primer pair, P9 (F: TGATAAGGAGCATTCGACGAT; r: CACCAGCA GTGTACGCAAGA), which corresponded to *Arabidopsis* gene At1G12910, allowed identification of an *Rfp*-linked single nucleotide polymorphism (SNP). The same primer pair was used to generate amplicons from a limited number of plants of the large backcross segregating population, and these were analyzed using a SNaPshot^{TM} assay with an ABI DNA analyzer.

Analysis of cosmid clones

Cosmids were isolated from the *B. rapa* cosmid library as described (Formanová et al. 2006), using as probes amplification products obtained with P9 primers on *Brassica* DNA. Four micrograms of Qiagen-purified cosmid DNA, in a volume of 300 µl, was sheared down to 5–8 kb using a hydroshear (GeneMachines). The fragmented DNA was subcloned into the pMOSblue cloning vector using the pMOSblue blunt-ended cloning kit (Amersham Biosciences

Corp.) according to the manufacturer's instructions. Ligated DNA was then transformed into Top10 chemically competent cells (Invitrogen). Sequencing and sequence assembly was performed by DNA LandMarks (St-Jean-sur-Richelieu) and Genome Quebec (Montreal) using the Applied Biosystems 3730XL DNA analyzer for capillary sequencing and the Phred/Phrap programs for sequence assembly, as described (Brown et al. 2003).

Results

Targeted mapping of *Rfp* with *Arabidopsis*-derived RFLP markers

The *Rfp* co-segregating marker cRF1, although it provided a genetic and physical anchor to the *Rfp* genomic region, did not allow this region to be genetically defined, since it did not identify recombination events in the populations of Jean et al. (1997). None of the clones from a contig of ~115 kb recovered from a limited chromosome walk initiated from the cRF1 marker (Formanová et al. 2006) gave rise to fertile transgenic plants upon their introduction into *pol* CMS *B. napus* by *Agrobacterium*-mediated transformation, indicating that more precise genetic localization of *Rfp* would be necessary for gene identification. The RFLP markers 5NE12 and 4ND7 found by Jean et al. (1997) to map closest to cRF1/*Rfp* on either side were separated by 15–30 cM, a distance too great to permit their use in positional cloning approaches. We therefore chose to explore a targeted mapping approach relying on the regional synteny or co-linearity between the *Arabidopsis thaliana* and *Brassica* genomes (Parkin et al. 2005) to identify flanking markers mapping closer to *Rfp*.

The sequences of cosmids containing the *Rfp*-linked polymorphic fragment detected by the cRF1 probe were found to show a high degree of similarity to a region of the *A. thaliana* genome in the vicinity of the 4.6-Mb coordinate of chromosome 1 (Formanová et al. 2006). It was reasoned that if synteny was maintained between the *Arabidopsis* genome and the *B. napus Rfp* region, *Brassica* RFLPs detected by *Arabidopsis* sequences surrounding the 4.6 Mb coordinate should also show linkage with *Rfp*. To evaluate this possibility, a set of potentially ordered and linked probes was generated by subcloning restriction fragments from a set of *Arabidopsis* BACs spanning the region from coordinates 4.1 to 4.9 Mb on chromosome 1 (Table 1). These probes were then surveyed for their capacity to detect *Rfp*-linked RFLPs by bulked segregant analysis (Michelmore et al. 1991) of pooled genomic DNAs of 15–20 fertile or sterile individuals from mapping populations of Jean et al. (1997).

Examples of *Rfp*-linked RFLPs detected by this method are shown in Fig. 1. Probe 2.7 HH which is located close to *Arabidopsis* coordinate 4.3 Mb, contains a 2.7 kb *Hin*dIII fragment subcloned from *Arabidopsis* BAC T15L9, and detects an EcoRI polymorphism between the *pol* CMS parent (Karat, P_S) and the nuclear restorer parent (Westar-Rf, P_F) as well as between pooled DNA samples from fertile (B_F) and sterile (B_S) segregants of the KW BC1 population (Jean et al. 1997). Similarly, probe 11.5BB, which maps near *Arabidopsis* coordinate 4.8 Mb, consists of an 11.5 kb BamHI *Arabidopsis* fragment and detects an EcoRI polymorphism between the parents as well as the bulked segregants of this cross. In total, 12 *Rfp*-linked RFLP markers, summarized in Table 1, were identified by this approach, of which 8 were used in further mapping experiments based on their position on the *Arabidopsis* chromosome and the robustness of their performance.

Table 1 *Arabidopsis* restriction fragments detecting *Rfp*-linked polymorphism

Marker	*Arabidopsis* BAC	Polymorphism detected
8.5BB[a]	T14N21	BamHI
6.1BB	T14N21	*Hin*dIII, EcoRI
3.5BB	T32B12	*Hin*dIII
5.3BB	T32B12	EcoRI
9.6BB	T32B12	BamHI
9.5HH	T13M17	EcoRI
2.7HH	T15L9	EcoRI, EcoRV
3.9BB	T15G19	BamHI, EcoRI
7.5BB	T8O4	*Hin*dIII
3.0BB	T9K10	EcoRI
11.5BB	T9K10	EcoRI
10.0BB	T26J15	EcoRV
4.4BB	T31J1	BamHI, *Hin*dIII
2.4BB	T31J1	EcoRI

[a] The marker name indicates the restriction endonuclease used as well as the obtained fragment size. (e.g. 8.5BB is an 8.5 kb BamHI fragment cloned from *Arabidopsis* BAC T14N21, that detects a BamHI polymorphism

To position *Rfp* relative to the various *Arabidopsis*-derived markers, we scored the alleles detected by these markers on a subset of individuals from the original BC1 mapping populations that were known to contain crossovers in the vicinity of *Rfp*. Recombination events on the 5NE12 side of *Rfp* were detected by three *Arabidopsis*-derived markers: 11.5BB, 10.0BB and 4.4BB. The number of crossovers between *Rfp* and each of these markers correlated with their position on *Arabidopsis* chromosome 1 (Fig. 2): marker 11.5BB, positioned near coordinate 4.7 Mb, detected 1 recombination event; marker 10.0BB, located near the 4.75 Mb coordinate, detected 2 events; and marker 4.4BB, located near coordinate 4.8 Mb, detected 3 events. Thus, genetic distance on the *B. napus* genome correlates with physical distance on the *Arabidopsis* genome, consistent with the premise of regional co-linearity. The

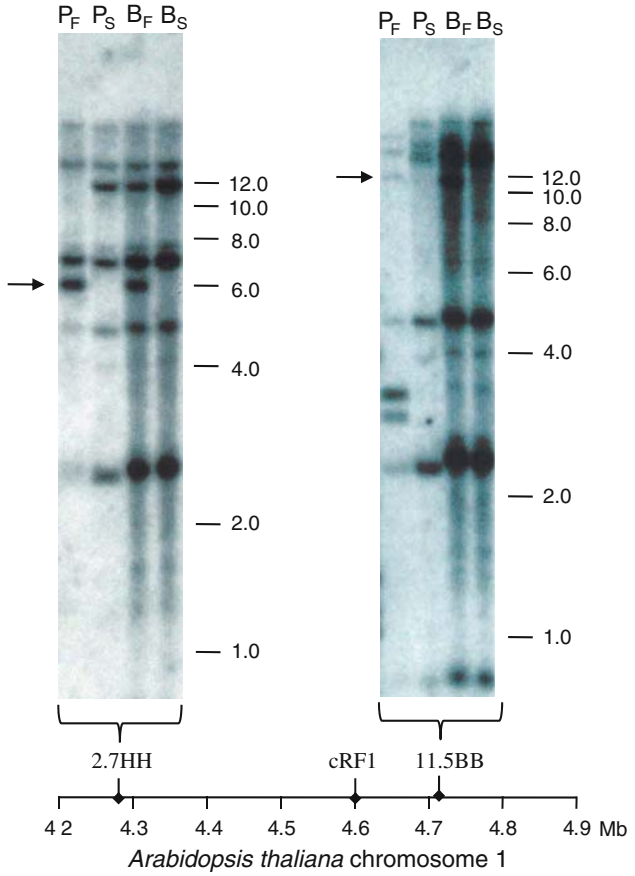

Fig. 1 Detection of *Rfp*-linked RFLPs using *Arabidopsis*-derived probes. *B. napus* DNA from the parental lines Westar-Rf (P_F) and *pol* CMS Karat (P_S) and pooled fertile (B_F) and sterile (B_S) progeny were probed with labeled fragments isolated from *Arabidopsis* BACs (Table 1). The position of the fragments on *Arabidopsis* chromosome 1 sequence is indicated on the *lower bar* (coordinates in Mb). *Left panel* probe 2.7HH detects an *Rfp*-linked 6.0 kb EcoRI fragment, indicated by an *arrow*, that is present in the fertile parent and fertile bulks, but not the sterile parent or sterile bulks. *Right panel* probe 11.5BB detects an *Rfp*-linked 11 kb EcoRI fragment, indicated by an *arrow*, that is present in the fertile parent and bulk but not the sterile parent or bulk. Sizes of molecular weight markers are in kb

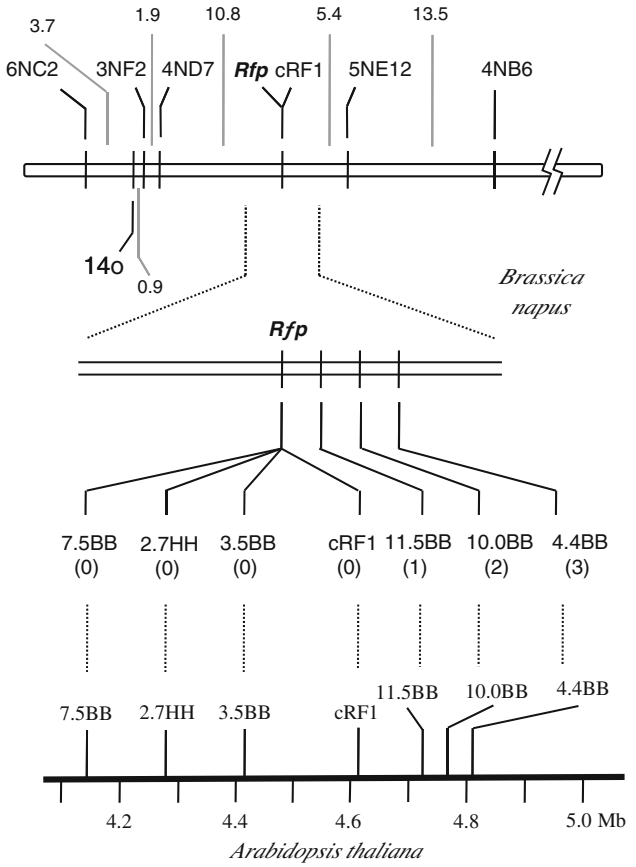

Fig. 2 Coarse mapping of *Arabidopsis*-derived RFLP markers. The *upper portion* of the figure is a *B. napus* genetic map that includes the 14o AFLP marker, the *Rfp* gene, and the anchoring cRF1 marker as well as other RFLP markers as determined from analysis of the KW population in the original mapping study (Jean et al. 1997). The *numbers* associated with each interval indicate genetic distance in centimorgans. The *bar* in the middle represents an expanded segment surrounding the *Rfp* gene where *Arabidopsis*-derived markers were ordered based on their position on *Arabidopsis* chromosome 1. The number of recombination events detected between a marker and the *Rfp* gene in the populations of Jean et al. (1997) is indicated in *parentheses*. The *thick bar* at the bottom represents a corresponding segment of the *Arabidopsis* chromosome with coordinates indicated in megabases (Mb)

RFLP marker 11.5BB genetically delimited one boundary of the *B. napus Rfp* region.

Surprisingly, no recombination event was detected on the opposite side of *Rfp,* despite the fact that four *Arabidopsis*-derived RFLP markers, extending from *Arabidopsis* coordinates 4.15–4.5 Mb, were used. This observation suggested that the frequency of recombination in the region flanking *Rfp* on the 4ND7 side might be low, and that to position *Rfp* to sufficient resolution it would be necessary to analyze a larger mapping population.

Development of PCR-based markers

To efficiently map genes using large populations, it is useful to first develop PCR-based markers that can be used to conduct a high-throughput screen for crossovers in the region of interest. This is followed by fine mapping in which all available markers are positioned on the genetic map of the targeted region. Two approaches were taken to identify PCR-based markers that would flank both sides of the *Rfp* gene. In one case, *Arabidopsis*-derived mapping probes described above were used to screen the *B. rapa* cosmid library of Formanová et al. (2006) to identify *Brassica* genomic clones carrying the *Rfp*-linked allele of the marker. The sequence in the vicinity of the polymorphic site was determined and PCR primers were designed that enabled an amplification of the *rfp*-linked allele and comparison of the two sequences. In a second approach, DNAs from the two parent lines of the KW BC1 population (Jean

et al. 1997) as well as bulked DNA samples from that population were screened for differentiating AFLPs.

Using the first approach it was possible to convert the 4.4BB RFLP marker which maps to the 5NE12 side of *Rfp* into a co-dominant PCR-based marker that detects a length polymorphism between the two alleles. An example of the variation detected by this marker is shown in Fig. 3b. The primer pair 3F and 4R (see "Materials and methods") amplified a band of approximately 600 bp from the sterile parent and a band of approximately 500 bp from the fertile parent. The F1 possessed both bands, as expected, as did the bulk of fertile BC1 individuals; the sterile bulk possessed only the large, sterile-specific band.

Several AFLPs detecting Rfp-linked polymorphisms were identified by bulked segregant analysis of individuals from the KW population. When these were used to analyze a subset of individuals from that population in which recombination in the vicinity of Rfp was known to occur, only one, designated 14o (Fig. 3a) was found to map sufficiently closely to *Rfp* on the 4ND7 side to be useful for screening of a large mapping population. Fourteen recombination events were found to occur in the KW BC1 population (121 plants) between the 14o marker and *Rfp*, indicating the loci are separated by a genetic distance of 13.6 cM. The 14o polymorphic band specific to the fertile line was excised from the gel, cloned and sequenced. A pair of primers designed on the basis of the sequence was found to amplify a band only from fertile plants. This Sequence Characterized Amplified Region (SCAR) marker was found to be useful in combination with the 4.4BB marker for identifying recombinants in a large mapping population (see below).

Fine mapping of *Rfp* with a large backcross population

To develop a large mapping population segregating for *Rfp*, an F1 individual generated by crossing a *pol* CMS plant (CV 'Karat' nuclear genotype) with the nuclear restorer line Westar-Rf (CV 'Westar' with an introgressed Rfp gene) was crossed as male with the recurrent parent, *pol* CMS 'Karat' (*rfp/rfp*; *pol*). Nine hundred and fourteen backcross seedlings were screened for crossovers in the *Rfp* region using the 14o and 4.4BB PCR-based markers (Fig. 3c). One hundred and nine plants that showed recombination between the two markers were grown to maturity and scored for male fertility/sterility. Of these, 57 were scored as male fertile and 52 as male sterile, consistent with the expected 1:1 ratio.

All 109 plants that contained crossovers between 14o and 4.4BB were further analyzed with RFLP markers cRF1, 3.5BB, 2.7HH and 7.5BB (Table 2; Figs. 2, 4). The number of recombination events detected between the markers cRF1, 3.5BB, 2.7HH and the *Rfp* gene, as indicated in Fig. 4 are consistent with the premise of co-linearity between the *Arabidopsis* and *Brassica* genomes in the *Rfp* region. The *Rfp* gene was found to be located between

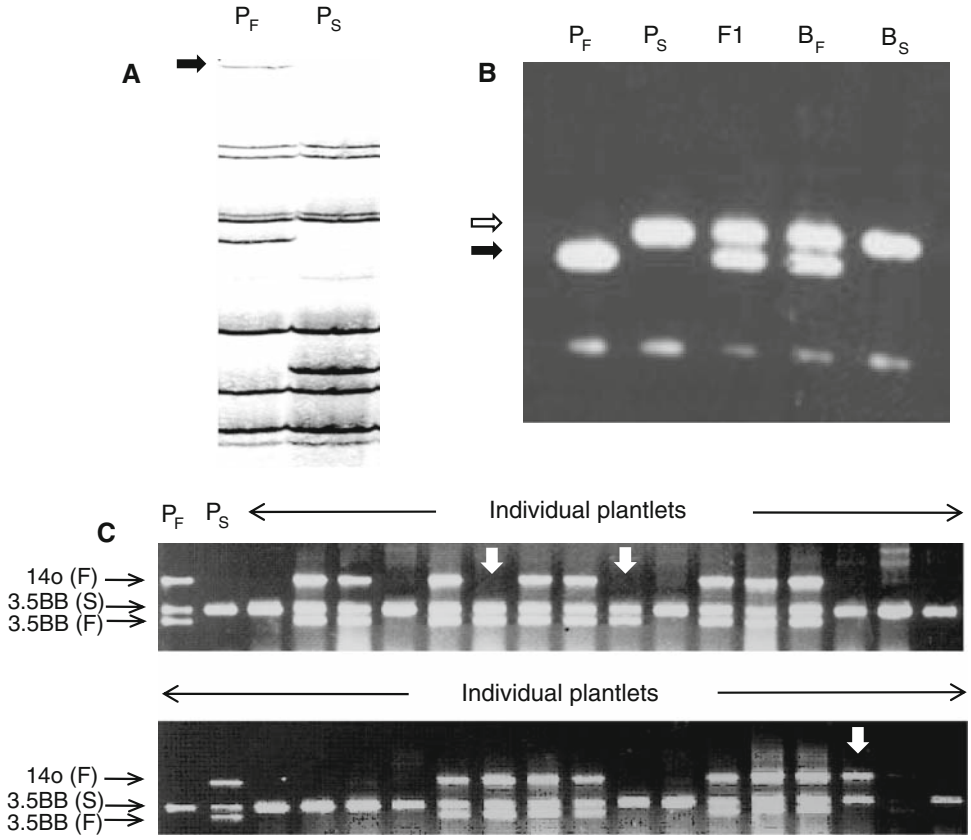

Fig. 3 Development of PCR-based markers for screening of a large mapping population. **a** AFLP marker 14o on fertile (P_F) and sterile (P_S) parental lines; the *filled arrow* indicates the polymorphic band that was subsequently converted into a SCAR marker. **b** 4.4BB polymorphism detected as different size amplicons on fertile (P_F) and sterile (P_S) parental lines, one F1 individual, and fertile and sterile bulks from the BC1 mapping populations (B_F, B_S); the *filled and open arrows* denote polymorphic bands specific to the fertile and sterile genotypes, respectively. **c** Screening of BC1 individuals with SCAR marker 14o and 4.4BB; *white arrows* indicate samples in which recombination was inferred to have taken place within the interval defined by the two markers

Table 2 Marker genotypes of individuals of the large KW population with crossovers in the vicinity of the *Rfp* gene

Plant	Phenotype	2.7HH	Rfp	3.5BB	7.5BB	cRF1	4.4BB
67A	S	A	A	A	A	H	H
195A	F	H	H	H	H	A	A
311A	S	A	A	A	H	H	H
33B	F	H	H	A	A	A	A
209B	F	H	H	H	H	A	A
221B	S	A	A	A	H	H	H
268B	F	A	H	H	H	H	H
313B	F	A	H	H	H	H	H
339B	F	ND	H	H	A	A	A
30C	S	A	A	A	ND	A	H
31C	S	A	A	A	A	A	H
80C	S	A	A	A	A	A	H
97C	S	A	A	A	ND	A	H
161C	S	A	A	A	A	A	H
209C	S	A	A	A	A	A	H
211C	S	A	A	ND	A	A	H

A homozygous for the allele of the sterile parent, *H* heterozygous, *F* fertile, *S* male sterile, *ND* genotype not determined

markers 2.7HH and 3.3BB, which are situated on *Arabidopsis* chromosome 1 at coordinates 4.28 and 4.44 Mb, respectively. Interestingly, the RFLP marker 7.5BB, which, based on its position in the *Arabidopsis* genome, was expected to map between markers 2.7HH and 14o on the 4ND7 side of *Rfp* (Fig. 2), was found instead to map genetically between the markers cRF1 and 3.5BB on the opposite (5NE12) side of *Rfp* (Table 2). Thus, although the 7.5BB marker is clearly linked to *Rfp*, its genetically determined position is not consistent with the premise of continuous co-linearity between the two genomes. This observation suggests that although there is overall synteny in the *Rfp* region between the *Arabidopsis* and *Brassica* genomes, deviations from co-linearity may exist as a result of local chromosomal rearrangements.

To identify additional *Rfp*-linked polymorphisms within the boundary of the region flanked by markers 2.7HH and 3.5BB, we examined *B. napus* EST databases for sequences homologous to those of the corresponding interval on *Arabidopsis* chromosome 1. Primers were then designed that would allow amplification of the *Brassica* orthologs. Upon sequence analysis of the products amplified by one such primer pair from the fertile and sterile bulks, we identified a SNP positioned within the *Brassica* ortholog of *Arabidopsis* gene At1g12910 (chromosome 1 coordinate 4.395 Mb). A single-base extension SNP genotyping assay was devised that allowed genotyping of all individuals with crossovers in the near vicinity of *Rfp*. This SNP marker, '12910', was found to score identically to 3.5BB with only one crossover between itself and the *Rfp* gene. Since At1g12910 is located between 2.7HH and 3.5BB on the *Arabidopsis* genome, it is likely that the *Brassica* ortholog of At1g12910 is now the closest marker flanking *Rfp* on the 5NE12 side. To determine the identity of the sequences sur-

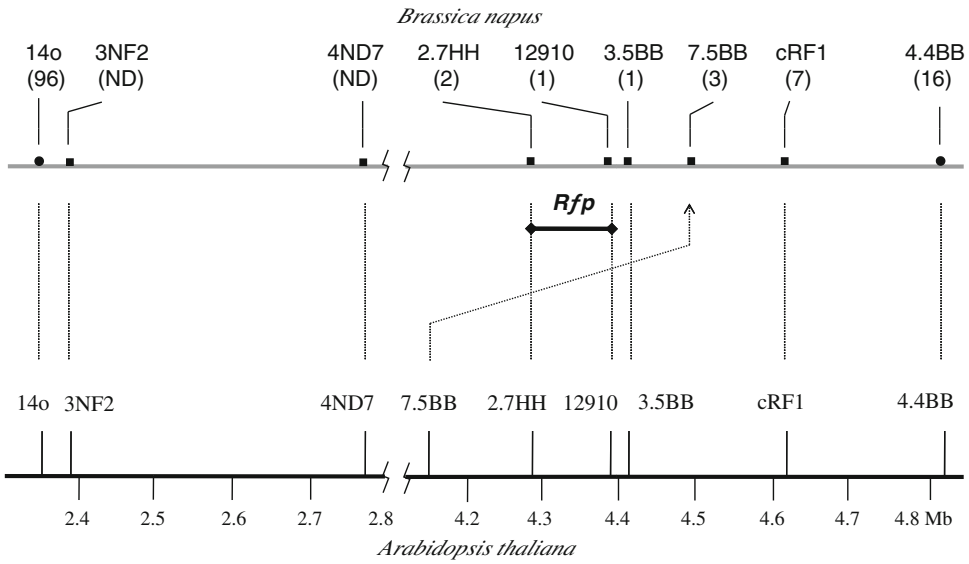

Fig. 4 Fine mapping of the *Rfp* gene with *Arabidopsis*-derived markers. *Upper bar* a schematic representation of the *B. napus Rfp* region; the placing of *Arabidopsis*-derived markers reflects their position on *Arabidopsis* chromosome. *Lower bar* a schematic representation of *Arabidopsis* chromosome 1 with coordinates in Mb. *Figures in parentheses* indicate the number of crossovers observed between the marker and the *Rfp* gene in the large BC1 mapping population. The genetically determined position of marker 7.5BB in *Brassica napus* differs from its location on the *Arabidopsis* chromosome, as indicated by an *arrow*. The marker designated 12910 is a SNP marker derived from a *B. napus* ortholog of the *Arabidopsis* gene At1g12910. The *bold line with filled diamonds* delimits the possible location of the *Rfp* gene

rounding the *Brassica* ortholog of At1g12910, the *B. rapa Rfp* cosmid library was screened with the *Brassica* amplicon containing SNP '12910'. One clone recovered from this screen, designated 2840A3, was found to contain the *Rfp*-linked SNP and thus to be anchored in the *Rfp* coding region. When this clone was sequenced, it was found to contain sequences corresponding to *Arabidopsis* chromosome 1 coordinates 4.376–4.406 Mb, including the At12910 ortholog. These fine mapping studies thus positioned the potential *Rfp* coding region to a site on linkage group 18 corresponding to *Arabidopsis* chromosome 1 coordinates 4.28–4.395.

Discussion

The highly duplicated nature of *Brassica* genomes poses a significant challenge to the isolation and identification of genes through positional cloning approaches. Initial comparative mapping studies indicated that, on average, a single region of the *Arabidopsis* genome is represented as three distinct regions in genomes of the related diploid *Brassica* species (Cavell et al. 1998; Lagercrantz 1998), a view that has been confirmed by physical analysis of *Brassica* genomes (O'Neill and Bancroft 2000; Town et al. 2006; Yang et al. 2006). By extension it can be expected that most *Arabidopsis* genome regions are represented an average of six times in the amphidiploid *Brassica* species such as *B. napus*, an estimate again congruent with recent comparative sequencing results (Cheung et al. 2009). Consistent with this view, the comparative mapping analysis of Parkin et al. (2005) indicates that the region of *Arabidopsis* chromosome 1 which is orthologous to the *Rfp* region is represented 5 times in the *B. napus* genome. The multiple restriction fragments hybridizing to the *Arabidopsis* probes observed in Fig. 1 reflect this highly duplicated nature of the *B. napus* genome.

Because of the complexity of *Brassica* genomes, it is common to recover clones representing multiple genomic regions from *Brassica* genomic libraries. To identify which clones correspond to a region specifying the trait of interest, it is necessary to physically link them to a mapped polymorphism genetically linked to that trait. Success in positional mapping therefore is critically dependent on the identification of tightly linked polymorphisms that can be used to recover corresponding genomic clones.

The possibility of exploiting genomic co-linearity by using sequences from a simpler genome as a means of obtaining tightly linked, co-linear markers for mapping in a related species (Bennetzen and Freeling 1993; Ramakrishna and Bennetzen 2003) has received widespread application in cereals (e.g. Kilian et al. 1997; Dunford et al. 2002; Perovic et al. 2004) and more limited applications in eudicot crop species (e.g. Ku et al. 2001; Kaló et al. 2004). As we show here, the *Arabidopsis* genome provides a powerful tool for comparative, targeted fine mapping in *Brassica* species that partially offsets the difficulties posed for positional cloning by the complexity of these genomes. The recovery of *Brassica* clones containing the tightly linked cRF1 polymorphism initially allowed the *Rfp* region to be physically anchored, and the recovery of the 2840A3 clone spanning a more tightly linked SNP, as described here, should facilitate the recovery of additional clones from the targeted region.

Although our mapping results strongly support the view that extended co-linearity exists between the *Rfp* region and the orthologous segment of the *Arabidopsis* genome, we identified a single exception to the co-linearity involving marker 7.5BB. The position of this marker on the *Arabidopsis* genome suggested that it would map between 2.7 HH and 14o on the 4ND7 side of *Rfp*. Instead, we found that 7.5BB mapped to the opposite side of *Rfp*, between markers 3.5BB and cRF1. This suggests that the sequences encompassing the 7.5BB polymorphism have undergone transposition since the divergence of the *Arabidopsis* and *Brassica* lineages. O'Neill and Bancroft (2000) have previously noted an apparent translocation of a gene in a comparison of co-linear *B. oleracea* and *Arabidopsis* genome segments. In addition, we have noted that genes encoding PPR proteins often appear in different chromosomal contexts within otherwise co-linear regions of *Arabidopsis* and *Brassica* chromosomes (Geddy and Brown 2007). No PPR gene is located in the region of the *Arabidopsis* genome from which the 7.5BB probe is derived. We have, however, found that sequences near the 7.5BB probe contain a helitron mobile element; this class of elements has been found to capture associated genes during transposition (Kapitonov and Jurka 2007). Additional insight into the molecular basis of this observation must await further analysis.

In summary, we have used *Arabidopsis* genome resources to directly obtain a set of ordered markers for high-resolution mapping of the *Brassica Rfp* gene. The *Rfp* gene has now been localized to a chromosomal region that is equivalent to a 115-kb segment of *Arabidopsis* chromosome 1. Since it has been estimated that the ratio of corresponding physical distances in *Brassica* and *Arabidopsis* genomes is <1.5 (O'Neill and Bancroft 2000), we expect that the markers defining the *Rfp* interval should be separated by less than 180 kb. This should now allow the physical interval spanning these markers to be isolated and characterized. The present study builds on our earlier work (Brown et al. 2003) and provides an additional example of the utility of this approach for fine mapping genes in *Brassica* species.

Acknowledgments This work was supported by funds from Discovery and Collaborative Research and Development Grants from the Natural Sciences and Engineering Research Council (NSERC), Canada, by BASF Plant Science and by the Industrial Research Assistance Program of the National Research Council (NRC), Canada. We express

thanks to Dr. Rod Wing, University of Arizona, for providing the *Arabidopsis* BACs and to Dr. Tom Bureau, Department of Biology, McGill University, for analysis of an *Arabidopsis* genome region for transposable element sequences.

References

Bennetzen JF, Freeling M (1993) Grasses as a single genetic system: genome composition, collinearity and compatibility. Trends Genet 9:251–253

Brown GG (1999) Unique aspects of cytoplasmic male sterility and nuclear fertility restoration in *Brassica napus*. J Hered 90:351–356

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

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

Cavell AC, Lydiate DJ, Parkin IA, 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 41:62–69

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

Cheung F, Trick M, Drou N, Lim YP, Park JY, Kwon SJ, Kim JA, Scott R, Pires JC, Paterson AH, Town C, Bancroft I (2009) Comparative analysis between homoeologous genome segments of *Brassica napus* and its progenitor species reveals extensive sequence-level divergence. Plant Cell 21:1912–1928

Dunford RP, Yano M, Kurata N, Sasaki T, Huestis G, Rocheford T, Laurie DA (2002) Comparative mapping of the barley *Ppd-H11* photoperiod response gene region, which lies close to a junction between two rice linkage segments. Genetics 161:825–834

Elina H (2007) Expression of a *Brassica napus* mitochondrial gene region associated with CMS: transcript initiation, splicing and nuclease protection. Ph.D. thesis, McGill University

Fan Z, Stefansson BR, Sernyk JL (1986) Maintainers and restorers for three male sterility-inducing cytoplasms in rape (*Brassica napus* L). Can J Plant Sci 66:229–234

Formanová N, Li X-Q, DePauw M, Keller WA, Landry BS, Brown GG (2006) Towards positional cloning in *Brassica napus*: generation and analysis of doubled haploid *B. rapa* possessing the *B. napus pol* CMS and *Rfp* nuclear restorer gene. Plant Mol Biol 61:269–281

Geddy R, Brown GG (2007) Genes encoding pentatricopeptide repeat (PPR) proteins are not conserved in location in plant genomes and may be subject to diversifying selection. BMC Genomics 8:13

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

Jean M, Brown GG, Landry BS (1997) Genetic mapping of nuclear fertility restorer genes for the 'Polima' cytoplasmic male sterility of canola (*Brassica napus* L.) using DNA markers. Theor Appl Genet 95:321–328

Kaló P, Seres A, Taylor SA, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis TH, Kiss GB (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. Mol Genet Genomics 272:235–246

Kapitonov VV, Jurka J (2007) Helitrons on a roll: eukaryotic rolling circle transposons. Trends Genet 23:521–529

Kilian A, Chen J, Han F, Steffenson B, Kleinhofs A (1997) Towards map-based cloning of the barley stem rust resistance genes *Rpg1* and *rpg4* using rice as an intergenomic cloning vehicle. Plant Mol Biol 35:187–195

Ku H-M, Liu J, Doganlar S, Tanksley SD (2001) Exploitation of *Arabidopsis*–tomato synteny to construct a high-resolution map of the *ovate* containing region in tomato chromosome 2. Genome 44:470–475

L'Homme Y, Stahl RJ, Li X-Q, Hameed A, Brown GG (1997) *Brassica nap* cytoplasmic male sterility is associated with expression of a mitochondrial gene region containing a chimeric gene similar to the *pol* CMS-associated *orf224* gene. Curr Genet 31:325–335

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

Landry BS, Hubert N, Etoh T, Harada JJ, Lincoln SE (1991) A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. Genome 34:543–552

Li X-Q, Jean M, Landry BS, Brown GG (1998) Restorer genes for different forms of *Brassica* cytoplasmic male sterility map to a single nuclear locus that modifies transcripts of several mitochondrial genes. Proc Natl Acad Sci USA 95:10032–10037

Menassa R, L'Homme Y, Brown GG (1999) Post-transcriptional and developmental regulation of a CMS-associated mitochondrial gene region by a nuclear restorer gene. Plant J 17:491–499

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 using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

O'Neill CM, Bancroft I (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homoeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. Plant J 23:233–243

Parkin I, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics 171:765–781

Pelletier G, Budar F (2007) The molecular biology of cytoplasmically inherited male sterility and prospects for its engineering. Curr Opin Biotechnol 18:121–125

Perovic D, Stein N, Zhang H, Drescher A, Prasad M, Kota R, Kopahnke D, Graner A (2004) An integrated approach for comparative mapping in rice and barley with special reference to the *Rph16* resistance locus. Funct Integr Genomics 4:74–83

Ramakrishna W, Bennetzen JL (2003) Genomic colinearity as a tool for gene isolation. Methods Mol Biol 236:109–122

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

Touzet P, Budar F (2004) Unveiling the molecular arms race between two conflicting genomes in cytoplasmic male sterility? Trends Plant Sci 9:568–570

Town CD, Cheung F, Maiti R, Crabtree J, Haas BJ, Wortman JR, Hine EE, Althoff R, Arbogast TS, Tallon LJ, Vigouroux M, Trick M, Bancroft I (2006) Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. Plant Cell 18:1348–1359

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

Yang TJ, Kim JS, Kwon SJ, Lim KB, Choi BS, Kim JA, Jin M, Park JY, Lim MH, Kim HI, Lim YP, Kang JJ, Hong JH, Kim CB, Bhak J, Bancroft I, Park BS (2006) Sequence-level analysis of the diploidization process in the triplicated FLOWERING LOCUS C region of *Brassica rapa*. Plant Cell 18:1339–1347

Zhang J, Stewart JM (2000) The economic and rapid method for extraction of cotton genomic DNA. J Cotton Sci 4:193–201