Theor Appl Genet (1998) 97 : 431–438

M. Jean · G. G. Brown · B. S. Landry

Targeted mapping approaches to identify DNA markers linked to the *Rfp1* restorer gene for the 'Polima' CMS of canola (*Brassica napus* L.)

Received: 6 June 1997 / Accepted: 12 February 1998

Abstract We have used two targeting approaches [pairs of nearly isogenic lines (NILs) and bulked segregant analysis] to identify DNA markers linked to the *Rfp1* restorer gene for the *pol* CMS of canola (*Brassica napus* L.). We were able to target the *Rfp1* locus as efficiently by comparing NILs as by bulked segregant analysis, and it was demonstrated in this instance that double-screening strategies could significantly improve the overall targeting efficiency. The chance occurrence of shared homozygosity at specific unlinked chromosomal regions in the bulks was found to limit the efficiency of bulked segregant analysis, while the efficiency of NIL comparison was limited by residual DNA from the donor cultivar at scattered sites throughout the genome of the NILs.

Keys words Targeted mapping · RFLP · RAPD · *Brassica napus* · Polima CMS · Nearly isogenic line · Bulked segregant analysis

Introduction

Finding genetic markers linked to a particular locus can be a laborious process if markers are tested at random. Two targeted mapping approaches have been described to ease this process: comparison of nearly isogenic lines (NILs) (Muehlbauer et al. 1988) and bulked segregant analysis (Michelmore et al. 1991). NILs are the natural result of introgressive and backcross breeding programs aimed at transferring important agronomic characteristics into elite cultivars. The donor of the trait is first crossed to the recipient cultivar. The F_1 hybrid generated is then crossed back to the recipient cultivar (designated as the recurrent parent), and the process is repeated several times until the unwanted portion of the genome from the donor is largely eliminated while the genome from the recipient strain is retained and the desired trait is introduced. The new breeding line is then nearly isogenic to the recurrent parent (the original elite cultivar) but still contains a limited amount of DNA from the donor, some of it flanking the introgressed gene, and some scattered at random sites in the genome.

Comparing recurrent parent, donor parent and their derived NILs to identify regions of the genome containing residual donor DNA can be an efficient way to identify markers linked to a gene of interest. This approach has been successfully used to find DNA markers linked to several genes (Young et al. 1988; Sarfatti et al. 1989; Martin et al. 1991; Messeguer et al. 1991; Paran et al. 1991; Diers et al. 1992; Schüller et al 1992; van der Beek et al. 1992; Barua et al. 1993). Unfortunately, NILs are often not available for important agronomic traits and producing them is a lengthy process.

Bulked segregant analysis allows marker targeting using any population segregating for a given characteristic. In this approach, the DNA from plants homozygous for a shared character (a phenotype or genomic interval) are mixed together to generate two DNA bulks representing alternate forms of the character. The DNA bulks will contain a random assortment of alleles for markers unlinked to the trait of interest but will have contrasting alleles for markers linked to the target, with the bulk representing a particular phenotype carrying the same allele as the donor of that phenotype. A growing number of publications have reported success using bulked segregant analysis to

Communicated by M. A. Saghai Maroof

M. Jean · G. G. Brown
McGill University, Biology Department, 1205 ave. Doctor Penfield, Montréal, Québec, Canada, H3A 1B1

B. S. Landry (✉)
DNA LandMarks, P.O. Box 6, St-Jean-sur-Richelieu, Québec, Canada, J3B 6Z1

target markers to particular genes or genomic intervals (Giovannoni et al. 1991; Reiter et al. 1992; Barua et al. 1993; Chalmers et al. 1993; Kesseli et al. 1993; Williams et al. 1993; Delourme et al. 1994).

Our main interest is the *Rfp1* restorer gene for the 'Polima' (*pol*) cytoplasmic male sterility (CMS) of *Brassica napus*. In the study presented here, we tested two targeting approaches (the comparison of NILs and the bulked segregant analysis) for their efficiency in identifying DNA markers linked to the *Rfp1* locus. We also analyzed the limitations of each approach and assessed what can be done to maximize the targeting efficiency.

Materials and methods

Plant materials and fertility scoring

The plant materials used for the present study were described in Jean et al. (1997). Cytoplasms and genotypes are designated by parenthesized italics following the cultivar name. Since a *pol* CMS and *pol* restorer lines derived from the same breeding program were not available, the NIL comparison was performed with a plant from the breeding line Westar-Rf (*Rfp1Rfp1*; *pol*) and a plant from the cultivar 'Westar' (*rfprfp*;*nap*). A plant from the cultivar 'Italy', the donor source of the *Rfp1* gene in Westar-Rf, was also used for comparison.

The backcrossed populations used for the bulked segregant analysis and the co-segregation analysis were generated as described in Jean et al. (1997). The KW population was made of 111 BC_1 plants derived from an intervarietal cross between a *pol* CMS Karat (*rfprfp*;*pol*) and *pol* restorer Westar-Rf (*Rfp1Rfp1*; *pol*) plants. The WW population was composed of 64 BC_1 plants derived from an intravarietal cross between the 2 'Westar' plants used for the NIL comparison (a 'Westar' (*rfprfp*;*nap*) and Westar-Rf (*Rfp1Rfp1*;*pol*) individuals). Different Westar-Rf (*Rfp1Rfp1*;*pol*) individuals were used to generate the two populations.

Plants were grown to maturity in growth chambers or greenhouses under standard conditions (16-h photoperiod, day/night temperatures of 22°/16°C). The scoring of fertility was as described in Jean et al. (1997).

Bulked segregant analysis

Bulked segregant analysis (Michelmore et al. 1991) was performed with plants from the KW and the WW populations. Bulked DNA samples were made for each population by combining DNA of sterile plants (S bulk) and fertile plants (F bulk). For the WW population, the S and F bulks each contained DNA from 15 plants. For the KW population, two S bulks and one F bulk were made, and different bulks were used for the restriction fragment length polymorphism (RFLP) and random amplified polymorphic DNA (RAPD) analyses. For the RFLP analysis, the S1, S2 and F bulks contained the DNA from 4, 9 and 30 plants, respectively. For the RAPD analysis, the S1, S2 and F bulks contained the DNA from 6, 12 and 15 plants, respectively. The S1 and S2 bulks were made of different plants; the plants from the S2 (RFLP analysis) and F (RAPD analysis) bulks were subsets of those from the S2 (RAPD analysis) and F (RFLP analysis) bulks.

Bulked segregant analysis is usually performed by comparing bulks of individuals homozygous for the alternate genotype, and markers potentially linked to a targeted locus are detected as a band specific to each bulk, usually similar to the one found in the parent with which it shares its phenotype. However, since we used a backcross population, only alleles from the donor parent segregated, and the only homozygous plants were those with the genotype of the recurrent parent. A marker potentially linked to the targeted locus was therefore detected as the absence of a band from the donor parent (non-recurrent) in the bulk of plants with the phenotype of the recurrent parent.

RAPD and RFLP analyses

Plant DNA extractions, RAPD and RFLP analyses, as well as the preparation of the probe OPF14, were performed as described in Jean et al. (1997). For RAPD analysis, random 10-mer primers (Operon Technologies, Alameda, Calif.; kits A to G) were tested on the parental DNAs. The probes used for RFLP analysis were hybridized on four restriction enzyme digests (*Bam*HI, *Eco*RI, *Eco*RV and *Hin*dIII) of the parental DNAs. Some of the probes tested in the present study were chosen from among those localized on the linkage map of *B. napus* genome (Landry et al. 1991), while the others were unmapped anonymous cDNAs prepared from a library of *B. napus* embryo cDNA clones (Harada et al. 1988). A RAPD marker for the *pol* CMS restorer (OPF14, Jean et al. 1997) was also used as probe. Probe preparation and hybridizations were performed as described in Jean et al. (1997).

Primers and probes identifying polymorphisms (additional DNA bands) from the Westar-Rf (*Rfp1Rfp1*; *pol*) parent were tested on the bulks. The segregation of the markers selected as potentially Rf-linked by bulked segregant analysis or by comparing the 2 'Westar' plants was then assessed in the BC_1 populations. These markers were also hybridized to restriction enzyme digests of DNA from the cultivar 'Italy'.

Genetic mapping

Loci are named after the probe or primer that revealed them, with lowercase letters (a, b . . .) designating separated loci revealed with the same probe or primer. Loci common to the present and previously published maps of *B. napus* genome (Landry et al. 1991) are labeled identically whenever possible.

The segregation of each locus was compared to Mendelian ratios expected in a BC_1 population using a Chi^2 for goodness-of-fit analysis, performed with Yate's correction as described in Little and Jackson-Hills (1978). Loci that did not fit a 1:1 ratio ($P > 0.05$) were not included in the present analysis. Co-segregation analyses were performed with the software MAPMAKER v. 2.01 (Lander et al. 1987, supplied by S. Tingey, DuPont) on a Macintosh LC III computer, as described in Jean et al. (1997). The statistical thresholds were set at a minimum LOD score of 4.0 and a maximum recombination fraction of 0.30. The Kosambi mapping function was used to convert the recombination frequency into centiMorgans.

Results and discussion

Targeting efficiency of NIL and bulked segregant analyses

Two targeting approaches, nearly isogenic line (NIL) analysis (Muehlbauer et al. 1988) and bulked segregant analysis (Michelmore et al. 1991), were tested for their efficiency in identifying DNA markers linked to the *Rfp1* restorer gene for the *pol* CMS. A total of 208 probes and 127 primers were screened by the analysis of DNAs from a pair of 'Westar' NILs. Potentially

Table 1 Targeting efficiency achieved by bulked segregant analysis, by NIL analysis and by double-screening strategies

	Percentage of screening achieved[a]		Percentage of targeting achieved[b]	
	RAPD analysis	RFLP analysis	RAPD analysis	RFLP analysis
Single-screening strategies				
– A single pair of bulks[c]				
– the F and S1 bulks	2.4% (3/127)	8.3% (20/242)	33.3% (1/3)	40.0% (8/20)
– the F and S2 bulks	5.5% (7/127)	8.3% (20/242)	12.5% (1/8)[e]	40.0% (8/20)
– A single pair of NILs[d]	3.1% (4/127)	7.2% (15/208)	0% (0/4)	40.0% (6/15)
Double-screening strategies				
– Two pairs of bulks from a regular population[c]	1.6% (2/127)	3.3% (8/242)	50.0% (1/2)	100% (8/8)
– One pair of bulks from a NIL-derived population[d]	0% (0/127)	2.9% (6/208)	– –	100% (6/6)

[a] The values in parentheses are the number of primers or probes displaying a polymorphism selected as potentially linked to the *Rfp1* locus over the total number of primers or probes tested
[b] The values in parentheses are the number of RAPD or RFLP markers truly linked to the *Rfp1* locus over the total number of RAPD or RFLP markers selected as potentially linked to the *Rfp1* locus
[c] These analyses were performed with plants from the KW population as described in Materials and methods and in Results and Discussion
[d] These analyses were performed with plants from the WW population as described in Materials and methods and in Results and Discussion
[e] One of the seven RAPD primers (OPA13) revealed two potentially Rf-linked loci

Rf-linked markers were detected with 15 (7.2%) RFLP probes and 4 (3.1%) RAPD primers (Table 1). Similarly, 242 probes and 127 primers were screened by the analysis of DNAs from the parents of the KW population and from pairs of DNA bulks derived from this population. Potentially Rf-linked markers were detected with 20 (8.3%) RFLP probes and 3 (2.4%) RAPD primers when the screening was performed with the F and S1 bulks, and with 20 (8.3%) RFLP probes and 7 (5.5%) RAPD primers when the screening was performed with the F and S2 bulks (Table 1).

Co-segregation analyses were performed with the WW population to confirm linkage between the *Rfp1* locus and the DNA loci selected as potentially Rf-linked by the analysis of the NILs. Among these, 6 (40.0%) RFLP markers and no (0%) RAPD markers proved to be linked to the targeted *Rfp1* locus (Table 1). Similarly, co-segregation analyses were performed with the KW population to confirm linkage between the *Rfp1* locus and the DNA loci selected as potentially Rf-linked by the analysis of the bulks. Among these, a total of 8 RFLP markers and 1 RAPD marker proved to be linked to the targeted *Rfp1* locus (Table 1). This corresponds to 40.0% of the RFLP markers and 12.5% of the RAPD markers selected with the F and S1 bulks and 40.0% of the RFLP markers and 33.3% of the RAPD markers selected with the F and S2 bulks. These results suggest that, with regard to the *Rfp1* locus, the screening and targeting efficiencies of NILs and bulked segregant analyses are similar.

Limitations to the targeting efficiency of NILs and bulked segregant analyses

In order to identify the main limitations to the targeting efficiency of NILs and bulked segregant analyses, the polymorphic DNA fragments detected in the Westar-Rf (*Rfp1Rfp1*;*pol*) breeding line for the markers selected by these targeting approaches were compared to the DNA fragments observed in cvs 'Westar' and 'Italy'.

NIL analysis

In the case of the markers selected by the NIL analysis, the polymorphic fragments observed in the Westar-Rf (*Rfp1Rfp1*;*pol*) breeding line were usually similar to DNA fragments that could be detected in the 'Italy' cultivar (Table 2). The polymorphisms still present in the BC_6 Westar-Rf (*Rfp1Rfp1*;*pol*) breeding line thus represent residual alleles from the 'Italy' cultivar, from which the *Rfp1* gene was originally introgressed. This

Table 2 Analysis of the markers selected by comparing the two 'Westar' NIL plants

Marker	Selected as Rf-linked when comparing the F bulk to the S bulk[a]	Percentage of sterile plants with a restorer-associated allele in the S bulk[b]	Presence of an allele similar to the Westar-Rf allele in the cultivar Italy[c]
Linkage group A-I (containing the *Rfp1* locus)			
5NB12	+	20	+
5NF2	+	20[a]	+
Rfp1	+	0	+
cRF1b	+	0	+
4ND7b	+	13[a]	+
3NF2	+	13	+
6NC2	+	13	–
Linkage group A-II			
OPD05	–	33[b]	N.A.
1ND1c	–	40	(+)
4NC1b	–	40	(+)
5NA2c	–	40	(+)
OPD08b	–	40	+
OPG13	–	40	+
Linkage group A-III			
4ND1	–	33[a]	(+)
6NC12	–	40	+
6ND9	–	40	N.A.
pCOT39	–	53	–
Unlinked markers			
4NF4d	–	53	+
5NB9b	–	53	+
OPF15	–	53[b]	N.A.

[a] The linkage groups correspond to those from Fig. 1A
[b] This percentage was calculated by dividing the number of recorded heterozygous individuals by the total number of individual tested. It must be considered the minimum possible value when the scoring from one or more individuals is missing. [a] Data missing for one individual; [b] data missing for three individuals
[c] For some markers, hybridization data for the restriction enzyme digest used to score the segregation of the marker were not available for all the cultivars tested. ( ) the correspondence of Rf-linked alleles was estimated from another common restriction enzyme digest; N.A., data not available

observation confirms that the targeting power of NILs is related to linkage drag, i.e. the presence of residual DNA from the donor cultivar at markers located near the introgressed gene.

Most of the markers selected by NIL analysis that were not linked to the *Rfp1* locus were clustered on two short chromosomal segments (Fig. 1). In these areas also, the Westar-Rf (*Rfp1Rfp1*;*pol*) breeding line still carried alleles similar to those of cv 'Italy' (Table 2). These chromosomal segments thus represent residual DNA from the donor cultivar that have not yet been eliminated by backcross breeding. This observation confirms that the targeting power of NILs is mainly limited by the occurrence of residual DNA from the donor cultivar at scattered sites around the genome of the NILs.

Bulked segregant analysis

In the case of the markers selected by the bulked segregant analysis performed with the KW population, the polymorphic fragments observed in the Westar-Rf (*Rfp1Rfp1*;*pol*) breeding line were usually similar to DNA fragments that could be detected in the regular 'Westar' cultivar (Table 3). The polymorphisms selected with bulked segregant analysis thus correspond to normal variations between the cultivars 'Westar' and 'Karat' (i.e. the parents of the KW population). Bulked segregant analysis is therefore insensitive to residual DNA from 'Italy', the original donor of the *Rfp1* gene. Indeed, the targeting power of bulked segregant analysis is related to the presence, among the plants of the phenotypic bulks, of a shared genotype in the chromosomal area containing the gene (or genes) responsible for the targeted phenotype and of a random assortment of genotypes for markers unlinked to the targeted regions.

Most of the markers selected with a single pair of bulks that were not linked to the *Rfp1* locus were clustered in seven chromosomal areas (Fig. 1). Four of these areas contained markers selected solely with the F and S2 bulks, while the other three contained markers selected with the F and S1 bulks. For all the false positive markers, 67% or more of the individuals included in the S bulk were of the same genotype as the 'Karat' *pol* CMS parent (Table 3). These results indicate that the most important limitations to bulked segregant analysis are the chance occurrence of shared homozygosity at specific unlinked chromosomal regions in the bulks. Indeed, with a segregating population derived only one generation after the original intercross (e.g. double haploid, F_2 and BC_1 populations), it is very likely that some genomic regions will be uncovered where the markers will have not yet been randomized through meiosis and recombination. Furthermore, a similar level of false positives was observed with the S1 bulk (containing 4–6 plants) and the S2 bulk (containing 9–12 plants) (see Table 1). Increasing the number of DNA samples combined to generate the bulks might therefore not be sufficient to eliminate all the false positives detected by bulked segregant analysis. Indeed, when a plant with a large genome size is being studied, such as *B. napus*, a number of such regions of shared homozygosity caused by linkage disequilibrium will likely be found in bulks of any convenient size.

Finally, it was noted that RFLP markers were selected as potentially Rf-linked by bulked segregant analysis only if 33% or fewer of the individuals from the S bulks carried the allele from the Westar-Rf (*Rfp1Rfp1*;*pol*) breeding line (Table 3; see also Table 2). This result indicates that the allele of the donor parent must be in a proportion of 16.5% or less to allow a variation of intensity to become identifiable. It was also observed that the S1 bulk (containing 4 BC_1

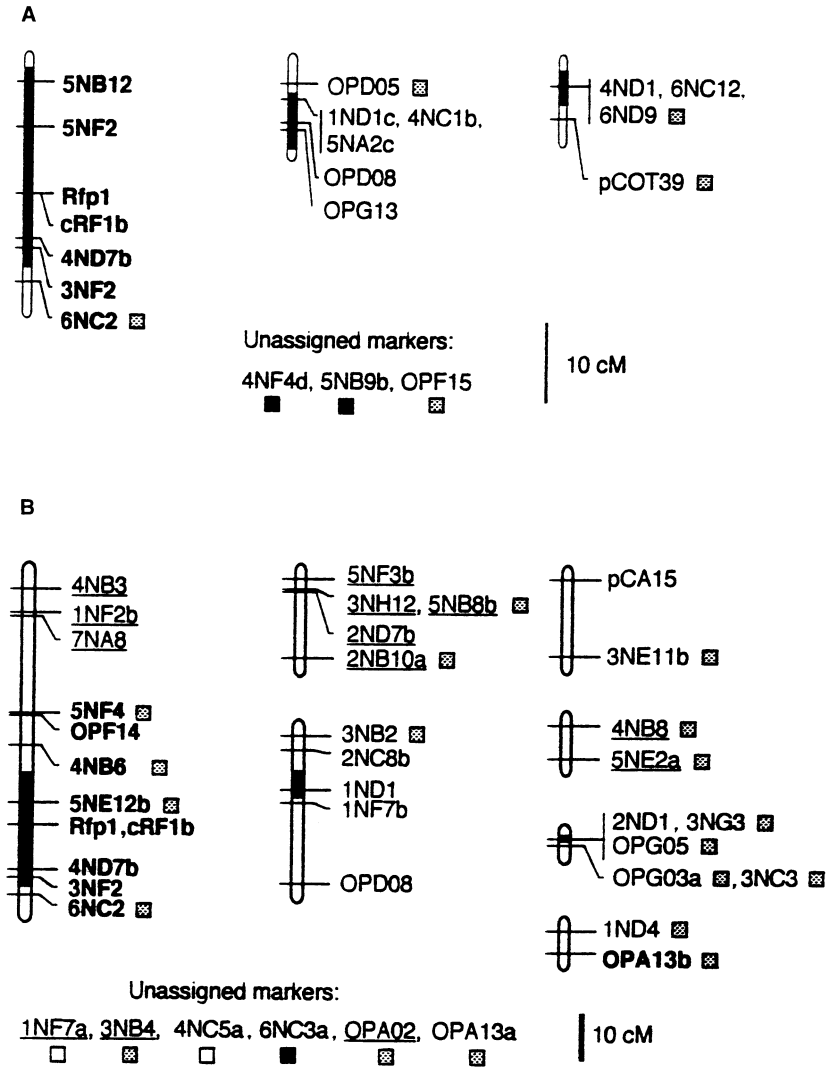

Fig. 1A, B Linkage assignment of the markers selected by various targeting approaches. In **A**, markers have been selected by NIL analysis (*normal type*); those in *bold* were also polymorphic in the analysis of bulks from the NIL-derived WW population (see Discussion). Co-segregation analyses were performed with the WW population as described in Materials and methods. In **B**, markers have been selected by the analysis of the F and S1 bulks (*normal type*) or the F and S2 bulks (*underlined type*) from the KW population; those in *bold type* were polymorphic with both pairs of bulks (see Discussion). Co-segregation analyses were performed with the KW population as described in Materials and methods. The scales represent 10 cM. Residual DNA segments from the cv 'Italy' are in *black* while DNA segments in *white* are from cv 'Westar' or of unidentified origin. *Gray boxes* designate markers for which the origin of the allele could not be identified (see Tables 2 and 3). The names (*right*) represent RFLP markers, RAPD markers (OP…) and seedling-specific cDNAs (pCA15 and pCOT39)

plants) was much easier to score than the S2 bulk (containing 9 BC_1 plants) (data not shown). The variation in intensity thus appears to be maximized by using bulks generated by pooling DNA from a small number of plants. Indeed, when a bulk contains a small number (e.g. 4) of BC_1 plants of the recurrent phenotype, a linked marker will be detected either as an absence of band or a faint band since none of the plants or only a single plant from the bulk will carry the allele of the donor parent. Such a faint band is easily discriminated from the more intense band of an unlinked marker as, in this case, 2–4 plants from the bulk would carry this allele. This also makes it improbable that a linked marker will be overlooked as a false negative result, since this could only happen if several misscored individuals or individuals with recombination events near the targeted gene were by chance included in the bulk.

Targeting efficiency of double-screening strategies

Two approaches involving a double-screening strategy were tested for their efficiency in targeting markers to the *Rfp1* gene. The first strategy was designed to take advantage of the targeting power of both NIL and bulked segregant analyses. One of the main drawbacks of NIL analysis is that pairs of NILs are not available for many agronomically important traits and producing them would be a lengthy process. Each step of a traditional backcross breeding program, however, results in the production of populations derived from crosses between very similar plants (the current breeding line and its recurrent parent). We thus hypothesized that performing a bulked segregant analysis with one of the generations of such a breeding program could be a very powerful targeting strategy. Indeed, all 6 (100%) RFLP markers selected with a pair of bulked

Table 3 Analysis of the markers selected by a bulked segregant analysis performed with the 'Karat' (*rfprfp;pol*) × 'Westar-Rf (*Rfp1Rfp1;pol*) BC_1 population

Marker[a]	Selected as Rf-linked when comparing the F bulk to the		Percentage of sterile plants with a restorer-associated allele in the[b]		Presence of an allele similar to the Westar-Rf allele in the cultivar[c]	
	S1 bulk	S2 bulk	S1 bulk	S2 bulk	Italy	Westar
Linkage group B-I (containing the *Rfp1* locus)						
4NB3	+	−	0[a]	44	−	+
1NF2c	+	−	0	44	(−)	+
7NA8	+	−	0	44	−	+
5NF4	+	+	0	22	(+)	+
OPF14	+	+	0	17	−	+
0NF14b	+	+	0	11	−	+
4NB6b	+	+	0	0	N.A.	+
5NE12b	+	+	0	0	+	+
Rfp1	+	+	0	0	+	−
cRF1b	+	+	0	0	+	−
4ND7b	+	+	0	22	+	−
3NF2	+	+	0	22	+	−
6NC2	+	+	0	22	−	−
Linkage group B-II						
OPD08a	−	+	83	25	−	+
1NF7b	−	+	100	11	−	+
1ND1	−	+	100	11[a]	(+)	−
2NC8b	−	+	100	11	−	+
3NB2	−	+	75[a]	11	+	+
Linkage group B-III						
2NB10a	+	−	25[a]	67	N.A.	+
2ND7b	+	−	0	67	(−)	(+)
3NH12	+	−	0	67	(−)	(+)
5NB8b	+	−	0	67	N.A.	+
5NF3b	+	−	0	78	−	+
Linkage group B-IV						
OPG05	−	+	67	17	N.A.	(−)
3NG3	−	+	75	22	N.A.	−
2ND1	−	+	75	22	+	−
3NC3	−	+	75	11	N.A.	−
OPG03a	−	+	50[a]	25	N.A.	+
Linkage group B-V						
3NE11b	−	+	50	22	+	+
pCA15	−	+	75	22	−	+
Linkage group B-VI						
4NB8	+	−	0	67	N.A.	+
5NE2a	+	−	0	67	N.A.	+
Linkage group B-VII						
1ND4	−	+	50	33[a]	N.A.	+
OPA13b	+	+	17	58[a]	N.A.	N.A.
Unlinked markers						
1NF7a	+	−	25	67[a]	−	+
3NB4	+	−	0	67	N.A.	+
4NC5a	−	+	75	33	(−)	+
6NC3a	−	+	75	11	+	−
OPA02	+	−	17	33	N.A.	+
OPA13a	−	+	83	25[a]	N.A.	N.A.

[a] The linkage groups correspond to those from Fig. 1B

[b] This percentage was calculated by dividing the number of recorded heterozygous individuals by the total number of individual tested. It must therefore be considered as representing the minimum possible value when the scoring from one or more individuals is missing. [a] Data missing for one individual

[c] For some markers, hybridization data for the restriction enzyme digest used to score the segregation of the marker were not available for all the cultivars tested. ( ), The correspondence of Rf-linked alleles was estimated from another common restriction enzyme digest; N.A., data not available

segregants from the WW population derived from the 'Westar' NILs proved to be linked with the *Rfp1* locus (Table 1). Our results thus indicate that bulked segregant analysis can efficiently screen out the false positive markers that slipped through NIL analysis, probably because it is insensitive to residual donor DNA unlinked to the target locus (see above).

The second strategy was designed to maximize the targeting achieved with bulked segregant analysis. It has been demonstrated that the targeting power of NIL analysis could be increased by combining the information from the analysis of several distinct pairs of NILs (Paran et al. 1991). This approach is thought to screen out the detection of unlinked residual donor DNA since distinct NILs are unlikely to carry the same residual regions. By analogy to this approach, we thus hypothesized that the comparative use of two or more pairs of bulks should improve the targeting efficiency of bulked segregant analysis. Indeed, all 8 (100%) RFLP markers and 1 of the 2 (50.0%) RAPD markers selected by combining the information from the analysis of two pairs of bulks from the KW population proved to be linked with the *Rfp1* locus (Table 1). Our results thus indicate that the comparative use of two or more pairs of bulks, even small ones, can efficiently screen out the false positive markers that slipped through a bulked segregant analysis performed with a single pair of bulks, probably by screening out the detection of regions of shared homozygosity unlinked to the targeted locus (see above). Such an approach also has the advantage of decreasing the negative impact resulting from the accidental inclusion in a bulk of misscored individuals or plants with recombination near the locus of interest. Indeed, 2 plants scored as CMS at an early stage of our experiment and included in the S2 bulk used for RAPD analysis in the KW population were found to be completely sterile (e.g. both female and male sterile), hence not CMS. These misscored individuals could be responsible for some of the false positives detected with this particular bulk. Since misscoring can be frequent if the trait studied is difficult to assay, affected by environmental conditions or conditioned by several genes and since misscored plants are often detected late in a mapping program, the use of more than one bulk during the screening process could prevent research efforts being wasted by the analysis of unreliable bulks.

Toward the restoration of isogenicity in the Westar-Rf breeding line

If the residual 'Italy' fragment found around the *Rfp1* locus is considered to be defined by the 6 markers 5NE12b, 5NB12, 5NF2, cRF1b, 4ND7b and 3NF2 (Table 3), its size can be estimated to 26.1 cM in the 'Westar-Rf (*Rfp1Rfp1*;*pol*) parent from the WW population, and 17.7 cM in the Westar-Rf (*Rfp1Rfp1*;*pol*) parent from the KW population (Fig. 1). This chromosomal segment, however, represents only 40.0% (6/15) of the total residual 'Italy' DNA still present in the Westar-Rf (*Rfp1Rfp1*;*pol*) breeding line (estimated to 7.2% (15/208), see Table 1). According to Muehlbauer et al. (1988), after six backcrosses, followed by an intercross and another backcross ($t = 8$), in a plant like *B. napus* ($n = 19$) with a predicted genome length of 2826 cM [two times the map of Landry et al. (1991), which was estimated to cover half of the *B. napus* genome], there should be about 1.3% of the donor genome (i.e. about 36 cM) left in the breeding line, 70% (25 cM) of which should be around the introgressed locus. The size of the residual 'Italy' fragment found in the BC_6 Westar-Rf (*Rfp1Rfp1*;*pol*) breeding line is thus in good agreement to this theoretical value. The proportion of unlinked residual 'Italy' DNA, however, appears to be much higher than it should at this stage of a backcross breeding program. Our results thus indicate that the six backcrosses yet performed have only been moderately effective in restoring isogenicity to the Westar-Rf (*Rfp1Rfp1*;*pol*) breeding line. Indeed, traditional backcross breeding programs rely solely on segregation to eliminate residual donor DNA unlinked to the introgressed gene and on recombination to decrease the size of donor DNA associated with the introgressed gene. The amount of progress achieved therefore heavily depends on the stochastic fluctuation in the amount of residual donor DNA present in the individuals chosen as parents for each backcross generation. Although most of the donor DNA can be effectively eliminated in the early backcrossing steps if parents are chosen on a random basis, the removal of donor DNA becomes increasingly inefficient if the process is continued into the later generations of an introgressive breeding program. This inefficiency to achieve an increase toward the isogenicity of the NILs applies both to residual donor DNA unlinked to the targeted locus, as demonstrated by the results of the present study, as well as to the residual donor DNA linked to a targeted locus, as already pointed out by Young and Tanksley (1989). By allowing the selection of the best progenitors from each generation, a marker-assisted breeding program would therefore be more efficient in restoring the recurrent genome, and the number of generations needed to achieve this goal could be reduced. As markers are identified for an ever-increasing number of agronomically important traits, breeding programs for various crops should benefit from the implementation of strategies based on marker-assisted selection.

Acknowledgements Special thanks are extended to Y. Dion and F. Vachon for their technical help with the hybridization experiments required for the initial screening of RFLP probes. This work was supported by the Natural Sciences and Engineering Research Council of Canada, grants to BSL and GGB.

References

Barua UM, Chalmers KJ, Hackett CA, Thomas WTD, Powell W, Waugh R (1993) Identification of RAPD markers linked to a *Rynchosporium secalis* resistance locus in barley using near-isogenic lines and bulked segregant analysis. Heredity 71: 177–184

Chalmers KJ, Barua UM, Hackett CA, Thomas WTB, Waugh R, Powell W (1993) Identification of RAPD markers linked to genetic factors controlling the milling energy requirement of barley. Theor Appl Genet 87: 314–320

Delourme R, Bouchereau A, Hubert N, Renard M, Landry BS (1994) Identification of RAPD markers linked to fertility restorer gene for Ogura radish cytoplasmic male sterility of rapeseed (*Brassica napus* L.). Theor Appl Genet 88: 741–748

Diers BW, Mansur L, Imsande J, Shoemaker RC (1992) Mapping *Phytophthora* resistance loci in soybean with restriction fragment length polymorphism markers. Crop Sci 32: 377–383

Giovannoni JJ, Wing RA, Ganal MW, Tanksley SD (1991) Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. Nucleic Acids Res 19: 6553–6558

Harada JJ, Baden CS, Comai L (1988) Spatially regulated genes expressed during seed germination and postgerminative development are activated during embryogeny. Mol Gen Genet 212: 466–473

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.). Theor Appl Genet 95: 321–328

Kesseli R, Witsenboer H, Stanghellini M, Vandermark G, Michelmore R (1993) Recessive resistance to *Plasmopara lactucae-radicis* maps by bulked segregant analysis to a cluster of dominant diseace resistance genes in lettuce. Mol Plant-Microbe Interact 6: 722–728

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SL, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174–181

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

Little TM, Jackson-Hills F (1978) Agricultural experimentation. Design and analysis. John Wiley & Sons, New York

Martin GB, Williams JGK, Tanksley SD (1991) Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near-isogenic lines. Proc Natl Acad Sci USA 88: 2336–2340

Messeguer R, Ganal M, de Vicente MC, Young ND, Bolkan H, Tanksley SD (1991) High resolution RFLP map around the root-knot nematode resistance gene (*Mi*) in tomato. Theor Appl Genet 82: 529–536

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

Muehlbauer GJ, Specht JE, Thomas-Comton MA, Staswick PE, Bernard RL (1988) Near-isogenic lines – a potential resource in the integration of conventional and molecular marker linkage maps. Crop Sci 28: 729–735

Paran I, Kesseli R, Michelmore RW (1991) Identification of restriction fragment length polymorphism and random amplified polymorphic DNA markers linked to downy mildew resistance genes in lettuce, using near-isogenic lines. Genome 34: 1021–1027

Reiter RS, Williams JGK, Feldman KA, Rafalski JA, Tingey SV, Scolnik PA (1992) Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Proc Natl Acad Sci USA 89: 1477–1481

Sarfatti M, Katan J, Fluhr R, Zamir D (1989) An RFLP marker in tomato linked to the *Fusarium oxysporum* resistance gene *I2*. Theor Appl Genet 78: 755–759

Schüller C, Backes G, Fischbeck G, Jahoor A (1992) RFLP markers to identify the alleles on the *MIa* locus conferring powdery mildew resistance in barley. Theor Appl Genet 84: 330–338

van der Beek JG, Verkerk R, Zabel P, Lindhout P (1992) Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: *Cf9* (resistance to *Cladosporium fulvum*) on chromosome 1. Theor Appl Genet 84: 106–112

Williams JGK, Reiter RS, Young RM, Scolnik PA (1993) Genetic mapping of mutations using phenotypic pools and mapped RAPD markers. Nucleic Acids Res 21: 2697–2702

Young ND, Tanksley SD (1989) RFLP analysis of the size of chromosomal segments retained around the *Tm-2* locus of tomato during backcross breeding. Theor Appl Genet 77: 53–359

Young ND, Zamir D, Ganal MW, Tanksley SD (1988) Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the *Tm-2a* gene in tomato. Genetics 120: 579–585