Theor Appl Genet (2005) 111: 446–455
DOI 10.1007/s00122-005-2017-5

ORIGINAL PAPER

J. L. Prieto · N. Pouilly · E. Jenczewski
J. M. Deragon · A. M. Chèvre

Development of crop-specific transposable element (SINE) markers for studying gene flow from oilseed rape to wild radish

Received: 3 December 2004 / Accepted: 30 March 2005 / Published online: 8 June 2005
© Springer-Verlag 2005

Abstract The screening of wild populations for evidence of gene flow from a crop to a wild related species requires the unambiguous detection of crop genes within the genome of the wild species, taking into account the intraspecific variability of each species. If the crop and wild relatives share a common ancestor, as is the case for the *Brassica* crops and their wild relatives (subtribe Brassiceae), the species-specific markers needed to make this unambiguous detection are difficult to identify. In the model oilseed rape (*Brassica napus*, AACC, $2n=38$)-wild radish (*Raphanus raphanistrum*, RrRr, $2n=18$) system, we utilized the presence or absence of a short-interspersed element (SINE) at a given locus to develop oilseed rape-specific markers, as SINE insertions are irreversible. By means of sequence-specific amplified polymorphism (SINE-SSAP) reactions, we identified and cloned 67 bands specific to the oilseed rape genome and absent from that of wild radish. Forty-seven PCR-specific markers were developed from three combinations of primers anchored either in (1) the 5′- and 3′-genomic sequences flanking the SINE, (2) the 5′-flanking and SINE internal sequences or (3) the SINE internal and flanking 3′-sequences. Seventeen markers were monomorphic whatever the oilseed rape varieties tested, whereas 30 revealed polymorphism and behaved either as dominant (17) or co-dominant (13) markers. Polymorphic markers were mapped on 19 genomic regions assigned to ten linkage groups. The markers developed will be efficient tools to trace the occurrence and frequency of introgressions of oilseed rape genomic region within wild radish populations.

Communicated by C. Möllers

J. L. Prieto · J. M. Deragon
CNRS UMR6547 Biomove, Université Blaise Pascal,
24 Avenue des Landais, 63177 Aubière Cedex, France

N. Pouilly · E. Jenczewski · A. M. Chèvre (✉)
UMR ENSAR-INRA, Station de Génétique
et Amélioration des Plantes, BP 35327,
35653 Le Rheu Cedex, France
E-mail: chevre@rennes.inra.fr
Tel.: +33-2-23485131
Fax: +33-2-23485120

Introduction

Gene flow from crops to wild related species has been the focus of intensive investigation during the last 10 years due to the release of genetically modified varieties and their possible impact on the diversity and biology of related weeds (Gepts and Papa 2003; Jenczewski et al. 2003). The first approach used to assess the probability of crop (trans)gene transfer into wild populations is based on a step-by-step strategy that focuses on (1) the presence of crops and close relatives growing and flowering nearby, (2) the extent of pollen and seed dispersal, (3) the production of viable, fit and fertile F_1 interspecific hybrids, (4) the production of viable, fit and fertile successive generations, (5) the opportunity for gene transmission, chromosome recombination and crop gene introgression into the wild genetic background and, finally, (6) the persistence of introgressed crop genes in natural populations. Most of studies analyzing gene dispersal was analyzed used controlled experimental designs in which both the cultivated plots and recipient wild/weed populations were known and controlled; the wild and cultivated plants selected generally had fixed alternate alleles (transgene, molecular or morphological markers, etc.) that enabled the identification of hybrids though progeny testing in the wild plants (Jenczewski et al. 2003). These studies have provided a wealth of important information on the genetic mechanisms that are involved in the gene flow process and have usually indicated the limiting step. However, they have only provided a short-term view of gene flow and have barely taken into account the genetic diversity of the sink and source populations and the different agronomic situations. A second approach is therefore needed to assess

the long-term stabilization of crop genes within the genome of weedy relatives, and this approach would ideally proceed by screening natural populations to detect introgressed crop genes. However, such a strategy, which would enable the tracing back of crop-to-wild gene flow in natural populations, is barely applicable because of the difficulty in identifying unambiguously crop-specific markers. This is particularly difficult when the crop and its weedy relatives share a common ancestor, as in the *Brassica* crops such as oilseed rape (*Brassica napus* L.) and its wild relatives (Song et al. 1988; Warwick and Black 1991) and display high levels of intraspecific diversity. Only herbicide tolerance, which is conferred by a transgene, has been used for the large-scale detection of oilseed rape gene hybridizations with wild populations (Warwick et al. 2003). This method allows the identification of short-term events (such as herbicide-tolerant varieties commercially released for fewer than 10 years; James 2004) but does not provide information on the potential effect that transgene position in the crop genome may have on its probability to be transferred by recombination into the genome of the wild relative.

Retroelements are one of the most promising sources of species-specific markers. Transposable elements (TEs) are usually classified in two major groups (Capy et al. 1998): class-I elements use an RNA intermediate and a reverse-transcriptase, while class-II elements use a DNA intermediate and a transposase. Short-interspersed elements (SINEs) are class-I TEs that have invaded the genome of most eukaryotes (Okada 1991). The SINEs are non-autonomous (i.e. they do not code for the proteins needed for their mobility) and must use the enzymatic machinery of autonomous LINEs (long-interspersed elements) for retrotransposition.

The S1 elements are short (180 bp) tRNA-related SINEs present at 500–1,000 copies per haploid genome in species of the *Brassiceae* tribe. These copies were generated through several waves of amplification at different evolutionary times (Deragon et al. 1994, 1996; Lenoir et al. 1997). The S1 elements have structural characteristic features found in SINEs, including a primary and secondary sequence homology to tRNA, a 3′-poly(A) region, an internal promoter recognized by the RNA polymerase III complex and flanking target site duplications (TSD).

Since new SINE insertions occur by a ‘‘copy and paste’’ (and not a ‘‘cut and paste’’) process, these insertions are effectively irreversible, and the presence or absence of a SINE at a given locus provides information that can be used to design molecular markers (reviewed in Cook and Tristem 1997; Shedlock and Okada 2000). A major advantage of this type of marker is that the probability of independent insertions at the same exact chromosomal site is virtually nil (Jurka and Klonowski 1996; Tatout et al. 1998), so that all organisms carrying a particular SINE insertion are derived from a unique event that occurred in their common ancestor. Consequently, SINE insertions have often been used to retrace phylogenetic relationships in animals (Murata et al. 1993; Batzer et al. 1994; Shimamura et al. 1997; Salem et al. 2003) and plants (Tatout et al. 1999; Cheng et al. 2003). The SINE markers are therefore well suited to detect unambiguously gene flow between closely related species and thus could be developed as a valuable tool for assessing the environmental impact of cultivated transgenic species on wild species.

Gene flow has already been demonstrated between oilseed rape (*B. napus*, AACC, $2n=38$), which is a partially allogamous species, and one of its progenitors, *Brassica rapa* L. (AA, $2n=20$) (see Chèvre et al. 2004 for review). Among the other wild relatives known to be frequent in oilseed rape fields, wild radish (*Raphanus raphanistrum* L., RrRr, $2n=18$) seems to be the most relevant weed to examine because of the frequency of interspecific hybridization under optimal conditions (Eber et al. 1994; Baranger et al. 1995; Darmency et al. 1998) and the possibility of gene exchange through recombination (Kerlan et al. 1993). Spontaneous hybridization between wild radish and oilseed rape has been reported under agronomic conditions (Chèvre et al. 2000; Rieger et al. 2001; Warwick et al. 2003), but no introgression of oilseed rape transgene within the wild radish genome has been demonstrated. However, we do not know whether the transgene could have been introgressed into the wild radish genome if it had been located in another region of the oilseed rape genome that is more prone to recombine. Crop-specific markers evenly distributed on the oilseed rape genome are needed to address this question.

In the study reported here, we used the S1 SINE to develop a set of molecular markers distinguishing *B. napus* from *R. raphanistrum*. Once mapped on the oilseed rape genome, we were able to use these markers as efficient tools to follow, from different genomic regions of the crop, the occurrence of crop gene introgression into the wild radish genome.

Materials and methods

Plant material

SINE markers were developed from a Canadian variety of *Brassica napus*, *Westar*, and from *Raphanus raphanistrum* populations collected in France and Canada. The oilseed rape variety and the Canadian wild radish population were provided by S. Warwick [Agriculture and Agri-Food Canada (AAFC)]. The polymorphism analyses were performed on seven oilseed rape varieties: three French winter types (*Darmor-bzh*, *Darmor* and *Samourai*), three spring types (*Westar*, *Stellar* and *Drakkar*) and a Korean variety, *Yudal*. Three double-haploid (DH) populations (62 individuals per population): DY (*Darmor-bzh* × *Yudal*), DS (*Darmor* × *Samourai*) and SD (*Stellar* × *Drakkar*), previously used to establish the oilseed rape map (Lombard

and Delourme 2001), were used to map the markers polymorphic between the parents of the segregating populations.

DNA extraction

For the identification of SINE markers, DNA was extracted from individual *B. napus* or *R. raphanistrum* seedlings by grinding the plant tissue in a small mortar containing 200 µl of 10 m *M* Tris-HCl (pH 8), 2 m *M* EDTA and 400 m *M* NaCl buffer and subsequently transfered to a 1.5-ml Eppendorf tube. Two microliters of RNAse (10 mg/ml) were added, and the resulting solution was incubated 15 min at 37°C. Five microliters of sodium dodecyl sulfate (SDS, 20%) was then added, followed by a 15-min incubation at 65°C. To eliminate cellular debris, we added 200 µl of 5 *M* NaCl to the solution, followed by centrifugation for 5 min (12,000 *g*). DNA was then precipitated with ethanol, dried and dissolved in 50 µl water. When necessary, an equivalent amount of DNA from 50 different individuals was pooled to generate a sample representative of each population.

For the polymorphism analyses and segregation studies, DNA was extracted from young leaves according to the method described by Doyle and Doyle (1990).

Sequence-specific amplified polymorphism (SSAP), inverse-PCR and PCR reactions

We first used an SSAP approach [i.e. a multiplex amplified fragment length polymorphism (AFLP)-like technique] to detect individual SINE insertions as a band on a sequencing gel. The SSAP reactions were performed as described previously (Ellis et al. 1998). Genomic DNA was digested either with the *Csp*6I, *Dpn*II or *Taq*I restriction enzymes, and the corresponding adapter was ligated to the fragments (see Ellis et al. 1998 for the design of adapters and adapter primers). A [P^{32}]-labeled SINE derived primer (5′-CCACTGGACTACGAGGTCC-3′), corresponding to positions 6 to 22 of the S1 consensus sequence, was used in combination with the adapter primer in the PCR reaction. Selected bands, corresponding to SINE insertions present in most or all of the *B. napus* individuals but absent in *R. raphanistrum* individuals, were excised out of the polyacrylamide gel and placed in TE buffer (10 m *M* Tris-HCl, pH 8 and 1 m *M* EDTA) for 16 h at 4°C to recover DNA. These fragments were cloned (using pGEM-T-Easy vector; Promega, Madison, Wis.) and sequenced. Inverse-PCR and PCR reactions were performed under standard conditions (Sambrook and Russell 2001). To determine the best conditions for the PCR amplification of SINE markers, we used three primer combinations for each marker: the first used two primers flanking the SINE (external primers), the second used one primer upstream of the SINE and one primer internal to the SINE (same primer as the one used for the SSAP reaction, see above) and the third used one primer downstream of the SINE and one primer internal to the SINE (5′-AGGACCTCGTAGTCCAGTGG-3′). A temperature gradient that started at the melting temperature (T_m) of the oligonucleotides, −15°C, and extended up to a T_m of +2°C (Robocycler 96; Stratagene, La Jolla, Calif.) was also used with each primer combination to determine the optimal annealing temperature of the primers in the PCR reactions. Two primers internal to the SINE sequence were also used as a positive control of amplification on each DNA sample.

For the polymorphic marker analyses and mapping, SINE amplifications were performed in 17-µl reaction mixes containing 20 ng DNA, forward and reverse primers (0.2 µ*M* each), 1.5 m *M* $MgCl_2$, 200 µ*M* dNTPs and 0.6 U *Taq* Polymerase (Eurobio, Courtaboeuf, France). The amplification conditions were 4 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at the annealing temperature (depending of the primer pair, see Table 1) and 30 s at 72°C, followed by 5 min at 72°C. Amplification products were separated on 2.5% agarose gels buffered with 1× TBE and visualized by UV fluorescence after ethidium bromide staining or on 6% acrylamide gels buffered with 1× TBE and revealed by silver nitrate staining.

For monomorphic bands among the varieties, we looked for single nucleotide polymorphism (SNP) between *Darmor-bzh* and *Yudal* only. The amplified bands were excised from dried acrylamide gel with a sterile scalpel, suspended in 50 µl ultra-pure, sterile water and submitted to four freezing/defrosting cycles. The reamplified products were sequenced by Genome Express (Meylan, France).

Genetic mapping

Polymorphic markers were located on previously established genetic maps (Lombard and Delourme 2001) using MAPMAKER/EXP (Lincoln et al. 1992).

Results

Identification of SINE markers

For a given restriction enzyme/adapter combination, the SINE-SSAP reaction generated from 50 to 75 bands that could be separated on a sequencing gel (Fig. 1). Between 30% and 40% of the bands were found in both *B. napus* and *R. raphanistrum*. Of the species-specific bands, most were present in a small number of individuals, and only a small proportion (around 20%) were found in all or most of the individual plants tested. Only *B. napus*-specific bands present at high frequency were considered. Sixty-seven SINE insertions that clearly distinguished *B. napus* var. *Westar* from two *R. raphanistrum* populations (Fig. 1) were identified using three

Table 1 Polymorphic *Brassica napus*-specific SINE markers, obtained from primer pairs (F, forward; R, reverse) designed either on upstream (5′), internal to SINE (Int) or downstream (3′) sequences

Marker	Primer name	Primer Sequence (5′ → 3′)	Primer type	T_m (°C)[a]	Size in basepairs (a1/a2)[b]	*Brassica napus* varieties[c]							
						W	Dbzh	Y	D	Sam	St	Dk	DH populations[c]
JLP001	JLP001F	AgCAACgATTAATCATCTCC	5′	55	390/211	a1	a2	a1	a2	a2	a1	a1	DY
	JLP001R	CTTACAATTACATTggCACg	3′										
JLP002	JLP002F	GgAAAACgAATATCAAAATggg	5′	50	375/190	a1/a2	a1	a1	a1	a1	a1	a1	
	JLP002R	AAggCACCATAAAgCATTAgC	3′										
JLP003	JLP003F	AggACCTCgTAgTCCAgTgg	Int	50	170	a1	a1	a1	a1	a1	a1	a1	
	JLP003R	CCgTTAgTTCCTCgTTACTg	3′										
JLP004	JLP004F	TTgTCAACCATAggTCTT	5′	50	360/170	a1	a1	a2	a1	a2	a1	a1	DS
	JLP004R	TCCCCTTTTATCCTgCTg	3′										
JLP005	JLP005F	gCAgTTTAgACTTTTAggCg	5′	50	270/260	a1	a1	a2	a1	a2	a1	a1	DY
	JLP005R	AgAAAAgAgACgTTAgCTCC	3′										
JLP006	JLP006F	CAgTTCAgCATATTgTgAg	5′	50	260	a1	a1	a1	a1	-	a1	a1	DS
	JLP006R	ACACAAgCCTATATTgAgC	3′										
JLP007	JLP007F	AgCTAgCgTgCTATCAATATC	5′	55	320/550	a1	a1	a1	a1	a1	a1	a2	SD
	JLP007R	ACCTTCgACTTAgACggACAgg	3′										
JLP008	JLP008F	ATTgCCTTggTTTAgTTggg	5′	55	410	a1	a1	a1	a1	a1	a1	a1	
	JLP008R	gggTgCAAgTTTTAAgAAgg	3′										
JLP009	JLP009F	GTACATTATTTggACAgTTTAC	5′	50	340/150	a1	a1	a1	a1	a1	a1	a2	SD
	JLP009R	AACgAgCCTAAACgCCTAC	3′										
JLP010	JLP010F	AggACCTCgTAgTCCAgTgg	Int	55	240/260	a1	a1	a2	a1	a1	a1	a1	DY
	JLP010R	ggTTAAAggTATTATgAgAcg	3′										
JLP011	JLP011F	gCTggCgCCgggCCTAgg	Int	55	210/220	a1	a1	a2	a1	a1	a1	a1	DY
	JLP011R	CATTACCgCATATgTTgCAgg	3′										
JLP012	JLP012F	CCTTgCAATAAgTgTgTTCTA	5′	55	225	a1	a1	a1	a1	a1	a1	a1	
	JLP012R	TTCCCAATTTTTCCACCACC	3′										
JLP013	JLP013F	AggACCTCgTAgTCCAgTgg	Int	50	260	a1	a1	a1	a1	-	a1	a1	DS
	JLP013R	gATCATgCACCAAgTTCgAAg	3′										
JLP015	JLP015F	CCAAgTTCCATCgTACgA	5′	55	370	a1	-	a1	-	-	a1	a1	DY
	JLP015R	ACCgTgTTTCAATCATTg	3′										
JLP016	JLP016F	TAgCgTAgAATTACTCATCA	5′	50	620	a1	a1	a1	a1	a1	a1	a1	
	JLP016R	CCACAAgTCTTATCTgTATCC	3′										
JLP017	JLP017F	CTgACAAgATTgTATCATCC	5′	50	570	a1	a1	a1	a1	a1	a1	a1	
	JLP017R	CTgAgCgATCATggCTCTT	3′										
JLP019	JLP019F	GTCCAgggCCCgTAAgAAgg	Int	50	210	a1	-	-	-	a1	-	a1	DS
	JLP019R	CATgTATgCCTATAAATgATTC	3′										
JLP020	JLP020F	AggACCTCgTAgTCTAgTgAT	Int	50	290	a1	a1	a1	a1	a1	a1	a1	
	JLP020R	AgTgACggCACTggCgg	3′										
JLP021	JLP021F	AggACCTCgTAgTCCAgTgg	Int	50	250	a1	-	a1	-	-	a1	a1	DY
	JLP021R	ATAAAgACTAAAgAgATATCAg	3′										
JLP022	JLP022F	ATAATCAgCgTCACACTgg	5′	50	210	a1	a1	a1	a1	a1	a1	a1	
	JLP022R	CCACTggACTACgAggTCC	Int										
JLP024	JLP024F	TgTTTgATTAACAgAgAAAgTgCA	5′	60	310/120	a1	a1	a1	a1	a1	a1	a2	SD
	JLP024R	ACgAgCCTAAACgCCTACACCA	3′										
JLP025	JLP025F	TCggACAACTAATggTTACAAAACTCT	5′	50	180	a1	a1	a1	a1	-	a1	a1	DS
	JLP025R	CTAggCCCggCgCCAgC	Int										
JLP026	JLP026F	TTCCAggTgTTgCTTgTTTCgCA	5′	55	190	a1	a1	a1	a1	a1	a1	a1	
	JLP026R	CCACTggACTACgAggTCC	Int										
JLP028	JLP028F	TCAACTCggCCgCCTAAC	5′	55	450/120	a1	a1	a1	a1	a1	a2	a1	SD
	JLP028R	gAATATTTACgCACTATCACg	3′										
JLP030	JLP030F	AggACCTCgTAgTCCAgTgg	Int	50	300	a1	a1	a1	a1	a1	-	a1	SD
	JLP030R	TTTCCgTgCCATCCTgACA	3′										
JLP031	JLP031F	AggACTTCTATAgTCCAgTgg	Int	48	200	a1	-	a1	-	-	a1	-	DY
	JLP031R	gAATATTACAggTTTgTgTCC	3′										
JLP032	JLP032F	gCTAgCCTTATCAACACCATgCT	5′	48	120	a1	a1	a1	a1	a1	a1	a1	
	JLP032R	CCACTggACTACgAggTCC	Int										
JLP033	JLP033F	TgTTCgATCgAgAAAAgCAg	5′	50	110	a1	a1	a1	a1	-	a1	a1	DS
	JLP033R	gAggTAgTACCACTggACTACA	Int										
JLP034	JLP034F	gggACCTTgTAgTCCAgTgg	5′	50	190	a1	a1	a1	a1	-	a1	a1	DS
	JLP034R	AtgAgACgTTAgCTCCAC	int										
JLP035	JLP035F	ggACCTCgTAgTCTAgTgg	5′	50	230	a1	a1	a1	a1	a1	a1	a1	
	JLP035R	AgACACCTTgCCggCCACA	Int										
JLP037	JLP037F	TCTCTTATgCAAggTAAgCT	5′	50	230	a1	-	a1	-	-	a1	a1	DY
	JLP037R	CgCCTgCgAATCgAACAgC	Int										

Table 1 (Contd.)

Marker	Primer name	Primer Sequence (5′ → 3′)	Primer type	T_m (°C)[a]	Size in basepairs (a1/a2)[b]	*Brassica napus* varieties[c] W	Dbzh	Y	D	Sam	St	Dk	DH populations[c]
JLP038	JLP038F	ACCCAggACCTCgTAgTCC	Int	50	295	a1	a1	a1	a1	a1	a1	a1	
	JLP038R	CgATATTCTTTgCATCATCCT	3′										
JLP039	JLP039F	ACCTTCCTCTACggCgCTTg	5′	50	310	a1	a1	a1	a1	a1	a1	a1	
	JLP039R	TCCTCAACTTACATCATggA	3′										
JLP040	JLP040F	gATCTTCATCTTATTTgTCgg	5′	45	350/160	a1	a1	a2		a1	a2	a1	DY;SD
	JLP040R	AgCTCTgACTTTTCTTAggC	3′										
JLP042	JLP042F	CTTTgCTgTAATATCTCCg	5′	48	350/160	a1	a1	a2	a1	a1	a2	a1	DY
	JLP042R	gAgACTTCCAgAAACTTCA	3′										
JLP044	JLP044F	CTTATgAAgTATCAgACCCACT	5′	50	160	a1	a1	a1	a1	a1	a1	a1	
	JLP044R	CCACTggACTACgAggTCC	Int										
JLP046	JLP046F	TTAATgATCAACTggCgATC	5′	50	250	a1	a1	a1	a1	a1	a1	a1	
	JLP046R	CCACTggACTACgAggTCC	int										
JLP047	JLP047F	CTCCATTggTTTTTTCAACTTAT	5′	50	300/120	a1	a1	a1	a1	a1	a2	a1	SD
	JLP047R	TTCTAACTgAACCAgACAgCC	3′										
JLP050	JLP050F	AggACCTCgTAgTCCAgTgg	Int	50	290	a1	a1	a1	a1	a1	a1	-	SD
	JLP050R	gTTgATCgAgTCCCTATgAg	3′										
JLP051	JLP051F	AATTCCCgTTCTTTCgggC	5′	50	190	a1	a1	a1	a1	a1	a1	a1	
	JLP051R	CCACTggACTACgAggTCC	Int										
JLP052	JLP052F	TAAATAgAAAAATggACCCCgg	5′	50	150	a1	-	a1	-	a1	-	-	DY
	JLP052R	CCACTggACTACgAggTCC	Int										
JLP053	JLP053F	AggACCTCgTAgTCCAgTgg	Int	50	300	a11	a1	-	a1	-	-	a1	DY
	JLP053R	gTTTAAAgACCCCCTTTAACg	3′										
JLP054	JLP054F	AggACCTCgTAgTCCAgTgg	Int	50	220	a1	a1	a1	a1	a1	a1	a1	
	JLP054R	TTTgTCTTACgCAACACCATg	3′										
JLP061	JLP061F	ACgCACTTCTCTTTTTCACg	5′	50	195	a1	-	a1	-	-	-	-	DY
	JLP061R	CCACTggACTATCgAgTCC	Int										
JLP065	JLP065F	GggCTTgATAgTCCAgTgg	Int	50	320	a1	-	-	-	-	a1	a1	
	JLP065R	TgAgATATTACATTCgAgATgTC	3′										
JLP066	JLP066F	gTCTgTgAATgATTgAgTTAgg	5′	50	200	a1	a1	a1	a1	-	a1	a1	DS
	JLP066R	CCACTggACTACgAggTCC	Int										
JLP067	JLP067F	AggACCTCgTAgTCCAgTgg	Int	50	240	a1	a1	a1	a1	a1	a1	a1	
	JLP067R	ACATCTgAAAgTCTCTCCTAg	3′										

[a]The optimal temperature of melting (Tm) is indicated
[b]An approximation of the molecular weight of the different PCR products is shown. The a1 product corresponds to a full site and the a2 product is the result of the amplification of an "empty" (or pre-integration) site with the exception of JLP005, JLP007, JLP010 and JLP011 where a2 represents a SINE-containing product of a different size than expected
[c]The presence of the above-mentioned markers in *B. napus* varieties: Westar (W), *Darmor-bzh* (Dbzh), *Yudal* (Y), *Darmor* (D), *Samouraï* (S), *Stellar* (St) and *Drakkar* (Dk) is shown. Polymorphic markers were mapped from the double-haploid (DH) populations: *Darmor-bzh* × *Yudal* (DY), *Darmor* × *Samouraï* (DS) or *Stellar* × *Drakkar* (SD)

different restriction enzyme/adapter combinations. These selected *B. napus*-specific bands were excised out of the gel, cloned and sequenced to obtain the genomic sequence upstream of each SINE element. An inverse PCR strategy was then carried out to characterize the corresponding downstream genomic region in each case. Accordingly, each selected *B. napus*-specific band on the SSAP gels was converted into a single locus marker that can be revealed by PCR; two examples are presented in Fig. 1. Forty-seven markers were derived by this approach (Table 1).

We then checked that all markers differentiated *B. napus* from *R. raphanistrum* individuals. To do so, we prepared three DNA pools using 50 individuals from either the *B. napus* var. *Westar* population or from the two *R. raphanistrum* populations (from France and Canada). Using these samples, we amplified by PCR each marker individually and verified that in all cases, and for various PCR conditions, amplification products were only observed in the positive control PCR amplification (not shown) and when the *B. napus* sample was used (Fig. 1). For each marker, two additional primer combinations, one using a primer upstream of the SINE in association with a primer internal to the SINE sequence and a second one using a primer downstream of the SINE in association to a second internal primer (see Table 1), were also tested. The same results were obtained; PCR products of the expected sizes were detected with the *B. napus* sample, but no product was amplified using the *R. raphanistrum* samples (data not shown). We therefore concluded that each SINE marker derived by our approach was able to discriminate *B. napus* from *R. raphanistrum* individuals. Surprisingly, we never amplified the corresponding orthologous "empty site" site in *R. raphanistrum* using external primer combinations (see Discussion).

The PCR conditions were optimized for each marker, and the most robust primer combination was selected (Table 1). Primers anchored in flanking 5′-and 3′-sequences were selected for 18 markers, 14 other

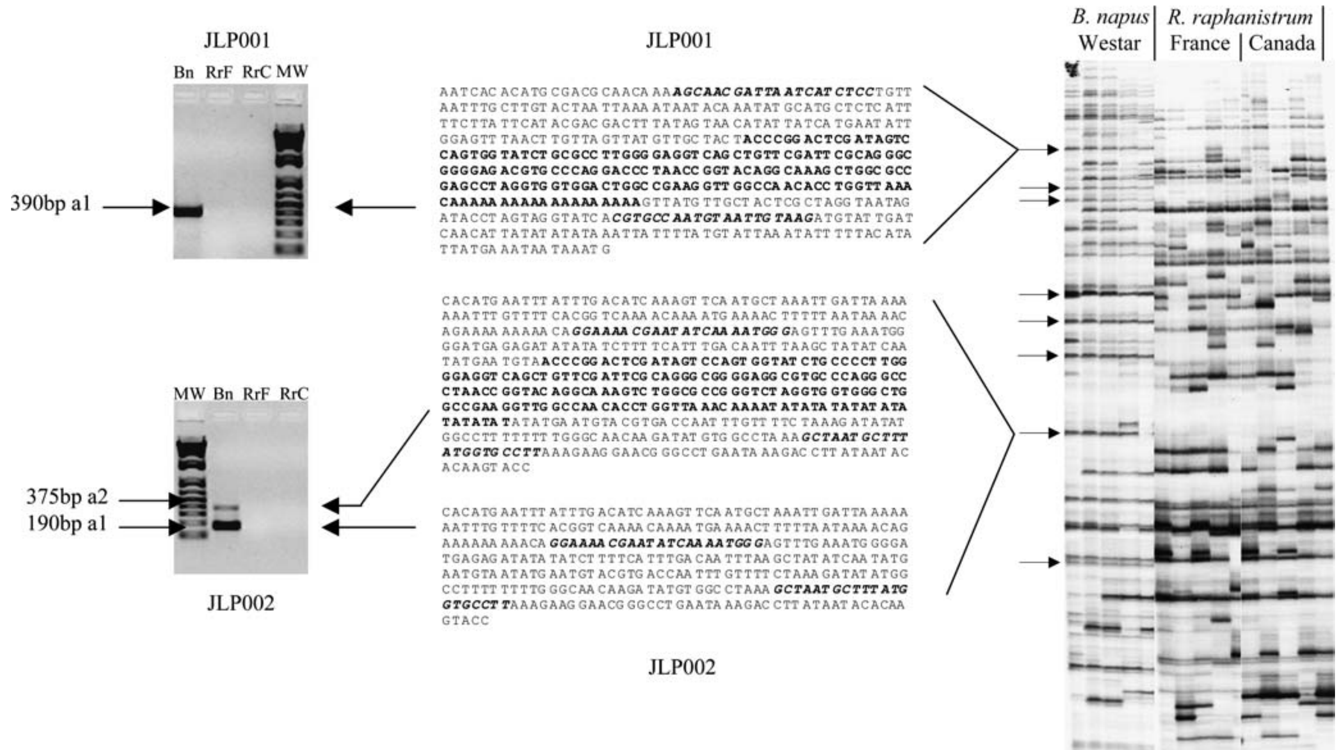

Fig. 1 Strategy adopted to produce the *Brassica napus*-specific SINE markers. Multilocus SINE-based SSAP analyses were initially carried out on DNA from individuals originating from the *B. napus* var. *Westar* population and from two *Raphanus raphanistrum* wild populations (from France and Canada). **Right panel** Representative SSAP result for five individuals from the three populations studied using *Csp6I* as the restriction enzyme. The SSAP bands representing SINE loci present in most or all *B. napus* individuals but absent in *R. raphanistrum* plants were selected (see *arrows*), excised out of the polyacrylamide gel, cloned and sequenced in order to obtain the genomic sequence upstream of the SINE (**middle panel**). Using this information, we performed inverse PCR reactions in each case to obtain the corresponding 3′ genomic sequences. Sequences for the two markers, JLP001 and JLP002, are presented (SINE sequences are in *bold*). Primers flanking the SINE elements were subsequently designed (*bold* and *italic*) to specifically and individually amplify each SINE locus. **Left panel** As examples, PCR products obtained following amplification of the JLP001 (*top*) and JLP002 (*bottom*) markers using a pooled DNA from sample of 50 individuals from each population are shown. *Bn* : *B. napus* var. *Westar*, *RrF*: *R. raphanistrum*, French population, *RrC*: *R. raphanistrum*, Canadian population, *MW*: molecular-weight marker. For the JLP001 marker, the SINE element is apparently fixed in the *B. napus Westar* population as a single PCR product of the expected size (390 bp). For the JLP002 marker, two PCR products were obtained, one corresponding to the SINE-containing allele (375 bp) and one corresponding to an "empty" (or pre-integrative) allele (190 bp) (see text). Therefore, the JLP002 marker probably results from a very recent SINE integration event and is not fixed in the *B. napus Westar* population; the two PCR products are only detected when using DNA sample from the *B. napus* population

markers were developed using primers targeting the 5′ and SINE internal sequences and 15 markers were generated using primers targeting the SINE internal and flanking 3′ sequences (Table 1). The efficiency of each marker to detect a *B. napus* locus within a *R. raphanistrum* population was tested by checking that the *B. napus* locus was detectable from a bulk containing DNA from 49 *R. raphanistrum* plants and one plant of *B. napus*; each plant was represented by 10 ng of DNA.

Polymorphism analyses among *B. napus* varieties

Among the 47 markers, 17 were present in all of the oilseed rape varieties analyzed (Table 1, a1 in all varieties). For 12 of these 17 monomorphic markers, bands amplified from *Darmor-bzh* and *Yudal* were sequenced, and these displayed no polymorphism between these two varieties; consequently, their locations on the oilseed rape map remain unknown.

Polymorphism between the oilseed rape varieties was detected directly for the 30 remaining markers (Table 1). For 13 of these, a second PCR product of lower molecular weight was amplified in at least one *B. napus* population (see Fig. 1, JLP002 for an example). For four of these loci (JLP001, JLP002, JLP004 and JLP009), we used sequencing to check that this second product corresponded to the amplification of an "empty" site (also called a pre-integrative site). The presence of these allelic "empty sites" was expected for recent SINE integration events that are not yet fixed in all *B. napus* populations (Tatout et al. 1999). In four cases (JLP005, JLP007, JLP0010 and JLP011), a second product was obtained with the primer combination in Table 1 that was larger than expected and therefore did not correspond to an "empty" site. Based on PCR results using the two other primer combinations (not

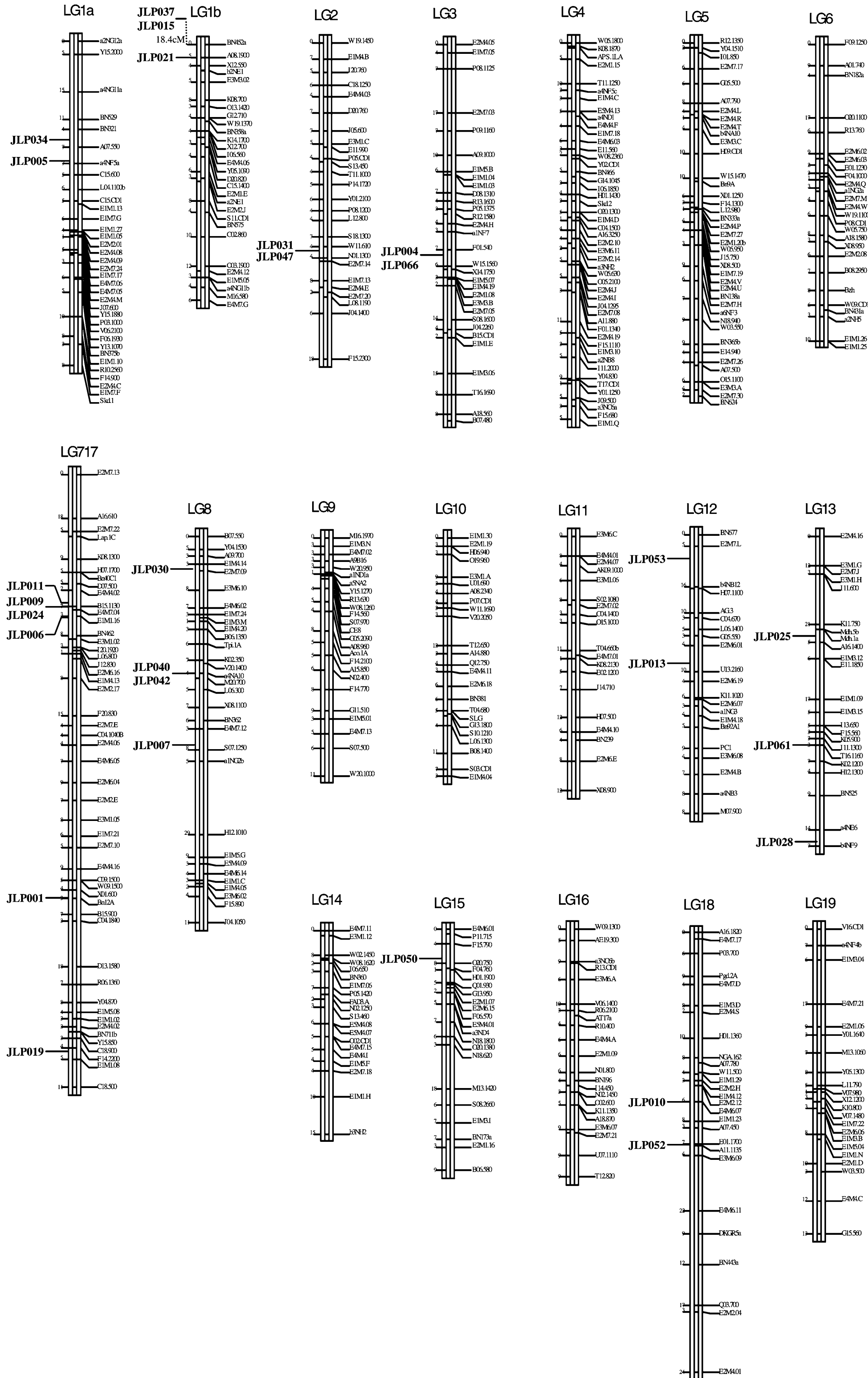
LG1a
JLP034
JLP005
JLP037
JLP015
18.4cM
JLP021
LG1b
LG2
JLP031
JLP047
LG3
JLP004
JLP066
LG4
LG5
LG6
LG717
JLP011
JLP009
JLP024
JLP006
JLP001
JLP019
LG8
JLP030
JLP040
JLP042
JLP007
LG9
LG10
LG11
LG12
JLP053
JLP013
LG13
JLP025
JLP061
JLP028
LG14
LG15
JLP050
LG16
LG18
JLP010
JLP052
LG19

Fig. 2 Location of 27 SINE markers on the oilseed rape genetic map. Linkage groups LG 1b, LG 4, LG 5, LG 6, LG 9, LG 10, LG 11, LG 14, LG 16, LG 19 correspond to the A genome, LG 1a, LG 2, LG 3, LG 8, LG 12, LG 13, LG 15, LG 717, LG 18 correspond to the C genome

shown), these bands corresponded to SINE-containing sites with a short insertion in the 5′ or 3′ flanking region.

The 17 remaining markers were dominant. The inability to amplify a PCR product for a given locus was either the result of the use of a SINE internal primer on DNA samples from varieties having only an ''empty'' site (15 cases) or due to the absence of primer hybridization with primers designed for the external borders (two cases) as a result of modifications to the flanking sequences.

Mapping of polymorphic markers

Two of the 30 polymorphic markers (JLP002 and JLP065) were not mapped because they were not polymorphic between the genotypes used to develop the DH mapping populations. The three DH populations were genotyped with the 28 remaining markers (Fig. 2). Twenty-four dominant (Fig. 3a) or co-dominant (Fig. 3b) markers were mapped on linkage groups (LGs) assigned to the C genome (Delourme, personal communication), whereas only one linkage group of the A genome (LG1b) carried three markers (JLP015, -021 and -037). Only one marker (JLP033) remained unassigned. Additional markers linked to JLP028 are needed to precisely map this marker as it is carried by a genomic region that might have translocated in some oilseed rape varieties (Lombard and Delourme 2001). Four clusters were identified on LG 1b, LG 3, LG 8 and LG 717, but all the other markers were well separated on the different chromosomes.

From the markers identified, it will be possible to follow 19 regions located on 10 of the 19 oilseed rape linkage groups. However, only nine of these will be detectable whatever oilseed rape varieties used as they carry co-dominant markers. The other nine will be useful only when all of the oilseed varieties cultivated nearby the wild radish populations carry the dominant markers.

Discussion

In the present study, we used the SINE S1 elements to develop a set of molecular markers that are able to distinguish *B. napus* from *R. raphanistrum* and can be used to assess the occurrence and frequency of introgressions of oilseed rape genomic regions within wild radish populations.

B. napus and *R. raphanistrum* are very closely related species belonging to the subtribe Brassicinae (Warwick and Black 1993) that both contain SINE S1 elements. However, we knew from previous studies (Lenoir et al. 1997) that quantitatively significant bursts of SINE S1 amplification arose independently (and therefore recently) in different species of the Brassicinae. Using the SSAP approach (Waugh et al. 1997), also called transposon display (Purugganan and Wessler 1995), we were able to identify those insertions that arose specifically in the *B. napus* lineage after its separation from the *R. raphanistrum* lineage. Since these events are necessarily very recent on an evolutionary time scale, it is not surprising to observe that nearly two-thirds of the markers we derived are not fixed in all of the *B. napus* populations tested. What is more surprising is that we

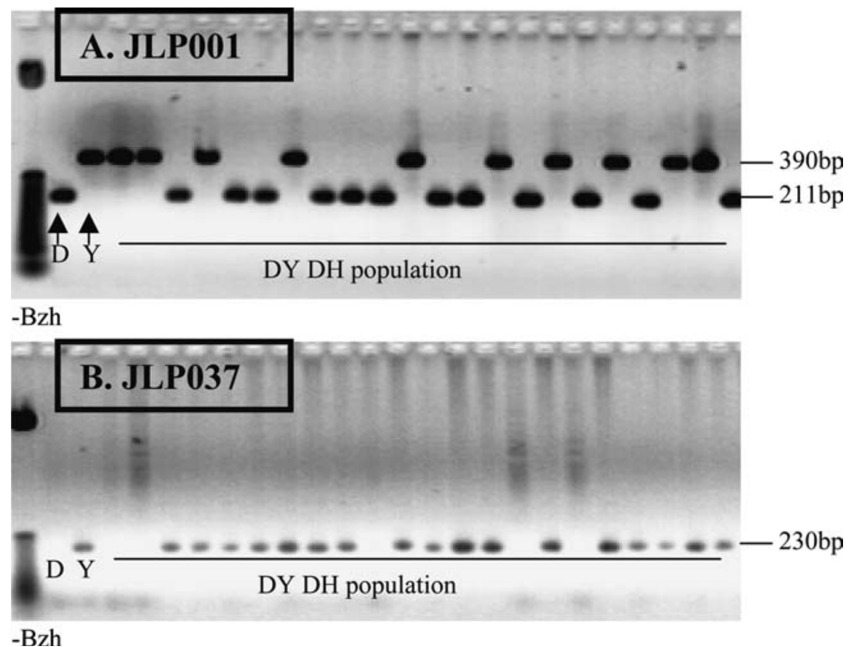

Fig. 3 Profile of amplified products from a DH oilseed rape population, DY [*Darmor-bzh* (*D*) × *Yudal* (*Y*)] for a co-dominant marker, JLP001 (**A**) and a dominant marker, JLP037 (**B**)

never succeeded in amplifying orthologous ‘‘empty’’ sites in *R. raphanistrum*. The SINEs reside in intergenic regions, and although these regions are less conserved than gene-coding regions, one would still expect to be able to amplify in many situations the orthologous *R. raphanistrum* ‘‘empty’’ site using the two external primers designed in *B. napus*. We observed recently in different *Arabidopsis thaliana* ecotypes that SINE-containing loci are often associated with small deletions in the SINE element and genomic regions flanking the SINE element (Lenoir et al. 2005). Therefore, the integration of a SINE element can rapidly lead to significant modifications of the upstream and downstream genomic regions. In our study, the rapid modification of genomic sequences flanking SINEs in the *B. napus* lineage could explain our inability to amplify ‘‘empty’’ orthologous sites of *R. raphanistrum* or some of *B. napus* loci in the case of dominant markers.

It is also intriguing that most of the SINE markers that were placed on the *B. napus* genetic map are located on the C genome (Fourman et al. 2002; Delourme, personal communication). We can possibly explain this observation in two ways. First, based on what we know of SINE S1 sequences in *B. rapa* (parental A genome) and *B. oleracea* (parental C genome), we can expect the internal SINE primer we used in the SSAP reaction to anneal more efficiently with SINE S1 from the parental C genome. In support of this, we observed previously that the two major SINE S1 subfamilies present in *B. rapa* are surprisingly quite divergent from the major SINE S1 subfamilies present in *B. oleracea* and that *B. rapa* SINEs are more related to SINEs from species of a more distant lineage, the *B. nigra* lineage (Lenoir et al. 1997). Second, although a small SINE S1 subfamily specific to *B. napus* was previously identified (Deragon et al. 1994, 1996; Lenoir et al. 1997), it is likely that SINE S1 retroposition was much more active in the *B. oleracea* parental lineage than in *B. napus* itself. Therefore, most if not all SINE loci described in this work are probably the result of retroposition events that took place in the parental *B. oleracea* lineage and not in *B. napus*.

It is also interesting to note that the only A linkage group (LG 1b) carrying SINE markers in a telomeric region is among those most homologous to a C linkage group (Parkin et al. 1995, Delourme, personal communication). It can be postulated that this A linkage group carries a translocation from the C genome (Udall et al. 2005).

The specific markers that we have developed will be efficient tools for assessing if crop gene introgression has occurred and is stabilized into the wild radish genome. Monomorphic markers will enable the detection of gene flow, although their genomic location will remain unknown. For dominant markers, their presence only will provide a clue for introgression. In all cases, co-dominant markers will efficiently track the introgression of large oilseed rape genomic regions within the wild radish genome. It has been shown that every time a cultivated radish genomic region is inserted into the oilseed rape genome no recombination occurs between the introgressed fragment and the recipient genome. Additionally, a large region of the oilseed rape genome species is generally replaced by the wild radish genome (Delourme et al. 1998). The same observation is expected in the reciprocal situation—i.e., oilseed rape genomic regions within the wild radish genome. Consequently, to confirm stabilized introgression from homozygous plants, it would be possible to check if the oilseed rape markers flanking the SINE marker are present and if wild radish markers have disappeared in the same region. The use of such markers will ascertain if some genomic regions are more susceptible to introgression than others, taking into account the initial location of these markers on the oilseed rape genome and the relative homology between the two species. If the number of specific oilseed rape markers were enhanced to get a better coverage of the oilseed rape genome, it would be possible to determine if some transgene insertion sites are safer than others, i.e. show a lower probability of gene flow from the crop to its wild relative oilseed rape.

In future studies, wild radish populations which have never been in contact with oilseed rape fields or which co-occur with oilseed rape fields from a relatively recent (Canada) to longer periods of time (France) will be collected and analyzed with the markers described herein to detect the occurrence and the frequency of oilseed rape genomic introgression.

Acknowledgements The work was supported by a post-doc grant of the National Institute of Agronomic Research and was funded by the French Research Ministry (Contract no. C002001).

References

Baranger A, Chèvre AM, Eber F, Renard M (1995) Effect of oilseed rape genotype on the spontaneous hybridization rate with a weedy species: an assessment of transgene dispersal. Theor Appl Genet 91:956–963

Batzer MA, Stoneking M, Alegria-Hartman M, Bazan H, Kass DH, Shaikh TH, Novick GE, Ioannou PA, Scheer WD, Herrera RJ, Deininger PL (1994) African origin of human-specific polymorphic Alu insertions. Proc Natl Acad Sci USA 91:12288–12292

Capy P, Bazin C, Higuet D, Langin T (1998) Dynamics and evolution of transposable elements. Springer RG Landes, Austin

Cheng C, Motohashi R, Tsuchimoto S, Fukuta Y, Ohtsubo H, Ohtsubo E (2003) Polyphyletic origin of cultivated rice: based on the interspersion pattern of SINEs. Mol Biol Evol 20:67–75

Chèvre AM, Eber F, Darmency H, Fleury A, Picault H, Letanneur JC, Renard M (2000) Assessment of interspecific hybridization between transgenic oilseed rape and wild radish under normal agronomic conditions. Theor Appl Genet 100:1233–1239

Chèvre AM, Ammitzboll H, Breckling B, Dietz-Pfeilstetter A, Eber F, Fargue A, Gomez-Campo C, Jenczewski E, Jorgensen R, Lavigne C, Meier MS, den Nijs H, Pascher K, Seguin-Swartz G, Sweet J, Stewart CN Jr, Warwick S (2004) A review on interspecific gene flow from oilseed rape to wild relatives. In: Den Nuijs HCM, Bartsch D, Sweet J (eds) Introgression from genetically modified plants into wild relatives. CABI Publ, Wallingford, pp 235–251

Cook J, Tristem M (1997) SINEs of the times–transposable elements as clade markers for their hosts. Trends Ecol Evol 12:295–297

Darmency H, Lefol E, Fleury A (1998) Spontaneous hybridization between oilseed rape and wild radish. Mol Ecol 7:1467–1473

Delourme R, Foisset N, Horvais R, Barret P, Champagne G, Cheung WY, Landry BS, Renard M (1998) Characterisation of the radish introgression carrying the *Rfo* restorer gene for the Ogu-INRA cytoplasmic male sterility in rapeseed (*Brassica napus* L). Theor Appl Genet 97:129–134

Deragon JM, Landry BS, Pelissier T, Tutois S, Tourmente S, Picard G (1994) An analysis of retroposition in plants based on a family of SINEs from *Brassica napus*. J Mol Evol 39:378–386

Deragon JM, Gilbert N, Rouquet L, Lenoir A, Arnaud P, Picard G (1996) A transcriptional analysis of the S1Bn (*Brassica napus*) family of SINE retroposons. Plant Mol Biol 32:869–878

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Eber F, Chèvre AM, Baranger A, Vallée P, Tanguy X, Renard M (1994) Spontaneous hybridization between a male-sterile oilseed rape and two weeds. Theor Appl Genet 88:362–368

Ellis TH, Poyser SJ, Knox MR, Vershinin AV, Ambrose MJ (1998) Polymorphism of insertion sites of Ty1-copia class retrotransposons and its use for linkage and diversity analysis in pea. Mol Gen Genet 260:9–19

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

Gepts P, Papa R (2003) Possible effects of (trans)gene flow from crops on the genetic diversity from landraces and wild relatives. Environ Biosafety Res 2:89–103

James C (2004) Global status of commercialized Biotech/GM Crops:2004. ISAAA Briefs 32

Jenczewski E, Ronfort J, Chèvre AM (2003) Crop-to-wild gene flow, introgression and possible fitness effects of transgenes. Environ Biosafety Res 2:9–24

Jurka J, Klonowski P (1996) Integration of retroposable elements in mammals: selection of target sites. J Mol Evol 43:685–689

Kerlan MC, Chèvre AM, Eber F (1993) Interspecific hybrids between a transgenic rapeseed (*Brassica napus*) and related species: cytological characterization and detection of the transgene. Genome 36:1099–1106

Lenoir A, Cournoyer B, Warwick S, Picard G, Deragon JM (1997) Evolution of SINE S1 retroposons in Cruciferae plant species. Mol Biol Evol 14:934–941

Lenoir A, Pélissier T, Bousquet-Antonelli C, Deragon JM (2005) Comparative evolution history of SINEs in *Arabidopsis thaliana* and *Brassica oleracea*: Evidence for a high rate of SINE loss. Cytogenet Cell Genet 110 (in press)

Lincoln S, Daly M, Lander E (1992) Constructing genetic linkage maps with MAPMAKER/EXP 3.0: a tutorial and reference manual, 3rd edn. Whitehead Institute Technical Report, Whitehead Technical Institute, Cambridge, Mass.

Lombard V, Delourme R (2001) A consensus linkage map for rapeseed (*Brassica napus* L): construction and integration of three individual maps from DH populations. Theor Appl Genet 103:491–507

Murata S, Takasaki N, Saitoh M, Okada N (1993) Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. Proc Natl Acad Sci USA 90:6995–6999

Okada N (1991) SINEs. Curr Opin Genet Dev 1:498–504

Parkin IAP, Sharpe AG, Keith DJ, Lydiate DJ (1995) Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). Genome 38:1122–1131

Purugganan MD, Wessler SR (1995) Transposon signatures: species-specific molecular markers that utilize a class of multiple-copy nuclear DNA. Mol Ecol 4:265–269

Rieger MA, Potter TD, Preston C, Powles SB (2001) Hybridisation between *Brassica napus* L and *Raphanus raphanistrum* L under agronomic field conditions. Theor Appl Genet 103:555–560

Salem AH, Kilroy GE, Watkins WS, Jorde LB, Batzer MA (2003) Recently integrated Alu elements and human genomic diversity. Mol Biol Evol 20:1349–1361

Sambrook J, Russell DW (2001) Molecular cloning, 3rd edn. Cold Spring Harbor Laboratory Press, Plainview

Shedlock AM, Okada N (2000) SINE insertions: powerful tools for molecular systematics. Bioessays 22:148–160

Shimamura M, Yasue H, Ohshima K, Abe H, Kato H, Kishiro T, Goto M, Munechika I, Okada N (1997) Molecular evidence from retroposons that whales form a clade within even-toed ungulates. Nature 388:666–670

Song KT, Osborn C, Williams PH (1988) Brassica taxonomy based on nuclear restriction fragment length polymorphism (RFLPs). 1. Genome evolution of diploid and amphidiploid species. Theor Appl Genet 75:784–794

Tatout C, Lavie L, Deragon JM (1998) Similar target site selection occurs in integration of plant and mammalian retroposons. J Mol Evol 47:463–470

Tatout C, Warwick SI, Lenoir A, Deragon JM (1999) SINE insertion as clade markers for wild Cruciferae species. Mol Biol Evol 16:1614–1621

Udall J, Quijada P, Osborn TC (2005) Detection of chromosomal rearrangements derived from homoeologous recombination in four mapping populations of *Brassica napus* L. Genetics 169:967–979

Warwick SI, Black LD (1991) Molecular systematics of *Brassica* and allied genera (subtribe Brassicinae, Brassiceae)—chloroplast genome and cytodeme congruence. Theor Appl Genet 82:81–92

Warwick SI, Black LD (1993) Molecular relationships in subtribe Brassicinae (Cruciferae, tribe Brassiceae). Can J Bot 71:906–918

Warwick SI, Simard MJ, Légère A, Beckie HJ, Braun L, Zhu B, Mason P, Séguin-Swartz G, Stewart JCN (2003) Hybridization between transgenic *Brassica napus* L and its wild relatives: *B rapa* L, *Raphanus raphanistrum* L, *Sinapis arvensis* L, and *Erucastrum gallicum* (Willd) OE Schulz. Theor Appl Genet 3:528–539

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BB, Powell W (1997) Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (*S*-SAP). Mol Gen Genet 253:687–694