

L. Tommasini · J. Batley · G.M. Arnold  
R.J. Cooke · P. Donini · D. Lee · J.R. Law · C. Lowe  
C. Moule · M. Trick · K.J. Edwards

## The development of multiplex simple sequence repeat (SSR) markers to complement distinctness, uniformity and stability testing of rape (*Brassica napus* L.) varieties

Received: 7 January 2002 / Accepted: 12 August 2002 / Published online: 22 October 2002  
© Springer-Verlag 2002

**Abstract** To assess the potential of multiplex SSR markers for testing distinctness, uniformity and stability of rape (*Brassica napus* L.) varieties, we developed three multiplex SSR sets composed of five markers each. These were used to measure the extent of diversity within and between a set of ten varieties using a fluorescence-based semi-automated detection technology. Also, we evaluated the significance of any correlation between SSRs, pedigree and five of the morphological characters currently used for statutory distinctness, uniformity and stability testing of rape varieties. An assignment test was allowed to identify 99% of the plants examined, with the correct variety based on the analysis of 48 individual plants for each variety. Principal coordinate analysis confirmed that a high degree of separation between varieties could be achieved. Varieties were separated in three groups corresponding to winter, spring and forage types. These results suggested that it should be possible to select a set of markers for obtaining a suitable separation. Diversity within varieties varied considerably, according to the variety and the locus examined. No significant correlation was found between SSR and morphological data. However, genetic distances measured by SSRs were correlated to pedigree. These results suggested that SSRs

could be used for pre-screening or grouping of existing and candidate varieties, allowing the number of varieties that need to be grown for comparison to be reduced. Multiplex SSR sets gave high-throughput reproducible results, thus reducing the costs of SSR assessment. Multiplex SSR sets are a promising way forward for complementing the current variety testing system in *B. napus*.

**Keywords** Simple sequence repeats (SSRs) · Multiplex PCR · Distinctness and uniformity testing · *Brassica napus*

### Introduction

*Brassica napus* L. is an important oilseed and fodder crop grown throughout the world in cool and moist climates. This species is an amphidiploid (AACC genome,  $2n = 38$ ) and is believed to have arisen by interspecific hybridization between *Brassica rapa* L. (syn. *campestris*) (AA genome,  $2n = 20$ ) and *Brassica oleracea* L. (CC genome,  $2n = 18$ ) (U 1935). *B. napus* is primarily inbreeding, and outcrossing rate ranges from 5 to 36% under field conditions (Huhn and Rakow 1979; Rakow and Woods 1987). Due to its economic importance, much effort has been put into devising methods for identifying varieties (Cooke 1999).

In countries that are members of the International Union for Protection of Plant Varieties (UPOV 1978, 1994, 1996), new crop varieties must undergo distinctness, uniformity and stability (DUS) testing that constitutes the basis of intellectual property protection in the form of Plant Breeders' Rights (PBR). Each candidate variety must be shown to be distinct from others of common knowledge, and to be sufficiently uniform and stable in the diagnostic characteristics used to demonstrate distinctness. The current testing system, which is based on the assessment of a range of standardised morphological characters, is expensive and time-consuming. It requires large areas of land to grow the crops and highly skilled personnel, making what are often subjective deci-

Communicated by P. Langridge

L. Tommasini · J. Batley · K.J. Edwards (✉)  
School of Biological Sciences, University of Bristol,  
Woodland Road, Bristol, BS8 1UG, UK  
e-mail: K.J.edwards@bristol.ac.uk  
Tel.: +44(0)-1275-392181 Direct +44(0)-1275-549431  
Fax: +44(0)-1275-394281

G.M. Arnold  
IACR – Long Ashton Research Station,  
Department of Agriculture, University of Bristol,  
Long Ashton, Bristol, BS41 9AF, UK

R.J. Cooke · P. Donini · D. Lee · J.R. Law · C. Lowe  
NIAB, Huntingdon Road, Cambridge, CB3 0LE, UK

C. Moule · M. Trick  
John Innes Centre, Brassica and Oilseed Research Department,  
Norwich Research Park, Norwich, NR4 7UH, UK

sions. An added complication with *B. napus* is the existence of various types of hybrid variety, some of which are also genetically modified. Furthermore, with the increasing number of members of UPOV, variety reference collections are expanding and it is becoming impossible to compare adequately new varieties with those of common knowledge using the limited number of morphological characters available. This may affect the quality of protection offered by PBR systems. It is therefore important to devise more rapid and cost-effective testing procedures to improve the current testing system (Cooke 1999).

Different biochemical technologies have proven potentially useful for fingerprinting *B. napus* varieties. These include the comparison of seed oil fatty acid profile by GLC analysis (White and Law 1991), HPLC analysis of leaf glucosinolates (Adams et al. 1989) and starch-gel electrophoresis of cotyledon isozymes (Mundges et al. 1989). However, biochemical and morphological markers can detect only a limited degree of polymorphism, and may be sensitive to environmental and developmental variation. By contrast, molecular markers are numerous, polymorphic and unaffected by the environment. Therefore, they offer potential advantages for plant variety profiling. RAPD (randomly amplified polymorphic DNA) and RFLP (restriction fragment length polymorphism) markers have been extensively used for fingerprinting *Brassica* varieties (Kresovich et al. 1992; Waugh and Powell 1992; Hallden et al. 1994; Mailer et al. 1994; Lee et al. 1996a, b). Nevertheless, the application of RFLPs or RAPDs to variety testing is limited by the cumbersome technique and low numbers of discriminating loci of RFLPs or by the dominant character and lack of reproducibility of RAPDs (Jones et al. 1997; Karp et al. 1997). Recently, simple sequence repeat (SSR) markers, consisting of short tandem base repeats (2–8 bp units), have gained increasing importance in plant variety testing (Cooke 1999; Donini et al. 2000). SSRs were studied in several crop species including *Brassicaceae* (Kresovich et al. 1995; Charters et al. 1996; SzewcMcFadden et al. 1996; Uzunova and Ecker 1999; Pleske and Strauss 2001; Saal et al. 2001). These studies showed that SSRs are a valuable tool for characterising germplasm in *Brassica* species, because they are numerous, highly informative, technically simple, robust and suitable for automated allele detection and sizing (Rafalski and Tingey 1993).

Until recently, the use of SSRs was limited by the slow data-throughput and the limited availability of markers coupled with the high cost of marker development. However, improvements in molecular-marker technology, such as multiplex PCR for DNA amplification, consisting of the simultaneous amplification of more than one SSR in a single reaction, combined with the utilisation of fluorescence-based automated DNA detection and fragment sizing, allow faster, more accurate and cost-effective acquisition of data. This offers a potential improvement to the efficiency and affordability of variety testing (Mitchell et al. 1997). Furthermore, due to the economic importance of cultivated *Brassica* species,

large investments have been made in the development of *Brassica* SSRs, many of which are available to the scientific community (<http://ukcrop.net/perl/ace/search/BrassicaDB>).

In outbreeding species, variation within varieties tends to be high and this lack of uniformity could hamper the use of molecular markers for distinguishing between varieties. Preliminary studies have suggested that it might be difficult to identify markers that are sufficiently uniform within varieties and, at the same time, are sufficiently variable between varieties to allow for variety discrimination (Mailer et al. 1994; Plaschke et al. 1995; Charters et al. 1996; Olufowote et al. 1997; Donini et al. 1998; Roldan-Ruiz et al. 2000). However, genetic diversity within *B. napus* varieties has not been studied extensively. Therefore, it is essential that intra-variety diversity is investigated before using DNA profiles from pooled or single samples for distinctness testing.

The aim of this research was to evaluate whether the use of multiplex PCR of SSRs combined with fluorescence-based automated technology could improve the distinctness and uniformity testing of rape varieties. To perform this, we developed three multiplex SSR sets and used these to evaluate the extent of genetic diversity within and between a set of ten varieties, based on the analysis of DNA from 48 individuals from each of these and on bulked DNA. We also assessed whether SSR data were significantly correlated to pedigree and morphological characters currently used for testing distinctness and uniformity of *B. napus* varieties.

## Materials and methods

### Plant material and DNA isolation

Seed of winter oilseed rape and forage rape varieties were obtained from the reference collections held at the NIAB, UK, while spring oilseed rape varieties were obtained from Bundessortment, Germany. This plant material represents ten varieties from the UK National List (Table 1) and was selected from different breeding

**Table 1** List of *B. napus* varieties surveyed with SSR markers. Variety classifications were obtained from listings compiled by NIAB, Cambridge, United Kingdom, and Bundessortment, Germany

Variety	Breeder	Type <sup>a</sup>	Quality <sup>b</sup>
Nimbus	Prodana Seeds A/S	S	DL
Falcon	NPZ Lembke	W	DL
Express	NPZ Lembke	W	LE
Marinka	Maribo UK Ltd	S	DL
Apex	GMBH	W	DL
Hobson	Sinclair McGill	F	LE
Bienvenue	Ets. Ringot	W	LE
Global	Svalof AB	S	DL
Cobol	Deutsche Staatveredelung	W	DL
Lirawell	Deutsche Staatveredelung	S	DL

<sup>a</sup> W, winter cultivar (i.e. biennial); S, spring cultivar (i.e. annual); F, forage rape

<sup>b</sup> DL, low erucic acid and glucosinolate content; LE, low erucic acid content

programmes in order to maximise the likelihood of polymorphism. These varieties have undergone statutory testing and proved to be distinct, uniform and stable.

Seedlings were grown in a greenhouse until the fourth true-leaf stage. DNA was extracted from a bulk of leaf tissue composed of 30 plants for each variety. Plants that made up the bulks were different from the plants analysed individually. For the bulk samples, 10 g of leaf tissue from each plant were ground in liquid nitrogen and phenol/chloroform extraction was performed (Ellis et al. 1988). Forty eight individual plants from each variety were analysed to determine the degree of diversity within and between each variety. Fifty milligrams of leaf discs from each plant were freeze-dried and DNA extracted by using a 'salting out procedure' (Genovar Diagnostics, Kent). The final DNA concentration ranged from 10 to 100 ng/ $\mu$ l.

#### Selection of markers for developing multiplex sets

Four hundred and six SSR primer sets developed from *B. napus*, *B. oleracea*, *B. rapa* and *Brassica nigra* were available for this study. Repeat units, primer sequences and in part the map position for the

**Table 2** PCR reagent concentration for the three multiplex SSR sets

Multiplex set no.	1	2	3
Template DNA (ng)	0.82	0.82	0.82
1 $\times$ PCR buffer ( $\mu$ l)	1	1	1
dNTP (nM)	0.33	0.33	0.33
	Primer set (ng)	Primer set (ng)	Primer set (ng)
	1 20	6 82	11 123
	2 82	7 82	12 123
	3 82	8 20	13 31
	4 20	9 20	14 31
	5 123	10 123	15 31
Amplitaq Gold (U)	0.4	0.4	0.4

loci under study have been previously published (<http://ukcrop.net/perl/ace/search/BrassicaDB>). Forty nine of these were chosen and used to screen bulked DNA samples from the ten varieties using radioactive labelling and standard denaturing gel-electrophoresis.

In order to label PCR amplification products, 75 ng of forward primer was radiolabelled with 0.88  $\mu$ Ci [ $\gamma$ - $^{32}$ P] ATP using 1.1 U of T<sub>4</sub> Polynucleotide-Kinase (Amersham/Pharmacia). The reaction was performed in 1  $\times$  One-Phor-All Buffer Plus (Amersham/Pharmacia) at 37 °C for 30 min and stopped at 65 °C for 10 min.

The PCR mixture (12  $\mu$ l total volume) contained 10–100 ng of template DNA, 75 ng of reverse primer, 100 ng of forward primer, 0.2  $\mu$ M of each dNTP (Promega), 0.25 U of AmpliTagGold® DNA polymerase (Perkin Elmer) and 1  $\times$  PCR buffer (Perkin Elmer). The PCR profile consisted of a hot start (94 °C for 10 min) followed by 35 cycles involving an initial denaturation step (94 °C for 60 s), a 60-s annealing step (54 °C), a 30-s elongation step (72 °C) and a final step of 72 °C for 10 min. Amplifications were performed using a GeneAmpPCR System 9700 (PE Biosystems) thermal cycler.

Using an M13 length standard, allele lengths were determined by electrophoresis of 3–4  $\mu$ l of PCR products on a 5.3% denaturing polyacrylamide gel. After the runs, the gels were transferred onto Chromatography Paper 3MM Chr (Whatman), covered with cling film (Dow Saran) and dried using a gel drier (BIO-RAD, Model 583) at 80 °C for 2 h. The dried gels were exposed to Kodak Biomax MR film (high sensitivity) for 12–48 h and scored manually.

#### Development of multiplex SSR sets

Fifteen of the 49 markers screened were chosen for developing three multiplex sets with five markers in each set (Table 2), based on: (1) ease of score; (2) the level of polymorphism detected in the ten varieties; (3) a compatible allele size range; and (4) similar optimal reaction conditions. These markers were composed of either di- or tri-nucleotide repeats. Thirteen of the multiplexed markers were mapped after they had been selected (<http://ukcrop.net/perl/ace/search/BrassicaDB>). These amplified from 18 loci and covered 13 (34%) of the chromosome arms (Table 3).

**Table 3** Allele size range and the number of bands for each of the 15 multiplexed SSRs in ten varieties of *B. napus* based on the analysis of DNA from bulked samples and individual plants; genotype diversity between varieties based on bulked DNA analysis

Primer set	Multi-plex set	Core motif	Dye label	Linkage group	Expected size (bp)	DNA from bulks of 30 plants			DNA from 48 individual plants from each variety	
						Allele size range (bp)	Number of bands	Genotype diversity between varieties	Allele size range (bp)	Number of bands
1-Na12 E02	1	TTG	FAM	–	127	124–129	2	0.66	98–171	8
2-Na12 A08	1	GA	FAM	3, 6	156	162–318	8	0.74	290–318	9
3-Ra2 EO3	1	CT	TET	19	279	262–294	5	0.34	252–294	7
4-Na12 A02	1	CT	TET	7, 10	190	161–197	7	0.68	164–202	10
5-Na12 AO7	1	GT	HEX	12	158	153–166	4	0.68	160–166	3
6-Ra2 A11	2	CT	TET	9	230	243–245	2	0.32	243–245	2
7-O110 B01	2	GA	TET	17	190	168–188	4	0.46	168–188	6
8-Na12 F03	2	GA	FAM	17	301	254–314	6	0.78	254–318	9
9-O110 B01	2	CT	FAM	–	131	91–142	6	0.82	88–142	12
10-Ra2 F11	2	CT	HEX	12, 13, 19	240	188–233	5	0.50	188–232	7
11-Na12 D04	3	CA	FAM	6	281	282–288	2	0.18	282–288	2
12-Ni4 D09	3	CT	HEX	9	206	167–207	6	0.80	167–207	7
13-Ra2 E11	3	CT	FAM	13	198	166–202	7	0.82	166–202	7
14-O110 F11	3	GGC	TET	11	149	138–150	3	0.64	131–144	3
15-O110 H02	3	GGC	TET	12, 13	220	184–214	3	0.54	205–214	2
Total							70			93
Mean							4.6	0.56		6.2

Fluorescently labelled primers were used for developing multiplex PCR sets. Three different dyes were used in each multiplex set. SSR bands were detected by attachment of a fluorescent dye molecule, and fluorescent oligonucleotides were synthesized (Sigma-Genosys). At the final step in each DNA synthesis, a phosphoramidite labelled with fluorescent dye, either 6-carboxyfluorescein (6-FAM), tetrachloro-6-carboxyfluorescein (TET) or hexachloro-6-carboxyfluorescein (HEX), was incorporated at the 5' end of the primer (Perkin Elmer/Applied Biosystems). Primers were re-suspended in double-distilled water at a concentration of 1 µg/µl. Primer working stocks were prepared at a concentration of 0.1 µg/µl.

First, single primer PCR amplifications were performed to check the primer set ease of score using fluorescent dyes, and to compare result reproducibility in single and multiplex reactions. Single primer PCR reactions were performed in 12 µl volumes containing 10–100 ng of genomic DNA, 0.075 µg of each primer, 1.25 µl of 1 × PCR buffer (Perkin Elmer), 0.2 mM of dNTPs and 0.27 U of *AmpliTag*Gold DNA® polymerase (Perkin Elmer). The same PCR profile and thermal cyclers as in the radiolabelled PCR were used.

To optimise multiplex reactions, first primers were added in equimolar amounts in the multiplex PCR reaction. Concentrations and thermal-cycler conditions were then optimised according to the level of amplification observed for each marker at a particular concentration, for obtaining a similar level of amplification in each multiplex set (Henegariu et al. 1997).

Amplified loci were detected by laser scanning during electrophoresis, using an ABI 377 Sequencer (Perkin Elmer/Applied Biosystems). Samples containing 1 µl of PCR product, 0.25 µl of GeneScan 500 internal lane standard labelled with N, N', N'' – tetramethyl-6-carboxyrhodamine (TAMRA) (Perkin Elmer/Applied Biosystems) and 1.25 µl of formamide were heated at 94 °C for 2 min, placed on ice and then loaded on 5% denaturing polyacrylamide [18 g of urea, 0.5 g of Dowex resin (Sigma), 5.6 ml of 29:1 Accugel (National Diagnostics), 5 ml of 10 × TBE and 25 ml of double-distilled water] gels (12 cm WTR plates).

Band sizes were generated automatically in comparison with a standard sizing ladder included in every sample prior to electrophoresis, using Genescan® and Genotyper® computer software. Band scoring was then checked manually. Banding-profile reproducibility was assessed by repeating experiments in independent single and multiplex PCRs and electrophoreses using bulked DNA. The ability of the multiplex set to identify varieties was tested by comparison of the banding profiles using all three multiplex sets.

#### Data analysis

SSR bands were recorded as present (1) or absent (0). Results from SSR analysis of bulked DNA represented a consensus of results from five independent single and multiplex PCR amplifications and electrophoreses. Inconsistently amplified bands were not included in any analyses. A similarity matrix was calculated using the Jaccard coefficient (Sneath and Sokal 1973) pooling together alleles across all loci.

To compare diversity between and within varieties, principal coordinates analysis (PCO; Gower 1966) was performed on the similarity matrix. The clustering method used was the unweighted pair group with arithmetic average clustering (UPGMA; Sneath and Sokal 1973).

Separation between varieties, based on analysis of DNA from individual plants and from bulks, was evaluated visually on the principal coordinate analysis (PCO) plot of the first two coordinate axes.

To quantitatively evaluate the ability of the multiplex sets to assign each individual plant to the correct variety, an assignment test (Peatkeau et al. 1995) was carried out based on genotype frequencies. This involved calculating the expected frequency of each individual's genotype in each of the ten varieties and subsequent assignment of each individual to the population where its

expected genotype frequency was highest. This calculation assumes random mating and linkage equilibrium within each variety.

Markers amplifying from one locus could be used for the estimation of allele frequencies, but markers amplifying from more than one locus allowed estimation of genotype frequencies only. To assess the robustness and the potential application of different measures of diversity, both gene and genotype diversity within varieties were calculated and compared. Gene diversity ( $D_i$ ) at a locus is calculated from allele frequencies and is defined as

$$D_i = 1 - \sum_u p_{iu}^2 \quad (\text{Weir 1990}),$$

where  $p_{iu}$  is the frequency of the  $u$ th allele for the  $i$ th locus, squared and summed across all alleles in the locus. Gene diversity is therefore restricted to those markers amplifying from one locus. In this paper we define genotype diversity ( $D_j$ ) in the same sense as Anderson et al. (1993) where a marker amplifies more than one locus. Genotype diversity is defined as

$$D_j = 1 - \sum_j p_{ij}^2 \quad (\text{adapted from Anderson et al. 1993}),$$

where  $p_{ij}$  is the frequency of the  $j$ th genotype for the  $i$ th marker, squared and summed across all genotypes for the marker. These methods for measuring diversity are similar to those used for measuring quantitative morphological characters, where mean and variance are calculated. An alternative method to measure genetic diversity is based on the estimation of the percentage of individuals possessing the most-frequent genotype within a variety, defined as conformity, which is a measure used for assessing qualitative morphological characters. Non-conformity ( $D_2$ ), defined as  $(100 - \text{conformity})/100$ , was calculated for assessing any significant correlation between measures of conformity, gene diversity and genotype diversity. Correlation was evaluated using (Spearman's) rank correlation. Genotype diversity and conformity within varieties were calculated for each variety and each locus, while gene diversity within varieties could be calculated only for primer sets amplifying from a single locus.

The typical uniformity of a variety was calculated as the mean uniformity across 15 markers based on genotype diversity values. Similarly, the typical uniformity of a marker was evaluated as the mean uniformity across the ten varieties.

Total separation efficiency (TSE), defined as the ratio of between within-variety genotype diversity, was calculated for each marker. Measures of genotype diversity between varieties were based on results from bulked DNA analysis, while mean genotype diversity within varieties was evaluated based on the analysis of 48 individual plants from each variety.

Mantel tests (Mantel 1967) were carried out for assessing the significance of any correlation between combined SSR data based on bulked DNA analysis, pedigree and morphological data. The similarities based on pedigree were calculated as the proportion of the shared genome according to pedigree information available from Technical Questionnaires held at NIAB. No pedigree information was available for the variety 'Falcon'. Five morphological characters were examined: erucic acid content, colours of petals, winter requirement for flowering, length of stem and time of flowering. These characters are a subset of those currently used in DUS testing (International Union for the Protection of New Varieties of Plants 1996) and were available from Technical Questionnaires held at NIAB. The number of morphological characters included in the analysis was limited by missing data. The variety 'Marinka' was not considered for the Mantel test because of missing information on three morphological characters. Similarity indices were computed using the Gower general coefficient of similarity (Gower 1966). One thousand randomisations were used for the Mantel tests to test for significance.

## Results

### Development of multiplex SSR sets

Forty six of the 49 primer sets (94%) screened amplified target sequences in the ten varieties; 27 of these (58%) showed clear banding patterns and could be scored unambiguously. Twenty of the scored markers (74%) were polymorphic (41% of those screened). Fifteen of these were selected, of which six were developed from *B. napus*, four from *B. rapa*, four from *B. oleracea* and one from *B. nigra*, for developing three multiplex sets, each composed of five markers. Within these, alleles close in size could be distinguished using different fluorescent dye labels.

Equimolar primer concentrations in multiplex PCRs showed uneven amplification, with some of the products scarcely amplified. Similar levels of amplification of each marker could be obtained by decreasing the quantity of primer for the strongly amplified fragments, increasing the amount of primers for the poorly amplified fragments and adjusting the concentration of the remaining PCR reagents accordingly (Table 2). The optimised thermal cycler conditions for multiplex set no.1 and 2 consisted of a hot start (94 °C for 10 min) followed by 35 cycles involving an initial denaturation step (94 °C for 1 min), a 1-min annealing step (54 °C), a 1-min elongation step (72 °C) and a final step of 72 °C for 10 min. Thermal cycler conditions for multiplex set no. 3 were identical to those used for single PCR amplification.

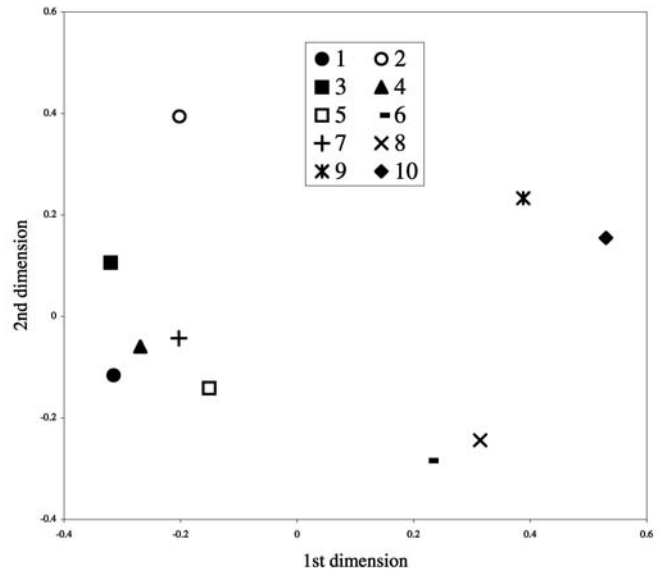
Banding profiles obtained from single and multiplex PCRs of bulked DNA were very similar (data not shown). Although slight band differences were found in repeated experiments, all ten varieties were identified correctly by comparing DNA profiles generated by all three multiplex sets to those generated by the respective primers amplified singly, as shown by results from PCO analyses based on single and multiplex amplifications of bulked DNA (Figs. 1 and 2 respectively).

### Analysis of diversity between varieties

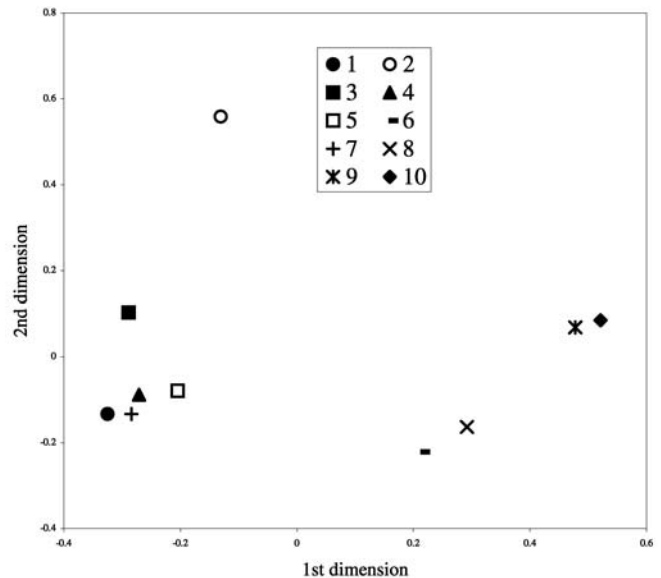
#### Analysis of bulked DNA samples

Table 3 shows that overall, 70 bands were detected using three multiplex sets, 58 of which were polymorphic, with a mean of 4.6 bands per marker. Genotype diversity between varieties for single markers varied considerably, ranging from 0.18 to 0.82, with a mean of 0.56.

PCO analyses showed that combined data from 15 markers could separate spring, winter and forage varieties into three distinct groups (Fig. 2). This is similar to results from the PCO analysis based on DNA from individual plants (Fig. 3). Unique DNA profiles in all ten varieties were observed using a minimum of two highly polymorphic markers (nine and 13). A progressive separation between varieties was observed by increasing the number of markers (data not shown).



**Fig. 1** PCO plot for the second vs first axes estimated for combined data from 15 singly amplified SSR markers on ten varieties of *B. napus* using bulked DNA from 30 plants. 1–10: individuals from variety 1 to 10. names were substituted with numbers 1 to 10 according to Table 6



**Fig. 2** PCO plot for the second vs first axes estimated for combined data from 15 multiplexed SSR markers on ten varieties of *B. napus* using bulked DNA from 30 plants. 1–10: individuals from variety 1 to 10. Variety names were substituted with numbers 1 to 10 according to Table 6

Results from the Mantel test comparing SSR with the pedigree similarity matrix showed strong evidence of a link between them ( $\rho = 0.5159$ ,  $p = 0.004$ ). The Mantel test did not show any correlation between SSR and morphological characters ( $\rho = 0.098$ ,  $p = 0.283$ ).



**Table 6** Genotype diversity ( $D_i$ ) within varieties for 15 SSR markers based on the analysis of 48 individual plants from each variety of the test array

Variety <sup>a</sup> Type <sup>b</sup>	1 W	2 F	3 W	4 W	5 W	6 S	7 W	8 S	9 S	10 S	Mean
Primer set											
5	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.01
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.04	0.01
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.01
3	0.00	0.44	0.12	0.00	0.29	0.00	0.04	0.00	0.04	0.00	0.09
6	0.00	0.00	0.52	0.04	0.00	0.00	0.04	0.41	0.00	0.00	0.10
15	0.00	0.00	0.00	0.00	0.00	0.45	0.12	0.25	0.26	0.00	0.11
10	0.00	0.04	0.00	0.12	0.00	0.00	0.00	0.34	0.08	0.51	0.11
1	0.00	0.00	0.08	0.04	0.40	0.38	0.08	0.04	0.04	0.04	0.11
7	0.00	0.00	0.04	0.31	0.00	0.40	0.04	0.04	0.12	0.22	0.12
12	0.00	0.00	0.08	0.00	0.00	0.19	0.29	0.26	0.12	0.25	0.12
8	0.00	0.04	0.04	0.16	0.00	0.41	0.35	0.08	0.56	0.16	0.18
2	0.00	0.08	0.00	0.00	0.04	0.04	0.52	0.57	0.19	0.64	0.21
13	0.00	0.00	0.08	0.08	0.04	0.22	0.63	0.33	0.81	0.68	0.29
9	0.00	0.04	0.19	0.32	0.68	0.12	0.43	0.08	0.89	0.89	0.36
4	0.49	0.00	0.08	0.46	0.39	0.00	0.30	0.65	0.69	0.65	0.37
Mean	0.03	0.04	0.08	0.10	0.12	0.15	0.19	0.20	0.27	0.27	0.15

<sup>a</sup> Variety names were substituted with numbers 1 to 10 to avoid assigning within-variety diversity values to particular varieties

<sup>b</sup> W, winter cultivar (i.e. biennial); S, spring cultivar (i.e. annual); F, forage rape

**Table 7** Gene diversity ( $D_i$ ) within varieties for 15 SSR markers based on the analysis of 48 individual plants from each variety of the test array

Variety <sup>a</sup> Type <sup>b</sup>	1 W	2 F	3 W	4 W	5 W	6 S	7 W	8 S	9 S	10 S	Mean
Primer set											
3	0.00	0.44	0.10	0.00	0.29	0.00	0.02	0.00	0.04	0.00	0.09
5	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.01
6	0.00	0.00	0.39	0.04	0.00	0.00	0.02	0.37	0.00	0.00	0.08
8	0.00	0.02	0.02	0.15	0.00	0.33	0.31	0.08	0.51	0.13	0.17
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.01
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.02	0.01
15	0.00	0.00	0.00	0.00	0.00	0.30	0.06	0.15	0.19	0.00	0.07
Mean	0.00	0.07	0.08	0.03	0.04	0.09	0.06	0.09	0.13	0.02	

<sup>a</sup> Variety names were substituted with the numbers 1 to 10 to avoid assigning within-variety diversity values to particular varieties

<sup>b</sup> W, winter cultivar (i.e. biennial); S, spring cultivar (i.e. annual); F, forage rape

**Table 8** Spearman rank correlation and  $P$ -value (in brackets) between gene diversity ( $D_i$ ), genotype diversity ( $D_i$ ) and non-conformity ( $D$ ) within varieties

Item	Gene diversity ( $D$ )	Genotype diversity ( $D_i$ )
Genotype diversity ( $D_i$ )	0.957 (<0.001)	–
Non-conformity ( $D_2$ )	0.964 (<0.001)	0.995 (<0.001)

could be calculated only at nine loci (3,5,6,7,8,11,12,13 and 14, Table 7). These three measures of diversity within varieties were highly correlated (Table 8) and consistent with results from the PCO analyses (Fig. 3).

## Discussion

### Evaluation of distinctness between varieties

The three multiplex sets were successful at discriminating between the ten varieties based on the analysis of bulks. Values of diversity between varieties based on bulks were in agreement with previous works on *Brassica* SSRs (Szwec-McFadden et al. 1996; Raybould et al.

1999; Plieske and Stauss 2001). The analysis of SSR profiles from bulked DNA is more complex than the analysis of single plants, either using data from single or multiplex amplifications. Moreover, the analysis of profiles from multiplex amplifications is more complex than the analysis of profiles from single amplifications. For this, it is more accurate to use data from single amplifications for genotyping bulked DNA. Nevertheless, PCO analyses of bulked DNA showed a high degree of similarity between results from single and multiplex amplifications. Additionally, both methods allowed to discriminate between all the varieties examined. The analysis of individuals had the advantage of allowing heterogeneity within varieties to be estimated and more reliable scoring than the analysis of bulks, as found in previous work (Donini et al. 1997; Olufowote et al. 1997; van Treuren 2001).

Ninety nine percent of the single plants examined could be identified with the correct variety based on the assignment test. Furthermore, a high degree of separation between the ten varieties examined could be achieved based on PCO analysis of individual plants. These were more clearly separated than in other studies based on AFLPs in an outbreeding crop species (Roldan-Ruiz et al. 2000). The high level of discrimination be-

tween varieties is likely to be due to the high uniformity within some varieties, reflecting inbreeding (Dillman et al. 1997), whilst the assessment of unselected varieties would probably result in poorer differentiation (Roldan-Ruiz et al. 2000). Some overlapping, evident when DNA from individual plants was analysed, did not allow assigning each plant to the correct variety. This was not only due to diversity within varieties, but also to outliers that may have resulted from seed mixing and cross-pollination during seed-production trials.

In general, new varieties may be considered “distinct” even in cases where they overlap with other varieties to a limited extent. However, it is necessary to quantify a threshold of the “minimum distance” between the new variety and any other variety from the reference collection, for the new variety to be considered distinct. For quantitative morphological characteristics, the distinction is considered clear if it occurs with 1% probability of an error; for example, on the method of the Least Significant Difference (International Convention for the Protection of new Varieties of Plants 1978). Similarly, the minimum distance could be quantified based on measured mean genetic distances between a set of approved varieties using individual plant DNA analysis. Nevertheless, statistical methods have to be developed for measuring genetic distances between varieties based on DNA from a group of single plants. Interesting approaches include the assignment test (Paetkau et al. 1995), which could be adopted to quantify a minimum distance, and therefore allow the definition of criteria for distinctness based on molecular data. However, distinctiveness as defined by legislation bears no relationship to genetic distance. As a consequence, isogenic lines with a single gene difference may be validly treated as ‘distinct’ according to the legislation, if the morphological distances make them distinct.

Considering that about 300 varieties constitute the EC Common Catalogue, and that there is an increasing number of new varieties that need to be tested, the morphological markers currently used in DUS testing or the 15 markers examined in this study would not be sufficient for discriminating varieties. The separation between varieties depends on the quality and number of markers. In fact, results showed that the total separation efficiency for a marker varied considerably and that progressive separation between varieties was obtained by increasing the number of markers used, similar to results found in previous work (Dillmann et al. 1997). This indicated that even markers with low total separation efficiency gave a contribution towards variety discrimination. Furthermore, these markers may be essential for enabling discrimination of a few closely related varieties (Tessier et al. 1999). When it is considered that 41% of the screened marker set amplified polymorphic bands from the test array, approximately 166 markers useful for distinctness testing should be available from the Biotechnology and Biological Sciences Research Council of the United Kingdom (BBRSC) Brassica SSR program. Therefore, it should be possible to select a set of SSRs that are poly-

morphic between varieties and sufficiently uniform within varieties for obtaining a suitable separation. This set could be calibrated in order to obtain the same separation as that obtained in the current testing system (Law et al. 1998).

#### Analysis of diversity within varieties

The heterogeneity found within varieties is in agreement with other preliminary studies on *Brassica* and other crop varieties using molecular markers (Poulsen et al. 1993; Olufowote et al. 1997; Donini et al. 1998; Roldan-Ruiz et al. 2000). Genetic heterogeneity in the ten varieties accompanies morphological heterogeneity in the descriptors used in the current testing system. It is not clear what proportion of morphological variation within *B. napus* varieties depends on genetic or on environmental variation, but our results suggested that the former might be significant. An additional problem in rape is the existence of various types of variety (single lines, synthetic populations, different types of hybrids) with varying genetic structures. Genetic heterogeneity might also be due to seed mixture and/or cross-pollination.

Since uniformity values in a particular variety differed according to the primer set used and the variety examined, a sufficient number of markers should be considered for estimating the typical (mean) uniformity of a variety. Similarly, to evaluate the typical uniformity of a marker, a representative set of current varieties should be considered. Currently, a cross-pollinated variety is considered uniform in a measured morphological characteristic, if its variance does not exceed 1.6-times the average of the variances of the varieties used for comparison (International Convention for the Protection of new Varieties of Plants 1994). Similarly, a threshold for genetic uniformity could be established based on the mean observed uniformity at combined SSR markers in a representative set of approved varieties.

Uniformity differences in different varieties may be due to differences in the breeding system used by different companies. Also, results suggested that diversity within spring varieties is higher than in winter varieties, although a larger set of varieties needs to be analysed before the causes of the observed differences can be estimated. Differences in uniformity for different markers might be due to the closeness of some markers to regions of the genome subject to breeding selection and, therefore, selected indirectly. This hypothesis is supported by previous studies showing that allele frequencies at some SSR loci are under functional constraints (Boland 1996). The lack of correlation between SSR and morphological data might be due to the limited number of molecular and morphological markers examined. More work is needed before conclusions on the causes of differences in uniformity at different SSR loci may be drawn. Understanding which factors affect the level of diversity within varieties at SSR loci could enable thresholds to be set for the genetic uniformity of new varieties. Markers



of genes regulating the morphological characters currently used in variety testing or other traits of interests for breeding, for instance Expressed Sequence Tags (EST) and Single Nucleotide Polymorphism (SNP) markers, are likely to show a higher degree of uniformity within varieties and consequently enable a higher degree of separation between varieties than neutral markers. Therefore, for plant variety testing, ESTs and SNPs offer potential advantages compared to SSR markers.

The correlation between the measures of diversity within varieties based on gene diversity, genotype diversity and conformity suggested that all the methods are robust and any of them can be used for uniformity testing. Therefore, markers which amplify from two or more loci from the *B. napus* genome can be used as well as markers amplifying from one locus only. The first are more informative, but only the latter can be used for estimating gene frequencies.

#### Correlation between SSR, pedigree and morphological data

The correlation between SSR and pedigree data suggested that the marker sets were able to determine relationships between varieties, in agreement with previous work (Law et al. 1999). Varieties could be separated into three distinct groups corresponding to spring, winter and forage varieties, as in former studies based on RFLP markers (Lee et al. 1996b). These results indicated that SSRs could be used for pre-screening or grouping of existing and candidate varieties, so that the number of varieties that need to be grown for comparison could be reduced. With the increasing size of reference collections and the number of candidate varieties, this would be a potentially attractive option. The lack of correlation between SSR and morphological data might be due the limited number of SSR markers and morphological characters available for comparison.

#### Efficiency of the technology

The definition of a "clearly distinct" variety based on morphological characters can cause discrepancies in interpretation due to differences in the character assessment, and variation caused by environmental interaction. Therefore, there is much interest in defining this concept more precisely (Schneider 1986). The marker sets used in this study showed clear banding patterns and allowed us to obtain highly reproducible results. The semi-automated fluorescence-based system permitted accurate scoring, reducing the reliance on subjective manual methods. Therefore the system developed here would in principle allow distinctness to be defined more objectively and precisely. However, for widespread use of DUS testing, it is important that it can be established that the markers can give comparable results irrespective of

the detection platform. Thus, some between-laboratory testing of markers is important.

## Conclusion

This research showed that SSR markers have a potential application for discriminating between *B. napus* varieties. Mean observed measures of uniformity for current varieties could be used to establish a baseline of uniformity for candidate varieties. Furthermore, the multiplex sets were able to determine the genetic relationships between varieties and therefore could be used for pre-screening and grouping candidate and existing varieties, reducing the number of varieties that need to be grown in field trial for comparison. Therefore, multiplex SSR sets coupled with fluorescence-based automated detection systems are a promising way forward for complementing statutory DUS testing of *B. napus* varieties.

A more comprehensive study on a larger set of varieties and on a larger number of markers should be carried out before finally developing a set of markers suitable for variety testing. This would enable better understanding and definition of the concept of 'minimum distance' between varieties, and also establish a threshold of uniformity within varieties. Statistical methods for quantifying genetic distances between varieties based on the analysis of DNA from a group of individual plants have to be optimised. Furthermore, research on marker stability across generations and reproducibility of results in different laboratories using different technologies is essential for the widespread application of SSR markers in DUS testing. Such studies will enable plant variety testing authorities to evaluate how tools offering high discrimination power should be used in DUS testing in the future.

Although there may be problems with the use of SSR or other molecular markers in DUS testing, the potential that they have for providing a rapid, cost-effective system deserves to be examined seriously. The large numbers of candidate varieties, and the existence of different hybrid types, along with the increasing size of the variety reference collections and of varieties of common knowledge, are all making the efficient-testing of new oilseed rape varieties very difficult. This in turn may ultimately undermine the quality of protection offered by PBR schemes. Even if the SSR markers described turn out to be not useful for DUS testing and variety registration purposes, they will have value in other areas of the production chain where it is necessary to be able to separate and/or identify varieties (e.g. for processing), and for consumer protection purposes. At the moment, such applications are practically impossible in oilseed rape trading, given the lack of morphological characteristics that can be applied to seeds.

**Acknowledgements** We thank the Department for the Environment, Food and Rural Affairs for financial support. IACR receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom (grant reference number 573 2301).

## References

- Adams H, Vaughan JG, Fenwick GR (1989) The use of glucosinolates for cultivar identification in swede, *Brassica napus* (L.) var *napobrassica* (L.) Peterm. *J Sci Food Agr* 46:319–324
- Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. *Genome* 36:181–186
- Boland CR (1996) Setting microsatellites free. *Nature Med* 2:972–974
- Charters YM, Robertson A, Wilkinson MJ, Ramsay G (1996) PCR analysis of oilseed rape cultivars (*Brassica napus* L. ssp. *oleifera*) using 5'-anchored simple sequence repeat (SSR) primers. *Theor Appl Genet* 92:442–447
- Cooke RJ (1999) Modern methods for the cultivar identification and the transgenic plant challenge. *Seed Sci Technol* 27:669–680
- Dillmann C, BarHen A, Guerin D, Charcosset A, Murigneux A (1997) Comparison of RFLP and morphological distances between maize *Zea mays* L. inbred lines. Consequences for germplasm protection purposes. *Theor Appl Genet* 95:92–102
- Donini P, Elias L, Bougourd S, Koebner RMD (1997) AFLP fingerprinting reveals pattern differences between template DNA extracted from different plant organs. *Genome* 40:521–526
- Donini P, Stephenson P, Bryan GJ, Koebner RMD (1998) The potential of microsatellites for high-throughput genetic-diversity assessment in wheat and barley. *Genet Resour Crop Evol* 45:415–421
- Donini P, Cooke RJ, Reeves JC (2000) Molecular markers in variety and seed testing. In: Arenciba AD (ed) *Plant genetic engineering: towards the third millennium*. Elsevier Science B.V. pp 27–34
- Ellis THN, Lee D, Thomas CM, Simpson PR, Cleary WG, Newman MA, Burcham KWG (1988) 5S ribosomal-RNA genes in *Pisum* – sequence, long-range and chromosomal organization. *Mol Gen Genet* 214:333–342
- Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 55:582–585
- Halden C, Nilsson NO, Rading IM, Sall T (1994) Evaluation of RFLP and RAPD markers in a comparison of *Brassica napus* breeding lines. *Theor Appl Genet* 88:123–128
- Henegariu O, Heerema NA, Dlouhy SR, Vance GH, Vogt PH (1997) Multiplex PCR: critical parameters and step-by-step protocol. *Biotechniques* 23:504–511
- Huhn M, Rakow G (1979) Einige experimentelle Ergebnisse zur Fremdbefruchtungsrates bei Winterrap (*Brassica napus oleifera*) in Abhängigkeit von Sorte and Abstand. *Z Pflanzenzucht* 83:289–307
- Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, vandeWiel C, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevski A, Marmiroli N, Aert R, Volckaert G, Rueda J, Linacero R, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol Breed* 3:381–390
- Karp A, Edwards KJ, Bruford M, Funk S, Vosman B, Morgante M, Seberg O et al. (1997) Molecular technologies for biodiversity evaluation: opportunities and challenges. *Nature Biotech* 15:625–628
- Kresovich S, Williams JGK, McFerson JR, Routman EJ, Schaal BA (1992) Characterization of genetic identities and relationships of *Brassica oleracea* L. via a random amplified polymorphic DNA assay. *Theor Appl Genet* 85:190–196
- Kresovich S, SzewcMcFadden AK, Bliet SM, McFerson JR (1995) Abundance and characterization of simple sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* (L.) (Rapeseed). *Theor Appl Genet* 91:206–211
- Law JR, Donini P, Koebner RMD, James CR, Cooke RJ (1998) DNA profiling and plant variety registration. III. The statistical assessment of distinctness in wheat using amplified fragment length polymorphisms. *Euphytica* 102:335–342
- Law JR, Cooke RJ, Reeves JC, Donini P, Smith JSC (1999) Most similar varieties as a grouping tool. *Plant Var Seeds* 12:181–190
- Lee D, Reeves JC, Cooke RJ (1996a) DNA profiling and plant variety registration. 1. Restriction fragment length polymorphism in varieties of oilseed rape. *Plant Var Seeds* 9:181–190
- Lee D, Reeves JC, Cooke RJ (1996b) DNA profiling and plant variety registration. 2. The use of random amplified DNA polymorphisms to discriminate between varieties of oilseed rape. *Electrophoresis* 17:261–265
- Mailer RJ, Scarth R, Fristensky B (1994) Discrimination among cultivars of rapeseed (*Brassica napus* L.) using DNA polymorphisms amplified from arbitrary primers. *Theor Appl Genet* 87:697–704
- Mantel N (1967) The detection of disease clustering and a generalised regression approach. *Cancer Res* 27:209–220
- Mitchell SE, Kresovich S, Jester CA, Hernandez CJ, SzewcMcFadden AK (1997) Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. *Crop Sci* 37:617–624
- Mundges H, Diederichsen E, Kohler W (1989) Comparisons of isozyme patterns in resynthesized amphihaploid rapeseed (*Brassica napus*) and their parental species *Brassica campestris* and *Brassica oleracea*. *Plant Breed* 103:258–261
- Olufowote JO, Xu Y, Chen X, Park WD, Beachell HM, Dilday RH, Goto M, McCouch S (1997) Comparative evaluation of within-cultivar variation of rice (*Oryza sativa* L.) using microsatellite and RFLP markers. *Genome* 40:370–378
- Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Mol Ecol* 4:347–354
- Plaschke J, Ganai MW, Roder MS (1995) Detection of genetic diversity in closely-related bread wheat using microsatellite markers. *Theor Appl Genet* 91:1001–1007
- Plieske J, Strauss D (2001) Microsatellite markers for genome analysis in *Brassica*. I. Development in *Brassica napus* and abundance in *Brassicaceae* species. *Theor Appl Genet* 102:689–694
- Poulsen GB, Kahl G, Weising K (1993) Abundance and polymorphism of simple repetitive sequences in *Brassica napus* L. *Theor Appl Genet* 85:994–1000
- Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant-breeding – RAPDs, microsatellites and machines. *Trends Genet* 9:275–280
- Rakow G, Woods DL (1987) Outcrossing in rape and mustard under Saskatchewan prairie conditions. *Can J Plant Sci* 67:147–151
- Raybould AF, Mogg RJ, Clarke RT, Gliddon CJ, Gray AJ (1999) Variation and population structure at microsatellite and isozyme loci in wild cabbage (*Brassica oleracea* L.) in Dorset (UK). *Genet Resour Crop Evol* 46:351–360
- Roldan-Ruiz I, Dendauw J, Van Bockstaele E, Depicker A, De Loose M (2000) AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). *Mol Breed* 6:125–134
- Saal B, Plieske J, Hu J, Quiros CF, Struss D (2001) Microsatellite markers for genome analysis in *Brassica*. II. Assignment of rapeseed microsatellites to the A and C genomes and genetic mapping in *Brassica oleracea* (L.). *Theor Appl Genet* 102:695–699
- Schneider F (1986) The concept of distinctness in plant breeders' rights. In: Styles BT (ed) *The systematics association special volume. Intraspecific classification of wild and cultivated plants*. Oxford University Press, Oxford, 29, pp 393–397
- Sneath PHA, Sokal RR (1973) *Numerical taxonomy*. WH Freeman and Co, San Francisco
- SzewcMcFadden AK, Kresovich S, Bliet SM, Mitchell SE, McFerson JR (1996) Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated *Brassica* species. *Theor Appl Genet* 93:534–538
- Tessier C, David J, This P, Boursiquot JM, Charrier A (1999) Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* (L.) *Theor Appl Genet* 98:171–177

- Treuren van R (2001) Efficiency of reduced primer selectivity and bulked DNA analysis for the rapid detection of AFLP polymorphisms in a range of crop species. *Euphytica* 117:27–37
- U N (1935) Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and its peculiar mode of fertilisation. *Jpn J Genet* 7:784–794
- UPOV (Union for the Protection of New Varieties of Plants) (1978) International convention for the protection of new varieties of plants test guide-lines. UPOV, Geneva, Switzerland
- UPOV (Union for the Protection of New Varieties of Plants) (1994) International convention for the protection of new varieties of plants. UPOV, Geneva, Switzerland C/28/10
- UPOV (Union for the Protection of New Varieties of Plants) (1996) Guidelines for the conduct of tests for distinctness, uniformity and stability. Rape seed (*Brassica napus* L. *oleifera*). UPOV, Geneva, Switzerland TG/36/6
- Uzunova MI, Ecke W (1999) Abundance, polymorphism and genetic mapping of microsatellites in oilseed rape (*Brassica napus* L.). *Plant Breed* 118:323–326
- Waugh R, Powell W (1992) Using RAPD markers for crop improvement. *Trends Biotechnol* 10:186–191
- Weir BS (1990) Genetic data analysis. Methods for discrete population genetic data. Sinauer Associates, Inc, Sunderland, Massachusetts
- White J, Law JR (1991) Differentiation between varieties of oilseed rape *Brassica napus* (L.) on the basis of the fatty acid composition of the oil. *Plant Var Seeds* 3:125–132