Theor Appl Genet (1999) 98: 1029–1035

I. Divaret · E. Margalé · G. Thomas

RAPD markers on seed bulks efficiently assess the genetic diversity of a *Brassica oleracea* L. collection

Received: 27 August 1998 / Accepted: 29 September 1998

Abstract The concept of a core collection was elaborated to fit the necessity of optimizing the management, for both conservation and use, of genetic resources in sizeable collections. This approach requires an analysis of how the genetic variability is structured among the accessions. The large number of heterogeneous populations in our collection of *Brassica oleracea* makes genetic diversity studies based on plant-to-plant analysis impracticable. To overcome this limitation, the variability analysis by RAPD on seed bulks was investigated for its efficiency in assessing the structure of the genetic diversity of this collection. The optimal bulk size and the bulking or sampling variation were evaluated with bulks of different size and with replicated samples. A mixture of known genotypes was also used to characterise the band detection in bulks, and to compare the plant-to-plant and the bulk methods. Forty seeds were chosen to represent each population. In such a bulk, the detection of bands depended on the proportion of the genotype they were derived from in the mixture. Intense and frequent bands were detected in the bulk with a 15% detection limit. The observed bulking or sampling variation within populations was smaller than the variation between populations, leading to an efficient separation of populations with a clustering of all samples of the same population. The distances calculated from bulk data were highly correlated with the distances based on the plant-to-plant analysis. We demonstrated that RAPD on seed bulks can be used to describe the genetic diversity between populations.

Key words *Brassica oleracea* L. · RAPD · Seed bulk · Genetic resources · Genetic variability

Communicated by H. F. Linskens

I. Divaret · E. Margale[1] · G. Thomas (✉)
Institut National de la Recherche Agronomique,
Station d'amélioration des Plantes,
Domaine de la Motte-BP 29,
F-35653 Le Rheu Cedex, France
Fax: +33 02 99 28 51 20
E-mail: gthomas@epi.roazhon.inra.fr

Present address:
[1] Serasem Recherche et sélection végétales,
60, rue Léon Beauchamp-BP 45,
F-59933 La Chapelle D'Armentieres Cedex, France

Introduction

Genetic resources can ideally be stored in *ex situ* collections whose objective is to maintain as many alleles as possible at the disposal of the plant breeder (Hayward et al. 1997). Often, the large size of a collection makes its management and use difficult, especially for outbreeding species. To reduce this size, a core collection consisting of a limited set of accessions chosen to represent the genetic spectrum of the whole collection has to be developed (Brown 1995). The first step in this approach is to structure the collection into groups from which stratified sampling can be established (Yonezawa et al. 1995).

To preserve the genetic resources of cultivated cole, *Brassica oleracea* L., endangered by the supremacy of hybrids and lines over populations in breeding programs, INRA (Rennes) has established a collection of French landraces. This collection re-groups 1000 very diversified local populations of cauliflower, cabbage and kale. These populations result from mass selection carried out by growers over many years, sometimes for more than a century (Hervé 1987). Up to now, these accessions have been incompletely described by growers, mostly in terms of quality and earliness, based on field observations made under different conditions. To optimise the conservation and the exploitation of this potential, it would be very helpful to structure the genetic diversity of these populations using neutral and reliable markers.

RAPD markers are widely employed to reveal genetic variation in many plant species. In *Brassica*, RAPD markers are considered to be as efficient as RFLP markers for estimating intraspecific genetic relationships among genotypes (Demeke et al. 1992; Thormann et al. 1992; Dos Santos et al. 1994). RAPD markers have also been used to quickly estimate genetic distances between *B. oleracea* accessions (Hu and Quiros 1991; Kresovich et al. 1992; Phippen 1997). Because of the efficiency and the convenience of this technique, we have decided to analyse our collection with RAPD markers.

As *B. oleracea* is a strictly allogamous species, populations are expected to be very heterogeneous. To analyse diversity in such outbreeding species, some authors studied 6–12 to 20, and up to 30 plants per accession (Crochemore et al. 1996; Lannèr-Herrera et al. 1996; Phippen 1997) according to the observed variability within accessions. Because of the large number of accessions in our collection it would be laborious to individually analyse several plants per population. Our aim is to determine whether RAPD analysis on bulks, especially on seed bulks, generates enough genetic information in a short time at a cheaper cost. To achieve this aim, we concentrate on two features. Firstly, we focus on the characteristics of RAPD patterns obtained from seed bulks. For that purpose, we tested the sampling methodology for the optimal bulk size, the detection limit of specific bands, and the representation of a population by its seed-bulk pattern. Secondly, we discuss the efficiency of this technique to analyse the genetic variability of our collection. We estimated the random variation occurring with the sampling or the bulking, and we compared genetic distances calculated from bulk data to genetic distances based on a plant-to-plant analysis.

Materials and methods

Plant material

A sample of seven populations of *B. oleracea* L. taken from the collection of landraces was employed. These populations were chosen to represent the different crops (cauliflower, cabbage and kale) of the collection, which are specifically characterised by their horticultural traits (Table 1). This sample of populations illustrates the diversity of the collection. We also used one F_1 hybrid and one doubled-haploid as known homogeneous genotypes.

DNA extraction

For plant-to-plant analysis, DNA was extracted from leaves using a modified Dellaporta et al. (1983) method. The same protocol was adapted for DNA extraction from seed bulks as follows. The seeds were ground in 2 ml of a cold buffer (20 mM Na metabisulphite, 0.5 M sorbitol, 0.1 M trizma-base, 70 mM Na_2EDTA, pH adjusted to 7.5 with HCl). After centrifugation, 2 ml of extraction buffer (100 mM TRIS-pH 8, 50 mM EDTA-pH 8, 500 mM NaCl, 1% SDS, 1% β-mercaptoethanol) was added to the pellet. The homogenate was treated as described in Dellaporta et al. (1983) bypassing the phenol-chloroform step. The DNA was precipitated by 5 M NaCl and absolute ethanol. The pellet was washed with 70% ethanol and, after air drying, was dissolved in TE buffer. For some samples (especially for kale), another DNA precipitation by 5 M NaCl and absolute ethanol was necessary.

DNA amplification

PCR was realised in a 10-µl vol containing 67 mM Tris-HCl (pH 8.8), 16 mM $(NH_4)_2SO_4$, 0.01% Tween 20, 2 mM $MgCl_2$, 0.15 mM of each dNTP, 0.25 µM of primer, 15 ng of genomic DNA and 0.4 units of *Taq* DNA polymerase (EUROBIOTAQ). Amplifications were performed in a DNA Thermal Cycler (Perking Elmer Cetus) programmed for an initial step of 30 s at 94°C, followed by 45 cycles with 30 s at 92°C, 1 min at 35°C, 2 min at 72°C, and 5 min at 72°C. Eight 10-mer primers from Operon Technology kits (Alameda, California) were used (OPA01, OPB07, OPC09, OPC13, OPD01, OPE08, OPE20). These primers were chosen because they generated simple (less than five RAPD) and complex (up to 15 RAPD) profiles.

Genetic distance calculation and classification

Only clearly resolved and repeatable bands were considered as RAPD markers after repeating several amplifications three times, with different thermocyclers and with duplicate samples in the same manipulation. For each pattern RAPD bands were scored as 0 for absent or 1 for present. The Jaccard (DJ) distance was

Table 1 Description of the plant material used in this study

Genotype[a]	Type	Crop	Horticultural characteristics	Convar. var
JF112	Population	Cauliflower	Winter (Jan-Feb)	*botrytis botrytis*
CM56.18A	Population	Cabbage	Savoy	*capitata sabauda*
CSA67.14	Population	Cabbage	Sauerkraut	*capitata capitata*
KM35.03	Population	Kale	Half-marrow stem	*acephala medullosa*
KB22.14	Population	Kale	Branchy	*acephala ramosa*
KL22.17	Population	Kale	Leafy	*acephala sabellica*
KY29.01	Population	Kale	Yellow	*acephala acephala*
H	F_1 hybrid	Kale	Half-marrow stem	*acephala medullosa*
G	Doubled haploid	Cauliflower	?	*botrytis botrytis*

[a] The population name is coded according to its horticultural characteristics and origin. For example, KL22.17 is a leafy Kale which has been collected in French department No. 22, (Côtes d'Armor). JF112 is an accession of cauliflower harvested in January-February

calculated using the following formula:

$$DJ = 1 - (1,1)/[(1,1) + (0,1) + (1,0)],$$

where (1, 1) is the number of bands shared by two patterns, and (0, 1) + (1, 0) the number of bands specific to one of the patterns. In the plant-to-plant method, the distance between two populations was given by the average of the ($n_x n_y$) distance values between the n_x and n_y individuals of the two populations x and y as described by Crochemore et al. (1996).

The distances were compared using a non-parametric analysis (proc npar1way option wilcoxon) of the Statistical Analysis System, SAS (1989). The rank correlation between distances was calculated with the "proc corr kendall" procedure of SAS. Dendrograms were generated from the distance matrices by the UPGMA option of the program NEIGHBOR or PHYLIP (Felsentein 1993) and drawn with TREEVIEW (Page 1996).

Results

Determination of the optimal seed bulk size

The optimal bulk size was determined using three populations of kale, the most variable crop of our collection as revealed by morphotype and molecular variability (Margalé et al. 1994). Bulks of 10, 20, 30, 40, 50, 60, 70 and 80 seeds of KM35.03, KY29.01 and KL22.17 were composed and referred to as DSN (different seed number) samples. The DSN sample patterns obtained with six primers (OPA01, OPB07, OPB15, OPC09, OPC13, OPE20) were compared for the three populations (Table 2). Only one given bulk size was studied for each population, but the three populations served as three repetitions. The number of bands in the DSN patterns was not significatively different ($P = 0.4767$).

Detection of bands specific to a genotype according to its proportion in a bulk of 40 seeds

The two known homogeneous genotypes, G and H, were mixed in different proportions to constitute different bulks and coded as GxHy, with x and y representing the number of seeds of each genotype in a 40-seed bulk (x = 0, 1, 5, 10, 15, 20, 30 or 40 seeds). Overall, 30 bands were obtained in the patterns of the GxHy samples with the same primers as previously employed. Among them, 14 bands were specific to H, 15 bands were specific to G, and one band was found to be present only in the bulks but absent in the initial genotypes. The number of detected bands specific to one genotype increased with the proportion of this genotype in the bulk (Fig. 1). Any genotype was preferentially amplified, i.e. the percentage of total bands detected was the same for a given proportion in the bulk (G10H30 and G30H10, 79% and 76% of total bands detected respectively). The detection depended on the intensity of the band, i.e. the bands detected in the bulk with a low threshold were intense. This technique allows one to detect rare bands in bulks; thus 33.3% of bands specific to G were detected when the proportion of this genotype in the bulk was 2.5% (1 seed out of 40 seeds). Half of the bands were detected when the proportion of the corresponding genotype in the bulk was 25%. More precisely, for genotype G, for which smaller proportions in the mixture were studied, 66.6% of the bands were detected when this proportion was 12.5%.

Representation of a population by a bulk of individuals

For individual analysis, 35–40 plants per population were studied for KL22.17, CM56.18A, JF112, KY29.01 and CSA67.14, with the eight primers. To compare the bulk method to an individual analysis for the same sample, equal amounts of DNA from each plant were mixed. A bulk of 40 seeds for each population was also used to compare bulking methods. In the individual analysis, we obtained a total 57 polymorphic bands with frequencies varying from 5 to 100%. In bulks, involving a mixture of DNA or a mixture of 40 seeds, only some of these bands were detected (65–85% of the

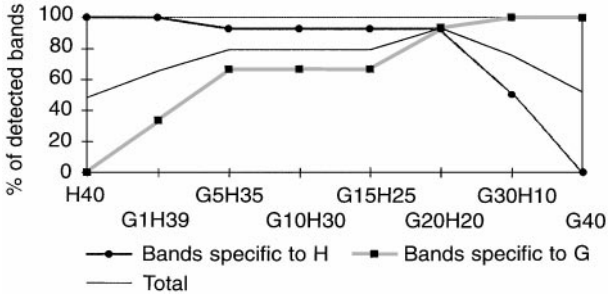

Fig. 1 Detection of specific bands according to the proportion of the genotype they were derived from in a 40-seed bulk. Two homogeneous genotypes, G and H, were mixed in different proportions (GxHy, with x seeds of G and y seeds of H). The percentage of detected bands specific to one genotype is calculated for different proportions of G and H in the bulk. The percentage of total bands detected in the two genotypes is also shown for each bulk

Table 2 Number of RAPD markers according to the bulk size for three populations

Population	10 seeds	20 seeds	30 seeds	40 seeds	50 seeds	60 seeds	70 seeds	80 seeds
KM35.03	21	21	21	21	21	18	21	21
KY29.01	19	19	20	20	20	19	20	19
KL22.17	20	20	21	21	20	20	19	20

total number of bands according to the population). The detection threshold was similar in the bulks of DNA as in the bulks of seeds: 15 to 27.5% according to the population (Table 3). It depended on the intensity of the bands that was different according to the population they derived from. For CM56.18A, one intense band present in 15% of the individuals was detected in the bulk, whereas a weak band present in 70% of the individuals was not detected (Table 3). However this observation was not systematic for all the detected bands but depended on the band considered. In bulks, the detected bands were those that were most intense or most frequent.

Effect of the variation due to sampling on the distances calculated within populations

The mean of the distances between DSNs was calculated for each DSN sample and each population (Table 4). For KY29.01 and KL22.17, no significant effect of the number of seeds was found on this distance ($P = 0.094$ and $P = 0.134$, respectively), but for KM35.03 a significant difference was found ($P = 0.002$). The bulk of 60 seeds was different from the others in this population. The distances between DSN samples within each population were significantly different from 0 ($P < 0.05$).

To investigate random variation occurring with the sampling or the bulking, five bulks of 40 seeds of the same population (KB22.14) were amplified with the eight primers. The number of bands was similar except for one sample (Table 5). The mean of the distances between the five samples was 0.107 ± 0.087 (0.053 ± 0.024 without sample 5). The variation between samples for one population (distance = 0.053) was found to be similar to the variation between DSN samples (average distance for the three populations = 0.047). Thus, there was an effect of sampling, but not of bulk size, on the distance calculations within populations.

Effect of the variation due to sampling on the distances calculated between populations

The distances between the three populations were calculated for all the DSN samples. The populations were clearly separated with a clustering of all the samples of

Table 3 Comparison of individual and bulk methods for polymorphism evaluation with RAPD markers

Marker	Individual method				Bulk methods					
	Number of detected bands	Number of bands with a frequency of		Average distance between individuals	Number of detected bands		Lowest pi[a] of detected bands		Highest pi[a] of undetected bands	
		< 10%	> 95%		DNA	Seeds	DNA	Seeds	DNA	Seeds
KL22.17	44	4	18	0.25	31	31	27.5	25	100	100
CM56.18A	50	3	13	0.36	36	33	15	15	70	70
JF112	34	2	12	0.25	28	29	17.5	17.5	70	65
KY29.01	46	1	7	0.44	30	30	17.9	17.9	48.8	48.7
CSA67.14	38	3	12	0.31	27	26	15.7	15.7	78.9	73.6
Total	57				39	40				

[a] pi = percentage of individuals sharing a particular band in the population as measured by the individual method

Table 4 Average distance of a given sample to all the other DSN[a] samples within three populations

Population	10 seeds	20 seeds	30 seeds	40 seeds	50 seeds	60 seeds	70 seeds	80 seeds	Mean	SE[b]
KM35.03	0.02	0.02	0.02	0.02	0.02	**0.12**	0.02	0.02	0.03	0.04
KY29.01	0.09	0.09	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.02
KL22.17	0.05	0.05	0.04	0.04	0.05	0.06	0.10	0.07	0.06	0.02

[a] DSN samples: different seed number samples, i.e. samples with different bulk sizes
[b] SE: standard error

Table 5 Number of RAPD markers for five samples of 40 seeds of KB22.14

Item	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
Number of bands	51	52	50	51	**40**

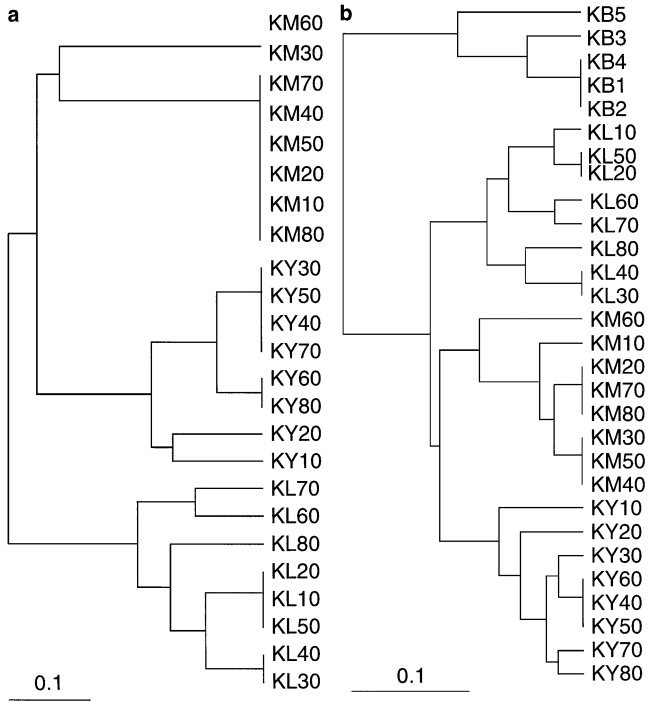

Fig. 2a, b UPGMA classification of different samples of three **a** and four **b** populations to evaluate the effect of the variation occurring with bulk sampling on diversity estimation. The scale represents an 0.1 unit of the Jaccard distance. The first two letters of the populations is followed: by the number of seeds in the bulk for KM35.03 (KM10 to KM80), KY29.01 (KY10 to KY80) and KL22.17 (KL10–KL80), and by the sample number for KB22.14 (KB1–KB5)

the same population, in the classification (Fig. 2a). The distances between the five samples of KB22.14 and the DSN samples were also calculated to test the effect of sampling variation on the distance between populations. The four populations were also clearly separated (Fig. 2b). However, when only one DSN sample per population was used, differences in the classification of the three populations were observed when using distance data from samples with 20 and 60 seeds.

Comparison of the distances calculated from data of plant-to-plant and bulk methods

The analysis of individuals gave added information through the intrapopulation distances (Table 3). The distances calculated from bulked DNA and seed samples were very highly correlated ($r^2 = 0.97$, $P = 0.0001$). The distances calculated with seed and DNA bulks were significantly correlated with the distances calculated by the individual method ($r^2 = 0.87$, $P = 0.009$ and $r^2 = 0.79$, $P = 0.006$, respectively). As expected, given these correlations, the UPGMA clustering obtained with both methods was similar (Fig. 3).

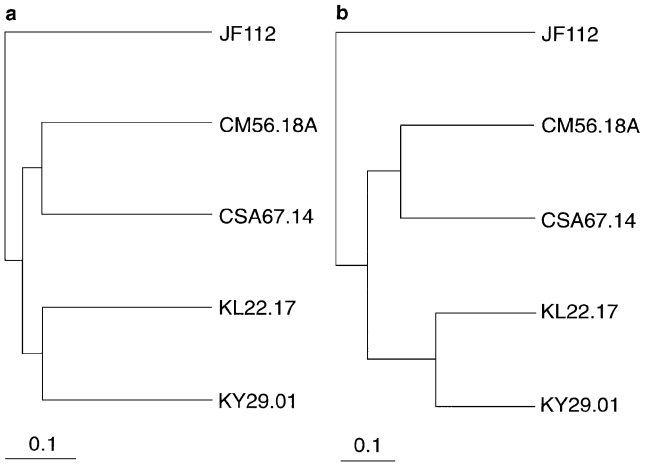

Fig. 3a, b Comparison of the UPGMA classification obtained by individual **a** and bulk **b** methods. The scale represents an 0.1 unit of the Jaccard distance

Discussion

Production of a representative pattern of a population by bulking 40 seeds

In this study, no major differences were observed between the patterns obtained with 10–80 seed samples from the same population. In the mixtures, the lowest detection threshold was 2.5% and 15% with two (G and H) and 40 individuals (for the populations), respectively. The detection of a specific band in a bulk depended on its intensity and on its proportion in the bulk or its frequency in the population.

RAPD analysis on seed bulks gave representative patterns for the populations. Scoring the most representative bands of one population is not necessarily equivalent to scoring all the bands existing in this population. The more frequent bands were characteristic of one population as well as the intense bands whose intensity varied with the population. Intense bands are more representative of a genotype than weak bands, since they result in a higher degree of homology. Indeed, intense bands in a bulk are not related to a high concentration of template in the bulk resulting from their widespread distribution in the population as postulated by Yu and Pauls (1993). RAPD marker intensity is associated with the degree of homology between primer and template or with the amplification of other fragments in the sample (Thorman et al. 1994). As a matter of fact, the intensity of bands did not increase with the proportion of these bands in the bulk. Bands found to be intense in the bulks were also intense in individuals and, inversely, frequent weak bands in the populations were still weak in the bulks. Additionally, the fact that some weak frequent bands were not amplified in the bulk may be due to the competition of hybridisation sites with a better homology sequence for

the primer. This difference in the degree of homology between intense and weak bands can explain why the most intense bands are considered to be more reliable and why Weeden et al. (1992) found a 4% error rate occurring with bands of faint or intermediate intensity. Dos Santos et al. (1994), using only bands classified as intense or medium for the analysis of *B. oleracea*, found that sampling errors were equivalent for RAPDs and RFLPs with respect to genetic-similarity determination.

Although no major difference was observed between the different bulk sizes, 40 seeds were used because variations were observed for small samples, i.e. for 10–30 seeds in a previous experiment with a cabbage population (data not shown). Moreover, bulks of 40 individuals provide enough DNA for further amplifications to avoid sampling or bulking variation.

In a bulk of this size, we did not find genotype-dependent competition as described by Halldén et al. (1996) in *Brassica napus*. The two genotypes, G and H, were equally amplified according to their proportion in the bulk. One band was noted only in the bulk. This band might be the result of an interaction between two allelic sequences of different length leading to heteroduplex formation (Halldén et al. 1996). A comparable case was not observed in the 40 individual bulks for which each individual was analysed. In bulks of 40 individuals, RAPD fragments with a frequency lower than 15% were not detected. This minimum detection limit is close to the 10% limit given by Michelmore et al. (1991). This limit was much lower (2.5%) when only two genotypes, G and H, were bulked. The relation between detection threshold and bulk complexity is, however, unclear. For populations, the intrapopulation variation reflects the bulk complexity; yet the lowest detection threshold was not observed for the population with the lowest intrapopulation variation. This detection threshold depended on the bands. Similar results were obtained by Virk et al. (1994) on rice where the minimum frequency required for a band to be visible in the bulk varied for different bands.

Application of bulk methods to genetic-variability analysis

Within a population, we found a variation between same-size samples that is similar to the variation between samples of different sizes. This variation was lower than the variation between populations and did not impede the clustering of all samples of the same population. Whatever the method employed (DNA or seed bulk, plant-to-plant analysis) the same classification was obtained even if the simplified bulk method detected less polymorphism. We also found that mixing 40 DNAs in equal amounts was equivalent to mixing 40 seeds. Using seeds instead of leaves is time saving and less expensive (McDonald et al. 1994). The plants do not have to be individually grown and DNA can be extracted directly from seeds with the technical precaution of eliminating polysaccharides by supplementary DNA precipitation. The use of RAPD markers from seeds in a stable manner for varietal identification has already been reported (McDonald et al. 1994) but, to our knowledge, this is the first time that RAPD marker patterns from seed bulks have been characterised for genetic diversity evaluation.

We can assume that the intrapopulation variation in populations more genetically related than the ones we used would also be sufficiently lower than the inter-population variation to not interfere with the classification. Indeed, as the distances are calculated with only polymorphic RAPD markers, the variation between populations is maximised. Thus, closely related accessions appear to be more distantly related in such an analysis where all the populations are compared on the same basis. Moreover, to limit the sampling or bulking error for the estimation of diversity, a large genetic basis, i.e. a large number of accessions, has to be employed to represent the studied diversity and enough markers have to be obtained. Another solution would be to use known replicates as a reference for duplicates. Weeden et al. (1992) even advise the use of duplicate samples from each individual tested and to score only the fragments amplified in both.

The aim of a classification is to assign each accession to only one class in so far as all the accessions belonging to the same class have an identical, or almost identical, allelic composition (Yonezawa et al. 1995). In an intraspecific comparison within the genus *Brassica*, the presence of a RAPD band in two individuals indicates a high level of sequence homology in a reliable and repeatable manner (Dos Santos et al. 1994; Lannèr-Herrera et al. 1996). In our study, we showed that RAPD analysis on bulks allows these comparisons to predict the diversity existing in the collection.

The constitution of a core collection needs the hierarchical stratification of the collection into groups. The seed-bulk method successfully described the genetic variability of our collection with a grouping of accessions consistent with the morphological, agronomic and geographical data (Margalé et al. 1995). Once this classification is established, the choice of accessions within the groups needs to be discussed, considering the fact that rare bands are not detected in bulk. In a core collection based on a neutral-allele model, the sampling procedure was developed to retain rare and widespread alleles (Hamon et al. 1995). But, as the neutrality principle may not hold for many genes, other sampling strategies exist (Yonezawa et al. 1995). These strategies take into account the redundancy, the genetic structure of the population, or the allelic diversity that can be estimated using RAPDs on bulks.

The seed-bulk method we described will enable us to reveal the structure of the genetic diversity of our collection. We expect to develop a core collection by

sampling accessions in the groups revealed by the classification based on RAPD variation. Moreover, RAPDs will give us information on the evolutionary history of groups of populations in order to understand how traditional breeding may have contributed to the structure of variability.

Acknowledgements The authors are indebted to F. Baron who is in charge of the management of the collection and who provided seeds of the populations and of H (P230 × 429). We thank P. Chatelet, for kindly providing seeds of G (G49.94). We also thank P. Barret for his valuable technical assistance. We are grateful to R. Delourme and N. Chantret for their corrections, as well as to M. Manzanarès-Dauleux and G. Charmet for their critical reading of the manuscript and helpful discussions.

References

Brown AHD (1995) The core collection at the crossroads. In: Hodgkin T, Brown AHD, Van Hintum ThJL, Morales EAV (eds) Core collections of plant genetic resources. John Wiley and Sons, Chichester, UK, pp 3–19

Crochemore ML, Huyghe C, Kerlan MC, Durand F, Julier B (1996) Partitioning and distribution of RAPD variation in a set of populations of the *Medicago sativa* complex. Agronomie 16: 421–432

Dellaporta S, Wood J, Hicks J (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1: 19–21

Demeke T, Adams RP, Chibbar R (1992) Potential taxonomic use of random amplified polymorphic DNA (RAPD): a case study in *Brassica*. Theor Appl Genet 84: 990–994

Dos Santos JB, Nienhuis J, Skroch P, Tivang J, Slocum MK (1994) Comparison of RAPD and RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotypes. Theor Appl Genet 87: 909–915

Felsenstein J (1993) PHYLIP (PHYLogenetic Inference Package) version 3.5 c (distributed by the author). Department of Genetics, University of Washington, Seattle, Washington, USA

Hamon S, Dussert S, Noirot M, Anthony F, Hodgkin T (1995) Core collections—accomplishments and challenges. Plant Breed Abstr 65: 1125–1133

Halldén C, Hansen M, Nillsson NO, Hjerdin A, Sall T (1996) Competition as a source of errors in RAPD analysis. Theor Appl Genet 93: 1185–1192

Hayward MD, Sackville Hamilton NR (1997) Genetic diversity, population structure and conservation. In: Callow JA, Ford-Lloyd, Newbury HJ (eds) Biotechnology and plant genetic resources, Cab International, pp 49–76

Hervé Y (1987) Evolution des crucifères cultivées et préservation des resources génétiques en France. CR Acad Agric Fr 73: 79–93

Hu J, Quiros C (1991) Identification of broccoli and cauliflower cultivars with RAPD markers. Plant Cell Rep 10: 505–511

Kresovich S, Williams JGK, Mc Ferson 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

Lannèr-Herrera C, Gustafsson M, Fält AS, Bryngelsson T (1996) Diversity in natural populations of wild *Brassica oleracea* as estimated by isozyme and RAPD analysis. Genet Resour Crop Evol 43: 13–23

McDonald MB, Elliot LJ, Sweeney PM (1994) DNA extraction from dry seeds for RAPD analyses in varietal identification studies. Seed Sci Technol 22: 171–176

Margalé E, Chèvre AM, Delourme R, Hervé Y (1994) Description de la diversité génétique chez *Brassica oleracea* à l'aide de marqueurs isoenymatiques et moléculaires. Genet Sel Evol 26, suppl 1: 137s–153s

Margalé E, Hervé Y, Hu J, Quiros CF (1995) Determination of genetic variability by RAPD markers in cauliflower, cabbage and kale local cultivars from France. Genet Resour Crop Evol 42: 281–289

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

Page RDM (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. Comput Appl Biosci 12: 357–358

Phippen WB (1997) Molecular characterization can quantify and partition variation among genebank holdings: a case study with phenotypically similar accessions of *Brassica oleracea* var. *capitata* L. (cabbage) 'Golden Acre'. Theor Appl Genet 94: 227–234

SAS institute (1989) SAS/STAT user's guide, version 6, 4th edn. SAS institute, Cary, North Carolina, USA

Thormann CE, Ferreira ME, Camargo LEA, Tivang JG, Osborn TC (1994) Comparison of RFLP and RAPD markers for estimating genetic relationships within and among cruciferous species. Theor Appl Genet 88: 973–980

Virk PS, Ford-Llyod BV, Jackson MT, Newbury HJ (1994) Use of RAPD for the study of diversity within plant germplasm collections. Heredity 74: 170–179

Weeden NF, Timmerman GM, Hemmat M, Kneen BE, Lodhi MA (1992) Inheritance and reliability of RAPD markers. In: Application of RAPD technology to plant breeding. Proc Symp, Joint Plant Breed Symp Series, Minneapolis, Minnesota, pp 12–17

Yonezawa K, Nomura T, Morishima H (1995) Sampling strategies for use in stratified germplasm collections. In: Hodgkin T, Brown AHD, Van Hintum ThJL, Morales EAV (eds) Core collections of plant genetic resources. John Wiley and Sons, Chichester, UK, pp 35–53

Yu K, Pauls KP (1993) Rapid estimation of genetic relatedness among heterogeneous populations of alfafa by random amplification of bulked genomic DNA samples. Theor Appl Genet 86: 788–794