Theor Appl Genet (1997) 94: 227–234

W. B. Phippen · S. Kresovich · F. G. Candelas
J. R. McFerson

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'

Received: 10 June 1996 / Accepted: 23 August 1996

Abstract To better characterize and conserve crop genetic resources, the assessment of genetic identity, relatedness, and structure among entries and collections becomes a priority. In the present study, a random amplified polymorphic DNA (RAPD) assay was applied as a quick, cost-effective, and preliminary screen to quantify and partition the molecular variation among accessions. Fourteen phenotypically uniform accessions of *Brassica oleracea* var. *capitata* L. (cabbage) similarly designated as 'Golden Acre' were tested with nine decamer oligonucleotide primers. These amplifications generated 110 fragments, of which 80 were polymorphic ranging in size from 370 to 1720 bp. The 80 polymorphic fragments were sufficient to distinguish between all 14 accessions. Data based on the partitioning of variation among accessions indicated that 'Golden Acre' entries could be reduced to as few as four groups, with the potential loss of variation being only 4.6% of the absolute current genetic variation in those holdings as estimated from RAPD analysis. This proposed grouping would concurrently save approximately 70% [$750–1000 (US) per accession] for each cycle of regeneration (approximately 20–25 years at most) which alternatively could then be used for other priorities in *B. oleracea* conservation and use. This case represents but one example where targeted use of a molecular-marker assay linked with rigorous statistical analysis will be useful for plant genebank management, particularly for questions at the intraspecific level. Molecular markers will provide genebank curators with additional sources of information to better plan and organize collection holdings and use finite financial support in a more effective manner.

Key words AMOVA · Conservation · Curation · Genetic markers · Molecular genetic screening · RAPD

Communicated by G. Wenzel

W. B. Phippen J. R. McFerson
USDA-ARS, Plant Genetic Resources Unit, Geneva, NY 14456-0462, USA

S. Kresovich (✉)
USDA-ARS, Plant Genetic Resources Conservation Unit, Griffin, GA 30223-1797, USA

F. G. Candelas
University of Georgia, Department of Genetics, University of Valencia, Valencia 46100, Spain

Introduction

Genebank curators must meet increasing demands from the user community, while resources allocated to support their activities continue to diminish. More comprehensive assessment of genetic identity, relatedness, and structure among entries and collections would allow curators to manage *ex situ* collections more efficiently (Kresovich et al. 1992). For example, genetic information could be applied to discriminate redundancies in a collection. Accessions considered redundant could then be bulked or eliminated, thereby decreasing maintenance costs. Although redundancies are ubiquitous among collections worldwide, curators have only circumstantial or preliminary evidence with which to address the problem. This situation is particularly troublesome in taxa like *Brassica* spp., where complex differences in growth form, life cycle, reproductive habit, and pollination systems require expensive seed-regeneration protocols (current cost ranging from $750–1000 U.S. for seed production and processing) on a recurring basis. However, molecular characterization can provide the genetic resolution necessary to resolve the complex web of relationships among genebank accessions (Soller and Beckman 1983; Epplen et al. 1991; Weising et al. 1991; Beyermann et al. 1992; Kaemmer et al. 1992; Kresovich et al. 1992, 1993; Nienhuis et al. 1993; Virk et al. 1995; Charters et al. 1996).

In the present investigation, a random amplified polymorphic DNA (RAPD) assay (Williams et al. 1990) was applied to characterize an array of phenotypically similar *Brassica oleracea* var. *capitata* (cabbage) accessions identified either as 'Golden Acre' or as similarly named selections therefrom. 'Golden Acre' is an historically impor-

tant, open-pollinated cultivar originally selected in the U.S. around 1920 from 'Copenhagen Market', a Danish cultivar originating in 1909 as a single plant selection from the German cultivar 'Ditmarscher'. 'Golden Acre' once enjoyed worldwide prominence as an early, round-headed type and was widely used in breeding programs. Seed companies maintained proprietary stocks while both private and public plant breeding programs developed named selections, e.g., *Fusarium* yellows-resistant lines from J.C. Walker and co-workers at the University of Wisconsin in the 1940s (P.H. Williams, personal communication). Worldwide, genebanks hold many accessions identified as 'Golden Acre' or as a selection of 'Golden Acre'. These 'Golden Acre'-types may be redundant to some degree; maintaining them as individual entries greatly taxes scarce financial and physical resources dedicated to maintenance and regeneration priorities. Our primary objectives were to employ a quick, cost-effective molecular method to characterize these apparently closely related, heterogeneous accessions and also be able to make curatorial decisions based on the data when linked with rigorous statistical analysis.

Materials and methods

Selection of plant material

Fourteen accessions of *B. oleracea* var. *capitata* 'Golden Acre' were utilized as the test array (Table 1). To reduce bias in data generation, analysis, or interpretation, individual accessions were coded to conceal identities. Identifying numbers of 'Golden Acre' accessions were not sequential because four entries had been omitted prior to analysis due to low germination. Two accessions, 'Copenhagen Market' (a progenitor of 'Golden Acre') and 'Jersey Wakefield' (an unrelated pointy-head cabbage), were also included as reference populations.

Plants were grown in the greenhouse employing recommended cultural practices. Three-week-old seedling tissue was harvested into liquid nitrogen and stored at –70°C. All leaf tissue was subsequently dried in a Labconco 6-l freeze drier (Labconco, Kansas City, Mo., USA) and stored at –20°C until needed.

DNA isolation

Genomic DNA extractions were performed using a ball bearing CTAB extraction method (Dellaporta 1983; Colosi and Schaal 1993). The amount of DNA extracted was quantified on a TKO 100 fluorometer (Hoefer Scientific Instruments, San Francisco, Calif., USA) and confirmed in an agarose gel following restriction enzyme digestion. The DNA extraction procedure, using 50 mg of dried leaf tissue per plant, produced from 20 to 100 mg of DNA.

DNA amplification

After an initial screen of 80 decamer oligonucleotide primers (Operon Technologies, Inc., Alameda, Calif., USA), nine primers were utilized for the amplification of random DNA sequences from the test array (see Table 2). Primer selection was based on the information content, clarity, and reproducibility of banding patterns in the 350–2000-bp range.

The protocol for DNA amplification reported by Kresovich et al. (1992) was followed. Reaction mixtures with a total volume of 25 µl contained 10 mM Tris-HCl (pH 8.3); 50 mM KCl; 2.0 mM $MgCl_2$; 0.001% gelatin; 0.2 mM each of dATP, dCTP, dGTP, and dTTP (Perkin Elmer Cetus, Norwalk, Conn., USA); 0.64 µM of a single decamer primer (Operon Technologies, Inc., Alameda, Calif., USA); approximately 25 ng of genomic DNA; and 0.625 units of DNA *Thermus aquaticus* (*Taq*) polymerase (Perkin Elmer Cetus, Norwalk, Conn., USA). Reaction mixtures were incubated in a Perkin Elmer Cetus 9600 DNA thermal cycler (Perkin Elmer Cetus, Norwalk, Conn., USA). The thermal cycler was programmed for 45 cycles of 1 min at 94°C, 5 min at 38°C, a 3-min ramp to 72°C, and 2 min at 72°C. The reaction mixture, without genomic DNA, was run with each amplification as a negative control. A positive control, lambda DNA (BRL, Bethesda, Md., USA), also was included. Amplification products were resolved by gel electrophoresis in 2.0% agarose in 1×TBE (89 mM Tris pH 8.0, 89 mM boric acid, 2 mM disodium EDTA). Molecular sizes of the amplification products were estimat-

Table 1 Phenotypically similar accessions of *B. oleracea* var. *capitata* L. 'Golden Acre' assayed for molecular variation

Assigned number	Cultivar	Origin	Source	Accession	Number of plants in study
1	Extra Early Golden Acre	USA	Joseph Harris Co.	NSL 22005.01	6
2	Golden Acre	GBR	National Vegetable Research Sta.	WGB 2006	11
3	Golden Acre	GBR	D T Brown and Co.	WGB 3709	8
4	Golden Acre	GBR	Sharpes International Seeds Ltd.	WGB 3710	12
5	Golden Acre	GBR	Sinclair McGill Ltd.	WGB 3348	12
6	Golden Acre 84	USA	Desert Seed	NSL 5969.01	8
7	Golden Acre 84	DNK	Chinese Agricultural Assoc.	PI 432718	6
9	Golden Acre sel 84	GBR	Asmer Seeds Ltd.	WGB 3762	12
10	Golden Acre sel Ditmarsh	GBR	Sharpes International Seeds Ltd.	WGB 2554	8
11	Golden Acre sel Earlibird	NLD	Holland Select BV	WGB 8333	8
12	Golden Acre sel Primo	GBR	Asmer Seeds Ltd.	WGB 3713	8
14	Golden Acre Yellows Resistant	USA	Northrup, King and Co.	NSL 5970.01	8
17	Resistant Golden Acre	USA	Ferry-Morse Co.	NSL 5994.01	8
18	Wisconsin Golden Acre	USA	Northrup, King and Co.	NSL 6005.01	7
CM	Copenhagen Market	USA	American Takii	G 30417	8
JW	Jersey Wakefield	USA	Asgrow	G 30752	8
Total					138

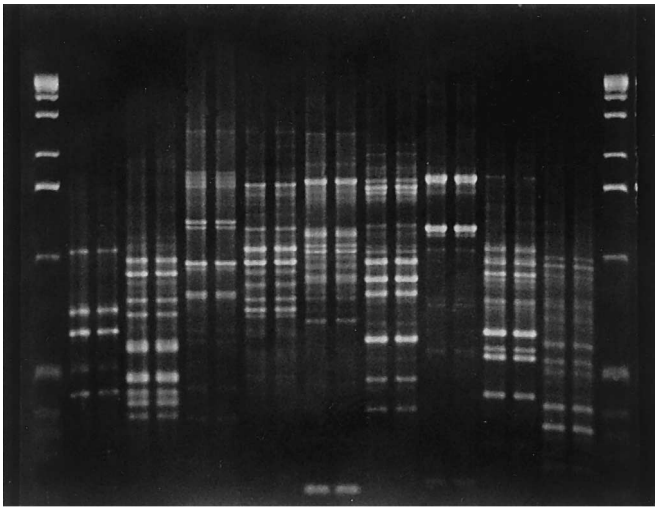

Fig. 1 Agarose gel depicting replicated DNA amplifications of 'Golden Acre' accession 1 with nine selected decamer primers. *Lanes 2,3* were amplified with OPB-04, *lanes 4, 5* were OPB-10, *lanes 6, 7* were OPB-19, *lanes 8, 9* were OPB-20, *lanes 10, 11* were OPV-04, *lanes 12, 13* were OPV-12, *lanes 14, 15* were OPV-13, *lanes 16, 17* were OPW-16, and *lanes 18, 19* were OPY-08. *Lanes 1 and 20* are 1-kb size standards

ed by utilizing a 1-kb DNA ladder (BRL, Bethesda, Md., USA) and a computer program, Size-It!, designed by Dr. Jeffrey Lawrence, Department of Genetics, Washington University, St. Louis, Mo., USA. The gels were stained with ethidium bromide and photographed with black and white film #667 (Polaroid, Cambridge, Mass., USA) under UV light. Reproducibility was checked by running replicate samples across amplifications for each primer (Fig. 1).

Data analysis

Differences within and among accessions were evaluated by means of the molecular analysis of variance (AMOVA) technique (Excoffier et al. 1992). This technique was adapted for usage with RAPD markers by Huff et al. (1993). The Euclidean distance metric between pairs of individuals *x* and *y*, defined by Excoffier et al. (1992) as

$$\varepsilon^2_{x,y} = n\left[1 - \frac{2n_{xy}}{2n}\right]$$

where n is the total number of polymorphic bands and n_{xy} is the number of bands shared by individuals *x* and *y*, was computed for all pairwise combinations of 138 individuals (ranging from 6–12 per accession) (Table 1). The AMOVA procedure was then used to estimate variance components for RAPD phenotypes, partitioning the variation among individuals/within accessions, among accessions/within groups, and among groups ('Golden Acre' accessions versus 'Copenhagen Market' and 'Jersey Wakefield'). The significance level for variance components was tested by means of a nonparametric re-sampling procedure as described in Excoffier et al. (1992). All analyses for the AMOVA procedure were made using the WINAMOVA program version 1.55 provided by L. Excoffier (ftp://acasun1.unige.ch/pub/comp/win/amova/amova155.zip). The Euclidean distance described previously was computed using the package RAPDistance version 1.03 (Armstrong et al. 1995, ftp://life.anu.edu.au/pub/molecular-biology/software/rapd103.zip).

The following genetic distance was derived from each population pairwise F_{st} values obtained in the AMOVA analyses:

$d = -\ln(1 - F_{st})$.

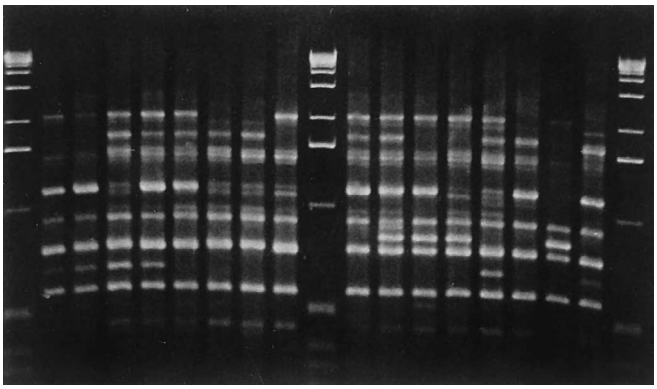

Fig. 2 Agarose gel depicting RAPD fragment polymorphisms seen within and between four accessions of *B. oleracea* var. *capitata* 'Golden Acre' when amplified with primer OPV-12. *Lanes 2–5* represent four individuals from accession 10, *lanes 6–9* accession 11, *lanes 11–14* accession 12, and *lanes 15–18* accession 14. *Lanes 1, 10,* and *19* are 1-kb size standards

This value represents the modified co-ancestry coefficient for each pair of populations, the corresponding distance matrix then was used to construct a dendrogram by the neighbor-joining method (Saitou and Nei 1987) using the program NEIGHBOR from the PHYLIP package (Felsenstein 1993). The significance of branching nodes was tested by randomization (100 samples) of the input order of accessions, and the subsequent construction of the consensus tree using the program CONSENSUS in the same package.

Results

On average, approximately 12 discrete DNA products were generated per primer, with a range of 8–22 bands for the nine selected primers (Table 2). Primer OPV-12 revealed the most polymorphism both within and among accessions (Fig. 2). A total of 110 markers was scored for the nine primers.

Measures of similarity and uniqueness were based on plant-to-plant comparisons (138 individuals, ranging from 6 to 12 individuals per accession based on intra-accession variation) (Table 3). The average, minimum, and maximum intra- and inter-accession distances for all pairwise comparisons have been summarized in Table 3. While individuals (eight) of the two reference accessions, 'Copenhagen Market' and 'Jersey Wakefield', grouped together uniquely in the corresponding neighbor-joining dendrogram (not shown), no accession of 'Golden Acre' was as well defined (individual plants aggregating as a distinct group). Moreover, pairwise analyses of 'Golden Acre' accessions showed that all were 'polyphyletic' in a populational sense (Avise 1993), i.e., for at least one individual 'X' from accession 'A' there was always at least one individual from a different accession that was closer to 'X' than any other individual from accession 'A'. An appropriate measure for this type of aggregation is provided by the index of classification, I_c (Estoup et al. 1995), and has been highlighted in Table 3(a). This index theoretically

Table 2 Characteristics of the nine selected oligonucleotide primers

Primer	Sequence	Total number of fragments scored	Number of polymorphic fragments	Size range of scored fragments (bp)
OPB-04	5′GGACTGGAGT	12	8	530–1400
OPB-10	5′CTGCTGGCAC	11	5	370–1130
OPB-19	5′ACCCCCGAAG	10	8	450–1660
OPB-20	5′GGACCCTTAC	15	13	440–1560
OPV-04	5′CCCCTCACGA	10	5	620–1570
OPV-12	5′ACCCCCCACT	22	16	370–1560
OPV-13	5′ACCCCCTGAA	8	8	740–1720
OPW-16	5′CAGCCTACCA	14	10	530–1420
OPY-08	5′AGGCAGAGCA	8	7	470–900
Total/range		110	80	370–1720

Table 3 (a) Average, minimum, and maximum distances for pairwise comparisons of *B. oleracea* individuals. Comparisons among individuals belonging to the same accessions are shown as intra-accession, and those among individuals from different accessions as inter-accession. The last column represents the index of classification (I_c) for the individuals belonging to each accession. (b) Summary of some relevant group and inter-accession comparisons averages. GA, JW, and CM denotes 'Golden Acre', 'Jersey Wakefield', and 'Copenhagen Market', respectively

(a)

Accession	Inter-accession			Intra-accession			I_c
	Average	Min	Max	Average	Min	Max	
GA-1	34.84	10	50	14.00	6	19	0.799
GA-2	36.29	4	55	14.98	1	30	0.922
GA-3	33.24	4	51	21.11	6	33	0.786
GA-4	35.49	9	51	21.82	3	35	0.830
GA-5	23.08	3	51	13.76	2	42	0.332
GA-6	22.95	4	46	9.79	5	15	0.825
GA-7	21.69	6	44	12.00	7	17	0.719
GA-9	22.24	4	46	8.44	3	15	0.974
GA-10	23.40	5	46	12.39	5	21	0.342
GA-11	21.81	3	45	9.96	3	16	0.503
GA-12	21.90	4	45	12.32	5	20	0.265
GA-14	23.08	7	50	14.00	6	20	0.277
GA-17	26.67	6	55	10.79	1	19	0.979
GA-18	24.26	6	51	13.71	4	20	0.901
JW	30.30	17	51	21.39	8	29	1.000
CM	25.41	9	46	8.50	3	15	1.000

(b)

Comparison	Average distance
Within 'Golden Acre' accessions	14.43
Within 'Golden Acre' accessions including 'Copenhagen Market'	14.12
Within accessions (all)	14.48
Among 'Golden Acre' accessions	26.59
Among 'Golden Acre' and 'Copenhagen Market'	25.03
Among 'Golden Acre' and 'Jersey Wakefield'	30.24
Among 'Golden Acre' plus 'Copenhagen Market' and 'Jersey Wakefield'	30.30
Among 'Copenhagen Market' and 'Jersey Wakefield'	31.13

varies from 1.0, representing a perfect grouping of all the individuals from one accession in one single group, to –0.5, which would represent the widest possible dispersion of individuals from a single accession or population. The two reference accessions, 'Copenhagen Market' and 'Jersey Wakefield', showed collective aggregation, while a range of values for I_c was observed among 'Golden Acre' accessions. Some of these entries (accessions 17, 9, 2, and 18) had high I_c values, whereas others (accessions 12, 14, 5, and 10) had much lower values, indicating that some of their individuals were quite scattered in the dendrogram among individuals belonging to different accessions.

Among the 'Golden Acre' holdings, accession 9 was the most homogeneous [average intra-accession Euclidean distance of 8.44, Table 3(a)], while accession 4 displayed the most variance in fragment patterns (average intra-accession distance of 21.82). The average pairwise distance estimate for 'Golden Acre' accessions was calculated to be 14.43. Of the two reference accessions, 'Copenhagen Market' exhibited an average dissimilarity of 8.50, one of the lowest among all entries in the study, while 'Jersey Wakefield' had one of the highest dissimilarities (21.39). The levels of variation detected within each accession may be attributed to the breeding system of cabbage (self-in-

Table 4 Summary of the results for three molecular analyses of variance (AMOVA) derived from RAPD analysis. (a) All 'Golden Acre' accessions as one group and 'Copenhagen Market' and 'Jersey Wakefield' as another. In this case a nested design was adopted with the following levels: among groups, among accessions within groups, and among individuals within accessions. (b) The second analysis was performed with the 14 'Golden Acre' accessions as one single group and the among-accessions and among-individuals-within-accessions components were estimated. (c) 'Golden Acre' accessions were assembled into four groups (accs. 1–4, accs. 5–9, accs. 10–14, and accs. 17–19) based on the results obtained by neighbor-joining clustering of the co-ancestry distances among pairs of accessions. The among-groups, among-accessions/within-groups, and within-accessions components of variation are shown. (*df*=degrees of freedom, SSD=sums of square differences, MSD=mean square differences, %Total=percentage of total variance, *P*-value=probability of obtaining a larger component estimate by chance alone)

Source of variation	*df*	SSD	MSD	Variance component	% Total	*P*-value
(a)						
Among groups	1	54.18	54.18	−0.22	−1.63	0.478
Accessions with groups	14	889.33	63.52	6.56	49.47	<0.001
Individuals within accessions	122	844.01	6.92	6.92	52.17	<0.001
(b)						
Among 'Golden Acre' accessions	13	817.14	62.86	6.45	48.53	<0.001
Individuals within 'Golden Acre'	108	739.39	6.85	6.85	51.47	
(c)						
Among groups	3	629.11	209.70	6.45	43.93	<0.001
Among accessions/Within groups	10	188.03	18.80	1.38	9.43	<0.001
Among accessions 1–4	3	87.03	29.01	2.19	19.21	<0.001
Among accessions 5–9	3	50.59	16.86	1.23	18.31	<0.001
Among accessions 10–14	3	35.84	11.95	0.73	10.75	<0.001
Among accessions 17–18	1	14.57	14.57	1.14	15.80	<0.001
Within accessions	108	739.39	6.85	6.85	46.64	<0.001

compatibility fostering outcrossing). Counterbalancing this was the strong selection pressure of plant breeding for uniform phenotype and the possibility of a genetic bottleneck incurred through limited population sizes during seed regeneration.

Results in Table 3(b) also include a summary of pairwise intra- and inter-accession average distances. The average distance between individuals belonging to 'Golden Acre' accessions, regardless of whether 'Copenhagen Market' (the putative progenitor) was or was not included, exhibited more variation which could be attributed to the intra-accession component as compared to the inter-accession value. To further test the significance of the contribution of the different levels of variation, a molecular analysis of variance (AMOVA) based on the same Euclidean distances as described previously was conducted. These results are summarized in Table 4.

Specifically, three different analyses of variance have been employed to establish further how variation was partitioned. The first analysis [Table 4, (a)] consisted of a nested design with two groups corresponding to: (1) all 'Golden Acre' accessions and (2) the reference cultivars, 'Copenhagen Market' and 'Jersey Wakefield'. There was no significant component of variation for differences among these two groups. However, there were significant differences among accessions within groups (49.47% of the total variance) and of individuals within accessions (52.17%). The negative value for the variance component corresponding to among-group differences can be explained by the disequilibrium in the number of accessions comprising both groups [Table 4(a)]. Subsequently an analysis was performed to estimate variation solely within individuals and among accessions of 'Golden Acre' [Table 4(b)]. The result showed a significant difference for among-accessions variation (48.53% of the total variation). This component of the total variance is slightly smaller than the within-accession component (51.47%). In other words, on average there is as much genetic variation within 'Golden Acre' accessions as there is among entries.

In order to further investigate the feasibility of grouping accessions of 'Golden Acre' for optimizing curatorial effort, four groups were established as follows. Based on the relationship between the variance components of each hierarchical level and the corresponding F-statistics (Wright 1951; Cockerham 1969, 1973), pairwise F_{st} values for all accessions were used to derive corresponding co-ancestry distances. The neighbor-joining clustering method then was used to construct a dendrogram grouping the different entries. The resulting dendrogram is presented as Fig. 3 and emphasizes that all 'Golden Acre' accessions could be aggregated as self-contained, distinctive groups. 'Golden Acre' accessions 1, 2, 3, and 4 constituted group 1, accessions 5, 6, 7, and 9 formed group 2, accessions 10, 11, 12, and 14 yield group 3, and lastly accessions 17 and 18 united as group 4. All of the branching points of the dendrogram defining these groups were supported by resampling as described previously.

The final analysis of molecular variance again represented a hierarchical design with all 'Golden Acre' accessions aggregated in the four previously defined subsets. The variation of the previous intra-accession cluster was then partitioned into the among-groups and the among-

accessions within-groups components [Table 4(c)]. The among-groups component accounted for 43.93% of the total variation, whereas the among-accessions within-groups represented only 9.43% of the total variation. The within-accessions component still represented 46.64% of the total variation. It may be noted that the four groups of 'Golden Acre' accessions contained similar levels of variation, ranging from 10.75 to 19.21% of the total. Most importantly, when comparing sources of variation associated with the groupings of 'Golden Acre' accessions [Table 4(b) and (c)] the reduction in accessions (from 14 to 4 groups) would constitute only a 4.6% reduction in absolute variation (48.53–43.93%).

Discussion

The RAPD assay provided a simple means to quickly compare genetic relationships among all 'Golden Acre' accessions (Fig. 3). As one might expect based on pedigree, passport, and descriptor information, all accessions of 'Golden Acre' were closely related. The putative progenitor of the 'Golden Acre' accessions, 'Copenhagen Market', was also embedded among entries (Fig. 3). However, the amount of variation exhibited by each accession varied significantly (Table 3). Following a partitioning of variation employing AMOVA, a clearer picture emerged regarding the specific relationships among individuals in accessions and it was established that more variation was present within the accessions than between entries (Table 4). Moreover, findings suggested that only certain accessions, or groups of accessions, might warrant individual maintenance. However, other entries clearly were genetically similar and phenotypically uniform, indicating that they may be placed on a lower priority for regeneration, bulked, or discarded. Armed with all possible sources of pertinent information, curators ultimately must be able to make difficult decisions on collection holdings. Finite resources, dedicated to long-term maintenance of genetic resources, warrant this action.

This type of molecular analysis, when integrated with pedigree, passport, and descriptor data, provides a basis for a curator to retain the broadest range of variation observed, while recognizing meaningful genetic differences among the entries. For example, aggregation of genetically similar accessions, as noted previously, may significantly reduce maintenance (particularly regeneration and processing) costs of the collection, thereby allowing for the planned addition of more valuable and unique cultivars or landraces. From the perspective of effective use of limited financial support, reducing the 'Golden Acre' holdings from 14 to 4 would reduce costs associated with production and processing by approximately 70% [$750–1000 (US) per accession in this specific case] per regeneration cycle (20–25 years). Though these accessions of 'Golden Acre' were selected for a particular character or phenotype, the genes or gene complexes conferring these traits could be conserved and easily made accessible in grouped samples. Data based on the partitioning of variation among accessions (Table 4) indicated that the reduction to four groups would only incur a potential loss of absolute variation of approximately 4.6%. It is critical to note that this level of aggregation effected only a relatively small loss of variation because the relationship between grouping and variation is asymptotic.

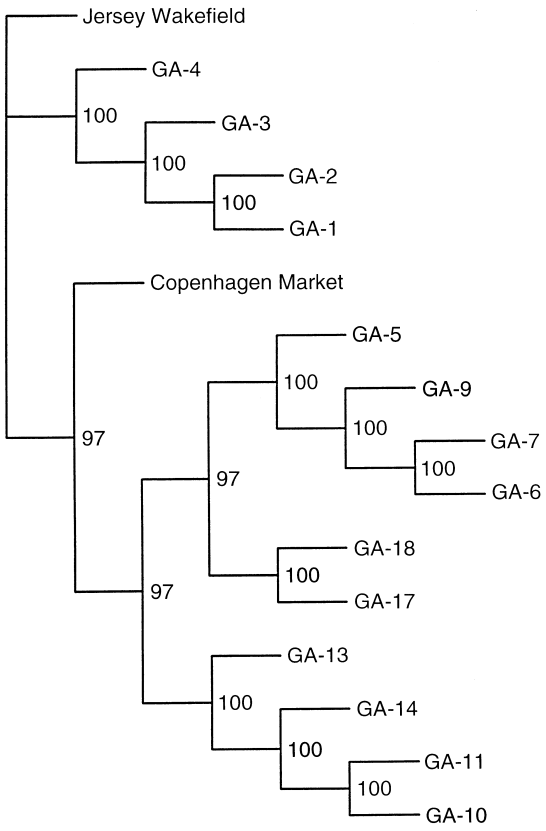

Fig. 3 Dendrogram obtained by the neighbor-joining method from the co-ancestry distances among pairs of *B. oleracea* accessions. The numbers in the branching nodes represent the number of times that the corresponding node is supported in 100 tree reconstructions with random and different orders of input of the accessions in the neighbor-joining algorithm. *GA*, *JW*, and *CM* denote 'Golden Acre', 'Jersey Wakefield', and 'Copenhagen Market', respectively

The loss of genetic variation due to drift will eventually deplete allelic variants from every accession, although this may be offset by mutation and/or migration. If one, single finite population (accession) is subdivided into several, smaller subpopulations with subsequent regeneration as separate units, the asymptotic loss of all allelic variation may be reduced so that genetic differences arising in the subdivision process (the among-accession component of variation) will be retained as fixed differences in the separate accessions. The converse is also true. When several small, independent subpopulations (accessions) are grouped into a single breeding population, initial among-population variation will eventually be lost by drift. In the present investigation, a method of analysis has been devel-

oped that can be employed to minimize the asymptotic loss of genetic variation by conversion of the among-accessions/within-groups component of variation into the within-groups component following the proposed aggregation, while simultaneously reducing the costs of maintenance of the collection.

The rate of loss of variation is, nevertheless, inversely proportional to the effective population size. However, for a given population size, assumed constant and equal to N individuals, the effective population size depends on the breeding structure, sex-ratio, and other factors. Among these, the variance in the contribution of individuals from one generation to their offspring (V_k) relates to the effective size (Hartl and Clark 1989) such that:

$N_e = (4N–2)/(V_k+2)$.

Therefore, a regeneration strategy and protocol that minimizes V_k (which eventually will equal zero if all individuals contribute exactly the same to the next generation) will maximize the effective population size (up to $N_e=2N–1$), and concurrently minimize the loss of genetic variation due to drift. Molecular techniques, such as the RAPD assay, are presently available for potentially determining the genotype of any individual in a finite population, and could be used as practical tools to achieve these goals.

In comparision with other molecular diagnostic methods, the RAPD assay has the advantages of low start-up and operating costs. Along with these benefits, the RAPD assay may be amenable to automation (Sederoff et al. 1992; Rafalski and Tingey 1993). With several hundred primers available commercially, the chances of identifying unique RAPD fragments to discriminate among closely related genotypes are quite high. Also, it is simpler and quicker to generate data by the RAPD assay than by many other molecular techniques. Conversely, issues regarding the reproducibility of the assay warrant concern and some error may be incurred when lab-to-lab comparisons are made. Nonetheless, this method may prove useful to curators for preliminary screening of the variation to be used for a more thorough characterization and maintenance of genetic resources, and to plant breeders for an assessment of seed purity. However, because of concerns with reliability it is unlikely that RAPD analysis will be favored for plant variety protection.

The large-scale assessment of genetic diversity in populations (whether in situ or ex situ) would be enhanced greatly with ready access to a sensitive, simple, timely, and cost-effective molecular technique such as the RAPD assay (Mailer et al. 1994; Margale et al. 1996). Undoubtedly, molecular screening techniques will increasingly be used to support curatorial and conservationist responsibilities. Such responsibilities include establishing priorities for conservation actions, and designing and monitoring management effectiveness and efficiency. Through the ready integration of molecular techniques with current practices, curators will be able to enhance both the accessibility and the value of the plant genetic resources which they are charged to conserve.

Acknowledgements The authors thank D. Astley for supplying selected accessions of ‘Golden Acre’. We also acknowledge R. E. Dean, R. R. Duncan, and A. L. Westman for constructive criticism of the manuscript. This investigation served as partial fulfillment of the requirements for a MS degree in Plant Breeding and Genetics at Cornell University. F.G.C. was supported through short-term fellowships from Ministerio de Educación y Ciencia and Generalitat Valenciana, Spain, and by project PB93-0350 from DGICYT.

References

Armstrong J, Gibbs A, Peakall R, Weiller G (1995) RAPDistance, package manual. Version 1.03. Australian National University, Canberra

Avise JC (1993) Molecular markers, natural history and evolution. Chapman-Hall, New York, pp 126–138

Beyermann B, Nurnberg P, Weihe A, Meixner M, Epplen JT, Borner T (1992) Fingerprinting plant genomes with oligonucleotide probes specific for simple repetitive DNA sequences. Theor Appl Genet 83:691–694

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

Cockerham CC (1969) Variance of gene frequencies. Evolution 23: 72–84

Cockerham CC (1973) Analysis of gene frequencies. Genetics 74: 679–700

Colosi JC, Schaal BA (1993) Tissue grinding with ball bearings and vortex mixer for DNA extraction. Nucleic Acids Res 21:1051–1052

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

Epplen JT, Ammer H, Epplen C, Kammerbauer C, Mitreiter R, Roewer L, Schwaiger W, Steimle V, Zischler H, Albert E, Andreas A, Beyermann B, Meyer W, Buitkamp J, Nanda I, Schmid M, Nurnberg P, Pena SDJ, Poche H, Sprecher W, Schartl M, Weising K, Yassouridis A (1991) Oligonucleotide fingerprinting using simple repeat motifs: a convenient, ubiquitously applicable method to detect hypervariability for multiple purposes. In: Burke T et al. (eds), DNA fingerprinting; approaches and applications. Birkhäuser Verlag, Berlin, Germany, pp 50–69

Estoup A, Garnery L, Solignac M, Cornuet JM (1995) Microsatellite variation in honey bee (*Apis mellifera* L.) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. Genetics 140:679–695

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Felsenstein J (1993) PHYLIP (phylogeny inference package). Version 3.5. University of Washington, Seattle

Hartl DL, Clark AG (1989) Principles of population genetics. 2nd edn. Sinauer, Sunderland

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing buffalo grass [*Buchloe dactyloides* (Nutt.) Engelm.]. Theor Appl Genet 86:927–934

Kaemmer D, Afza R, Weising K, Kahl G, Novak FJ (1992) DNA amplification fingerprinting of wild species and cultivars of banana (*Musa* spp). Bio/Technology 10:1030–1035

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, Lamboy WF, Szewc-McFadden AK, McFerson JR, Forsline PL (1993) Molecular diagnostics and plant genetic resources conservation. AgBiotech News Info 5:255–258

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

Margale E, Herve Y, Hu J, Quiros CF (1995) Determination of genetic-variability by RAPD markers in cauliflower, cabbage and kale local cultivars from France. Genetic Resources and Crop Evolution SEP, V42, 3:281–289

Margale E, Herve Y, Hu J, Quiros CF (1996) Characterization by RAPD markers of a local cole crop cultivar collection from France: optimal sample size and number of markers. Euphytica (in press)

Nienhuis J, Slocum MK, DeVos DA, Muren R (1993) Genetic similarity among *Brassica oleracea* L. genotypes as measured by restriction fragment length polymorphisms. J Am Hort Sci 118:298–303

Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genet 9:275–279

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sederoff R, Grattapalia D, Wilcox P, Chaparro J, O'Malley D, McCord S, Whetten R, McIntyre L, Weir BS (1992) Use of PCR-based RAPD markers for genetic mapping in conifers. In: Abstracts, 203rd ACS National Meeting, San Francisco, California, BTEC73

Soller M, Beckmann JS (1983) Genetic polymorphism in varietal identification and genetic improvement. Theor Appl Genet 67: 25–33

Virk PS, Newbury HJ, Jackson MT, Ford-Lloyd BV (1995) The identification of duplicate accessions within a rice germplasm collection using RAPD analysis. Theor Appl Genet 90: 1049–1055

Weising K, Ramser J, Kaemmer D, Kahl G, Epplen JT (1991) Oligonucleotide fingerprinting in plants and fungi. In: Burke T et al. (eds), DNA fingerprinting; approaches and applications. Birkhäuser Verlag, Berlin, Germany, pp 313–329

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Wright S (1951) The genetical structure of populations. Ann Eugen 1:323–334