12-2006

Analysis of Bulked and Redundant Accessions of Brassica Germplasm Using Assignment Tests of Microsatellite Markers

Von Mark V. Cruz
Iowa State University

John D. Nason
Iowa State University, jnason@iastate.edu

Richard Luhman
Iowa State University

Laura F. Marek
Iowa State University, lmarek@iastate.edu

Randy C. Shoemaker
United States Department of Agriculture
See next page for additional authors

Analysis of Bulked and Redundant Accessions of Brassica Germplasm Using Assignment Tests of Microsatellite Markers

Abstract
This study was conducted to determine if *Brassica* germplasm bulks created and maintained by the USDA-ARS North Central Plant Introduction Station (NCRPIS) were made with genetically indistinguishable component accessions and to examine newly identified putative duplicate accessions to determine if they can be bulked. Using ten microsatellite primer pairs, we genotyped two bulks of *B. rapa* L. ssp. *dichotoma* (Roxb.) Hanelt comprising four accessions and three bulks of *B. rapa* L. ssp. *trilocularis* (Roxb.) Hanelt comprising fourteen accessions, as well as four pairs of putatively duplicate accessions of *B. napus* L. Assignment tests on ten individual plants per accession were conducted using a model-based clustering method to arrive at probabilities of likelihood of accession assignment. The assignment tests indicated that one of the two bulks of *B. rapa* ssp. *dichotoma* involves genetically heterogeneous accessions. It was observed in the *B. rapa*ssp. *trilocularis* bulks that the component accessions could be differentiated into groups, with misassignments observed most frequent within groups. In *B. napus*, only one of the four pairs of putative duplicates showed significant genetic differentiation. The other three pairs of putative duplicates lack differences and support the creation of bulks. The results of the assignment tests were in agreement with cluster analyses and tests of population differentiation. Implications of these results in terms of germplasm management include the maintenance and/or re-creation of some*Brassica* germplasm bulks by excluding those accessions identified as being unique in this study.

Keywords
Brassica, Duplicate, Individual assignment, Genetic resources, SSRs, Rapeseed, Rationalization

Comments
This article is from *Euphytica* 152 (2006): 339, doi:10.1007/s10681-006-9221-5.

Authors
Von Mark V. Cruz, John D. Nason, Richard Luhman, Laura F. Marek, Randy C. Shoemaker, E. Charles Brummer, and Candice A. C. Gardner

Euphytica (2006) 152:339–349
DOI 10.1007/s10681-006-9221-5

Analysis of bulked and redundant accessions of *Brassica* germplasm using assignment tests of microsatellite markers

Von Mark V. Cruz · John D. Nason · Richard Luhman · Laura F. Marek · Randy C. Shoemaker · E. Charles Brummer · Candice A. C. Gardner

Received: 23 March 2006 / Accepted: 16 June 2006 / Published online: 30 August 2006

Abstract This study was conducted to determine if *Brassica* germplasm bulks created and maintained by the USDA-ARS North Central Plant Introduction Station (NCRPIS) were made with genetically indistinguishable component accessions and to examine newly identified putative duplicate accessions to determine if they can be bulked. Using ten microsatellite primer pairs, we genotyped two bulks of *B. rapa* L. ssp. *dichotoma* (Roxb.) Hanelt comprising four accessions and three bulks of *B. rapa* L. ssp. *trilocularis* (Roxb.) Hanelt comprising fourteen accessions, as well as four pairs of putatively duplicate accessions of *B. napus* L. Assignment tests on ten individual plants per accession were conducted using a model-based clustering method to arrive at probabilities of likelihood of accession assignment. The assignment tests indicated that one of the two bulks of *B. rapa* ssp. *dichotoma* involves genetically heterogeneous accessions. It was observed in the *B. rapa* ssp. *trilocularis* bulks that the component accessions could be differentiated into groups, with misassignments observed most frequent within groups. In *B. napus*, only one of the four pairs of putative duplicates showed significant genetic differentiation. The other three pairs of putative duplicates lack differences and support the creation of bulks. The results of the assignment tests were in agreement with cluster analyses and tests of population differentiation. Implications of these results in terms of germplasm management include the maintenance and/or re-creation of some *Brassica* germplasm bulks by excluding those accessions identified as being unique in this study.

Keywords. *Brassica* · Duplicate · Individual assignment · Genetic resources · SSRs · Rapeseed · Rationalization

V. M. V. Cruz · R. Luhman · L. F. Marek · R. C. Shoemaker · E. C. Brummer · C. A. C. Gardner (✉)
Department of Agronomy, Iowa State University, Ames, IA 50011, USA
e-mail: gardnerc@iastate.edu

J. D. Nason
Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011, USA

R. C. Shoemaker
USDA-ARS Corn Insect and Crop Genetics Research Unit, Ames, IA 50011, USA

C. A. C. Gardner
USDA-ARS Plant Introduction Research Unit and North Central Regional Plant Introduction Station, Ames, IA 50011, USA

Introduction

Brassica napus (rape) and *B. rapa* (turnip rape) are important crops which have been traditionally

grown for animal fodder and as sources of inedible and edible oils (McNaughton 1995). At present, they are also used to produce biodiesel, as well as rubber and plastics (Kimber and McGregor 1995; Houmiel et al. 1999). *Brassica rapa* ($2n = 20$, AA) is believed to be one of the diploid progenitor species of the amphidiploid *B. napus* ($2n = 38$, AACC). Germplasm collections of oilseed *Brassica* in the U.S. are maintained at the USDA-ARS North Central Plant Introduction Station (NCRPIS), Ames, IA. There are currently more than 2000 accessions of oilseed *Brassica* at the NCRPIS, with the bulk comprised of *B. napus* and *B. rapa*. As part of effective management of these collections, probable duplicate accessions are routinely identified by the curator, both through the examination of passport data and via field characterization. Rationalization of germplasm collections by reducing the number of duplicate accessions is important in order to alleviate constraints on storage space and financial resources, and to simplify germplasm regeneration procedures in general (Hintum and Visser 1995; Treuren and Hintum 2003). For *Brassica*, regeneration costs are relatively high because many forms are biennial, requiring vernalization, and because many accessions also require pollination control in order to preserve the original genetic profile due to high levels of outcrossing. *Brassica rapa*, in particular, is an obligate outcrosser with genetic self-incompatibility (Eastham and Sweet 2002). Similarly, *B. napus*, though often treated as autogamous, can outcross at a rate of 20–30% (Khachatourians et al. 2001; Rakow and Woods 1987)

Bulking of *Brassica* germplasm was done by NCRPIS in 1999 and 2001. Numerous accessions of *B. rapa* in the collection received from a single source were bulked because of morphological and/or phenological similarity and a lack of definitive passport data. The morphological characterization and final bulking were performed with the help of a crop breeder. The *Brassica* germplasm bulks were formed by mixing equal portions of pure live seed from each original accession. As a result, eleven new accessions of *B. rapa* ssp. *dichotoma* and six new accessions of *B. rapa* ssp. *trilocularis* were created. Each bulk has two to sixteen component accessions. The original component accessions of these *B. rapa* bulks are available; their use allowed us to test the hypothesis of accession homogeneity and determine whether the original bulks were created with genetically similar accessions.

In a similar case, apparent duplicate pairs of *B. napus* accessions were evaluated to test for similarity using molecular characterizations. These recently identified pairs have been considered for bulking because of morphological similarity, identical varietal names and origins, and consecutive genebank numbers. Similar approaches to rationalize germplasm collections by identifying and reducing the number of probable duplicates in genebanks have been conducted in other species such as *B. oleracea* L. (Hintum et al. 1996) and *Linum usitatissimum* L. (Treuren et al. 2001).

Ten microsatellite or simple sequence repeat (SSR) marker loci were used in the molecular characterization of the *Brassica* accessions. Many classes of molecular markers in *Brassica* are available, but SSRs have gained popularity because of cost effectiveness, speed, reproducibility, and especially polymorphism (Lund et al. 2003; Phippen et al. 1997; Snowdon and Friedt 2004). Analysis of SSR data was conducted by using an assignment test to answer the question of accession homogeneity. This test uses the multilocus genotypes of representative individuals from each accession and determines if fixed differences between accessions exist. The method was first implemented by Petkau et al. (1995) and has been used successfully in population and conservation biology studies to assign individuals to specific source populations with as few as seven polymorphic marker loci (Primmer et al. 2000). The details of this method have been extensively reviewed (Cornuet and Luikart 1996; Davies et al. 1999; Pritchard et al. 2000; Waser and Strobeck 1998). Here, we introduce assignment tests as a tool in plant genetic resource management to assist curatorial decisions by enabling the coarse classification of, and discrimination among, putatively duplicate accessions in *B. napus* and validate whether old bulks in *B. rapa* are composed of genetically similar accessions.

Methodology

Plant materials

Four accessions of *B. rapa* ssp. *dichotoma*, fourteen accessions of *B. rapa* ssp. *trilocularis*, and eight accessions of *B. napus* (Table 1) were selected for analysis. The *B. rapa* accessions comprise the original component accessions of old bulks whereas the *B. napus* accessions represent four pairs of putative duplicates selected from several pairs of accessions with consecutive accession numbers, identical varietal names, and similar origins as indicated in the passport data. Ten to fifteen plants per accession were grown in the greenhouse using Rootrainer™ trays (Spencer-Lemaire Ind. Ltd., Edmonton, AB) with Fafard® Canadian Growing Mix no. 2 (Conrad Fafard, Inc., Agawam, MA).

Tissue collection and DNA extraction

Leaf tissue was harvested when the plants were at the 3–4 leaf stage. Ten plants per accession were sampled. DNA was obtained from each plant using FTA® cards (Whatman Inc., Sanford, ME).

Table 1 List of selected component accessions of five bulks in *B. rapa* ssp. *dichotoma* and *B. rapa* ssp. *trilocularis* and putative duplicates in *B. napus*

GH ID*	Accession	Variety	Source**
B. rapa ssp. *dichotoma*			
	Ames 26162 bulk		
14	Ames 9288	–	United States
15	Ames 9291	–	United States
	PI 633167 bulk		
16	Ames 9410	–	United States
17	Ames 9414	–	United States
B. rapa ssp. *trilocularis*			
	Ames 26168 bulk		
18	Ames 9698	–	United States
19	Ames 9699	–	United States
20	Ames 9701	–	United States
21	Ames 9702	–	United States
22	Ames 9718	–	United States
23	Ames 9719	–	United States
	Ames 26170 bulk		
24	Ames 9889	–	United States
25	Ames 9895	–	United States
26	Ames 9900	–	United States
	Ames 26171 bulk		
27	Ames 9893	–	United States
28	Ames 9898	–	United States
29	Ames 9919	–	United States
30	Ames 9920	–	United States
31	Ames 9925	–	United States
B. napus			
5	PI 469724	Aomori	South Korea
6	PI 469725	Aomori	South Korea
7	PI 469789	Fonto	South Korea
8	PI 469790	Fonto	South Korea
9	PI 469808	Gokstad	South Korea
10	PI 469809	Gokstad	South Korea
11	PI 470046	Titus	South Korea
12	PI 470047	Titus	South Korea

* Greenhouse flat number

** Source in USDA-ARS GRIN database; original *B. rapa* components from India

Leaf extracts were collected by moderately pounding the leaves on the sample area of the cards. The cards were then allowed to dry and were stored at 24°C in a dessicator cabinet containing Drierite (W.A. Hammond Drierite Co. Ltd., Xenia, OH). To elute the DNA from the FTA® cards, 3 mm diameter discs were punched out of the sample area. Each disc was put in a 0.5 ml micro-centrifuge tube and washed with 200 µl TE (10 mM Tris, pH 8.0, 0.1 mM EDTA). After washing, the paper discs were incubated for 10–15 min at room temperature with 20 µl TE, and the eluted DNA used in PCR.

PCR amplification of microsatellites

PCR was performed in 96-well Microseal™ polypropylene microplates (Bio-Rad Lab. Inc., Hercules, CA). Ten SSR primer pairs from BrassicaDB (http://brassica.bbsrc.ac.uk/BrassicaDB) were used in this study (Table 2). Selection of the primers was based on presence of polymorphism observed in a previous study of *Brassica* germplasm (Cruz 2006). The reaction mix was composed of 0.1 µl of eluted DNA, 9 µl of sterile ddH2O, 1 µl 10× PCR buffer (200 mM Tris–HCl, pH 8.4, 500 mM KCl), 0.5 µl dNTPs (2 mM), 0.3 µl MgCl (50 mM), 0.1 µl of primers (50 mM), and 0.05 µl *Taq* polymerase (5 U/µl) (Invitrogen Corp., Carslbad, CA). Thermal cycling was done using DNA Engine® (PTC-200™) thermal cyclers (Bio-Rad Lab. Inc., Hercules, CA) with the following conditions: 94°C for 2 min, then followed by 35 cycles of amplification at 94°C for 30 s, 55°C for 30 s, 72°C for 30 s, and final extension at 72°C for 4 min. The PCR products were separated in a 4.0% agarose gel prepared in 1× TAE (40 mM Tris–acetate, 1 mM EDTA, pH 8.3) with incorporated ethidium bromide (0.46 µg/ml). Visualization of amplified products was done with a UV light box and photographed on a digital gel-documentation system.

Data analysis

Assignment tests involving putative duplicate and previously bulked accessions were conducted with Structure v.2.0 (Pritchard et al. 2000), a software package that uses a model-based clustering method and Bayesian approach to determine probabilities of assignment for samples from individual accessions. The clustering method employed by this software does not assume a particular mutation process. In conducting the simulations, each plant was tagged with its original accession number to serve as prior population information in making inferences on re-assignment. After conducting trial simulations, final program settings were used with a 30,000 iteration burn-in period and 10^5 Monte Carlo Markov Chain iterations for data gathering. The null hypothesis is that the putative duplicate (or previously bulked) accessions are genetically distinct, in which case Structure should correctly group plants from the same accession. Alternatively, if putative duplicate (or previously bulked) accessions are homogeneous, Structure will not be successful in grouping individual plants from the

Table 2 List of the ten microsatellites analyzed, band sizes observed and polymorphism information content (PIC)

SSR Name	Repeat	Primer sequence (forward, reverse)	Size range observed (bp)	PIC
Na12-A02	$(CT)_{16}$	AGCCTTGTTGCTTTTCAACG, AGTGAATCGATGATCTCGCC	150–200	0.452
Na12-A08	$(GA)_{28}$	AACACTTGCAACTTCATTTTCC, CATTGGTTGGTGAATTGACAG	150–320	0.622
Na12-C08	$(CT)_{50}$	GCAAACGATTTGTTTACCCG, CGTGTAGGGTGATCTAGATGGG	275–350	0.537
Na14-C12	$(AG)_{17}$	CACATTTTGGTTCAATTCGG, TACGACGCTGGTTTCGATTC	190–200	0.340
Na14-D07	$(CCG)_3$	GCATAACGTCAGCGTCAAAC, CTGCGGGACACATAACTTTG	150–175	0.419
Ni4-D09	$(CT)_{25}$	AAAGGACAAAGAGGAAGGGC, TTGAAATCAAATGAGAGTGACG	170–200	0.555
Ol11-H02	$(AAC)_{18}$	TCTTCAGGGTTTCCAACGAC, AGGCTCCTTCATTTGATCCC	180–210	0.473
Ra2-E03	$(CT)_{18}$	AGGTAGGCCCATCTCTCTCC, CCAAAACTTGCTCAAAACCC	225–315	0.509
Ra2-E07	$(GA)_{19}$	ATTGCTGAGATTGGCTCAGG, CCTACACTTGCGATCTTCACC	100–170	0.677
Ra2-F11	$(CT)_{34}$	TGAAACTAGGGTTTCCAGCC, CTTCACCATGGTTTTGTCCC	190–300	0.793

same accession. For diploid *B. rapa*, we also computed the probability of population differentiation between pairs of bulked accessions using GENEPOP v.3.4 (Raymond and Rousset 1995). An experiment-wide error rate of $P < 0.05$ was maintained and the significance of population differentiation values determined using critical values obtained by sequential Bonferonni adjustment (Rice 1989). Overall relationships among accessions were visualized by using cluster analysis of genetic distances. Nei's (1972) genetic distance, which assumes a model of mutation and drift for populations of constant effective size, was computed for pairs of accessions using GenAlEx v.6 (Peakall and Smouse 2006). Clustering of Nei's distances was performed by using the UPGMA-algorithm in NTSYS-pc v.2.2 (Rohlf 2005).

Results

All ten microsatellite primer pairs amplified bands in the *Brassica* accessions. However, species-specific amplification was observed with the Ra2-E07 and Na14-C12 primer pairs: no amplification product was obtained with Ra2-E07 in *B. napus*, with Na12-C08 in *B. rapa* ssp. *dichotoma*, or with Na14-C12 in *B. rapa* ssp. *trilocularis* (Table 2). The polymorphism information content (PIC) value of the ten microsatellite loci ranged from 0.340 (Na14-C12) to 0.793 (Ra2-F11), with an average of 0.538. An average of three bands was amplified per primer pair over all accessions in the species analyzed. Posterior probabilities derived from assignment tests on the ten plants of each accession are shown in Figs. 1 to 5 together with dendrograms indicating relationships among the component accessions of bulks in *B. rapa* and the putative duplicates in *B. napus*. Specific information indicating correct assignments as well as misclassifications of plants in each accession is discussed below.

Component accessions of bulked *B. rapa* ssp. *dichotoma*

Results obtained from samples of the two pairs of original accessions that comprised the two bulks of *B. rapa* ssp. *dichotoma* are shown in Fig. 1. The UPGMA dendrogram distinguished between the two bulks and grouped their component accessions

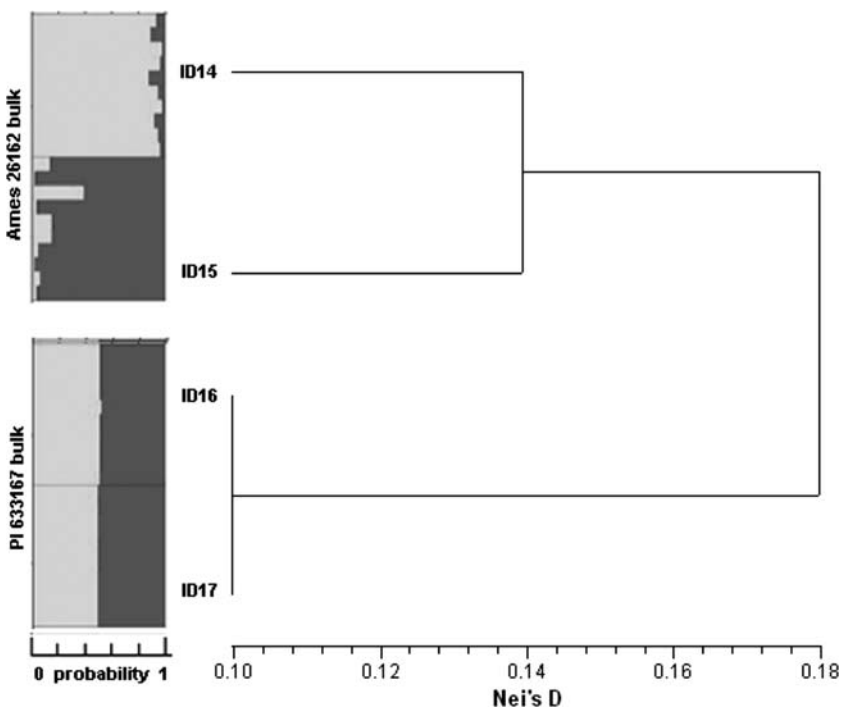

Fig. 1 Relationships among the component accessions of two bulks of the Ames 26162 bulk and the PI 633167 bulks of *B. rapa* ssp. *dichotoma*, and the probabilities of assignment to accessions within bulks. Pairwise comparison of differentiation between component accessions was significant between ID14 and ID15 ($P = 0.002$) and between ID16 and ID17 ($P = 0.004$)

according to their bulking designation. In the PI 633167 bulk, component accessions had nearly equal probabilities of assignment and the genetic distance between them was small. In contrast, in the Ames 26162 bulk, a marked difference was observed between its member accessions, with accession specific probabilities of assignment generally greater than 0.8. Despite this difference between the two bulks, no plant was misclassified for either pair of the component accessions and tests of population differentiation indicated that component accessions in both bulks are genetically distinct ($P < 0.05$).

Component accessions of bulked *B. rapa* ssp. *trilocularis*

The Ames 26168 bulk could be differentiated into groups of four (Ames 9699 [ID19], Ames 9702 [ID21], Ames 9718 [ID22], and Ames 9719 [ID23]) and two (Ames 9698 [ID18] and Ames 9701 [ID20]) component accessions (Fig. 2). Also, a few of the plants in the latter two accessions had low probabilities of assignment to their source, suggesting the presence of admixing among the component accessions. Pairwise tests of population differentiation were significant for comparisons between the following component accessions: both ID18 and ID20 versus ID19, ID21, ID22, and ID23, and ID23 versus both ID19 and ID22 ($P < 0.01$ in all comparisons). Misassignments were most frequent for ID21, ID22, and ID23. with almost half of the plants in each of these accessions misassigned.

The Ames 26170 bulk had probabilities of assignment that were almost equally distributed among the three component accessions with random misassignments observed in all three members of the bulk (Fig. 3). Despite this similarity, the UPGMA tree of Nei's distances

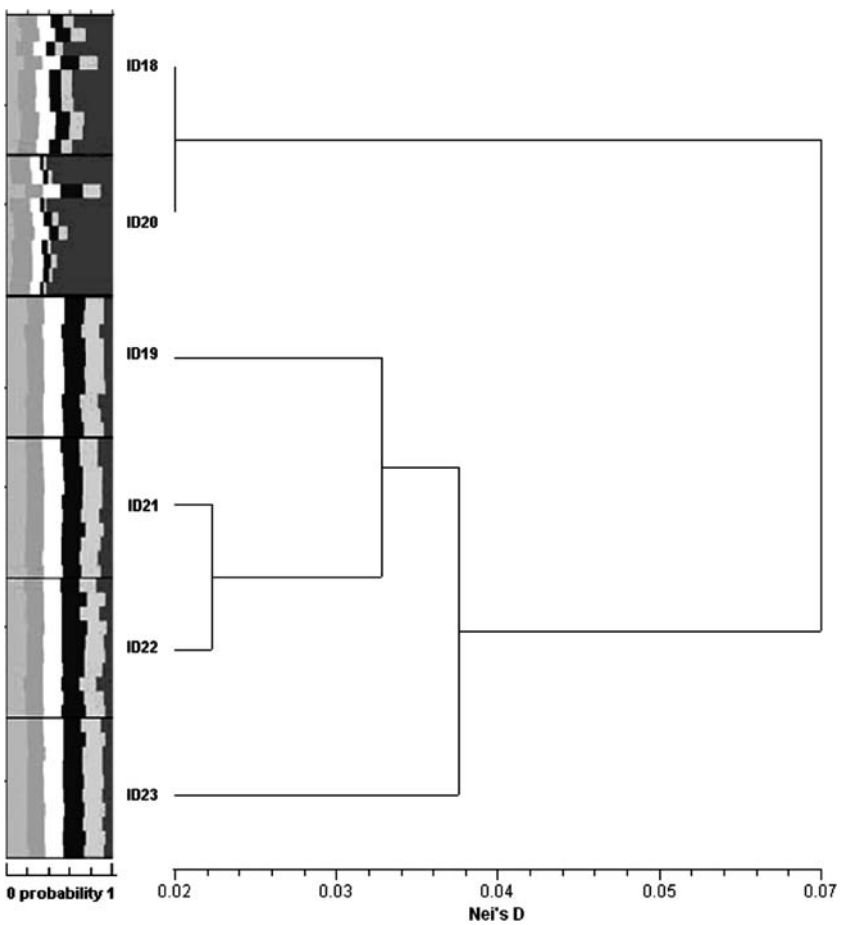

Fig. 2 Relationships among the component accessions of Ames 26168 bulk (*B. rapa* ssp. *trilocularis*) and the probabilities of assignment of individual samples to each accession. Colors portrayed (L to R) correspond to probabilities of assignment to accessions ID23, ID22, ID21, ID20, ID19, and ID18, respectively. Pairwise comparison of differentiation between component accessions were non-significant in ID18 and ID20 ($P = 0.173$), ID19 and 22 ($P = 0.058$), ID19 and 21 ($P = 0.737$), ID21 and ID22 ($P = 0.672$), and ID21 and ID23 ($P = 0.185$); all other comparisons have significant *P*-values ($P < 0.01$)

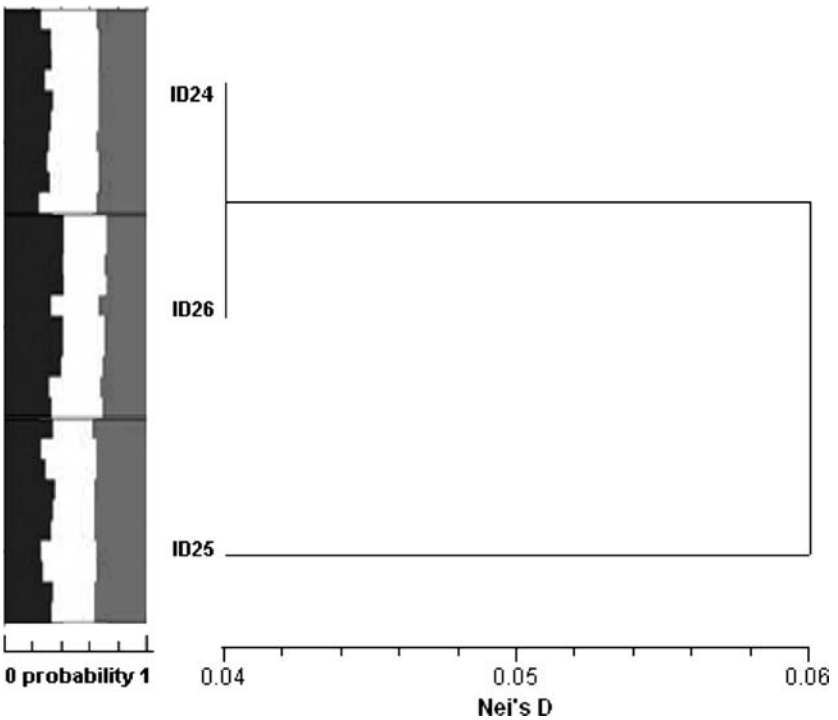

Fig. 3 Relationships among the component accessions of Ames 26170 bulk (*B. rapa* ssp. *trilocularis*) and the probabilities of assignment of the samples within each accession. Colors (L to R) correspond to probabilities of assignment to accessions ID26, ID25, and ID24, respectively. Pairwise comparison of differentiation between component accessions was non-significant between ID24 and ID26 ($P = 0.543$) and highly significant for all other comparisons ($P < 0.016$)

grouped accessions ID24 and ID26 relative to ID25, and pairwise tests of population differentiation were significant for ID25 versus ID24 and ID26 ($P < 0.016$ in both cases).

In the bulk accession Ames 26171, two distinct groups were found: the first comprising Ames 9893 [ID27] and Ames 9898 [ID28], and the second Ames 9919 [ID29] Ames 9920 [ID30], and Ames 9925 [ID31] (Fig. 4). Two plants were misassigned between ID27 and ID28, and three plants between ID29 and ID30. No misassignment was observed in ID31. The component accessions ID29 and ID30 were closely grouped in the UPGMA tree, had similar probabilities of assignment and did not exhibit significant population differentiation ($P = 0.065$). The remaining three component accessions (Ames 9893 [ID27], Ames 9898 [ID28], and Ames 9925 [ID31]), in contrast, exhibited greater genetic distances, heterogeneous probabilities of assignment, and significant pairwise differentiation from each other and from ID29 and ID30 ($P < 0.01$ in all cases). Whereas misassignments were observed in ID27, ID28, ID29 and ID30, the high probabilities of individual samples being re-assigned back into ID31 and the absence of misclassified plants suggest that this accession in particular was different from the rest of the bulk components.

Probable duplicates in *B. napus*

The UPGMA tree of Nei's distances grouped accessions consistent with their variety names (Fig. 5). Among the four pairs of *B. napus* accessions, only the probable duplicates labeled as Aomori (PI 469724 [ID5] and PI 469725 [ID6]) showed evidence of genetic differentiation. Analysis conducted using Structure indicates that plants from these two accessions could be reassigned back to their original source accessions with probabilities of 0.6–0.8. Moreover, Structure did not misclassify a single plant between the two Aomori accessions.

In the case of the other putative duplicates—Fonto (PI 469789 [ID7] and PI 469790 [ID8]), Gokstad (PI 469808 [ID9] and PI 469809 [ID10]), and Titus (PI 470046 [ID11] and PI 470047 [ID12])—there was an average of 50% probability of assignment of samples to either accession in a pair (Fig. 5). Also, several individual plants were misclassified in these pairs: two in Fonto, four in Gokstad, and four in Titus. The

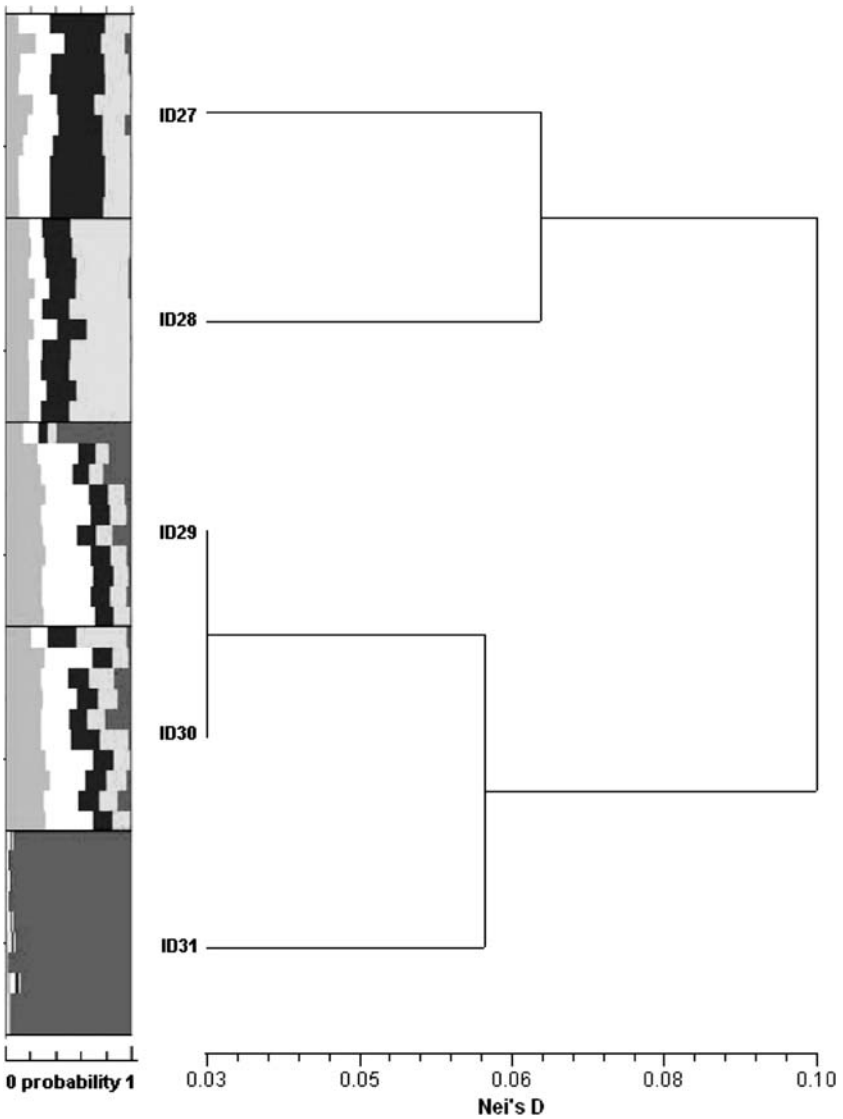

Fig. 4 Relationships among the component accessions of Ames 26171 bulk (*B. rapa* ssp. *trilocularis*) and the probabilities of assignment of the samples within each accession. Colors (L to R) correspond to probabilities of assignment to accessions ID31, ID30, ID29, ID28, and ID27, respectively. Pairwise comparison of differentiation between component accessions was non-significant between ID29 and ID30 ($P = 0.065$) and highly significant for all other comparisons ($P < 0.01$)

nearly equal probabilities of assignment, frequent misassignment, and the small genetic distance between these three accession pairs suggest that they are more likely candidates for bulking than are the two Aomori accessions.

Discussion

Microsatellites have been demonstrated to be a good source of additional information for curators to clarify relationships among phenotypically similar accessions (Phippen et al. 1997; Dean et al. 1999). In our study, differences among accessions comprising bulks and among the putative duplicates were detected using ten polymorphic microsatellite primer pairs on ten plants per accession. Our marker analyses revealed that some of the component accessions in the bulks are genetically dissimilar despite morphological and other similarities that led to their bulking. In general, as the number of component accessions in a *B. rapa* bulk increased, so did the number of genetically unique components, as observed in Ames 26168 and Ames 26171. It is recommended that these bulks be re-formulated using only genetically homogeneous accessions and excluding the unique ones.

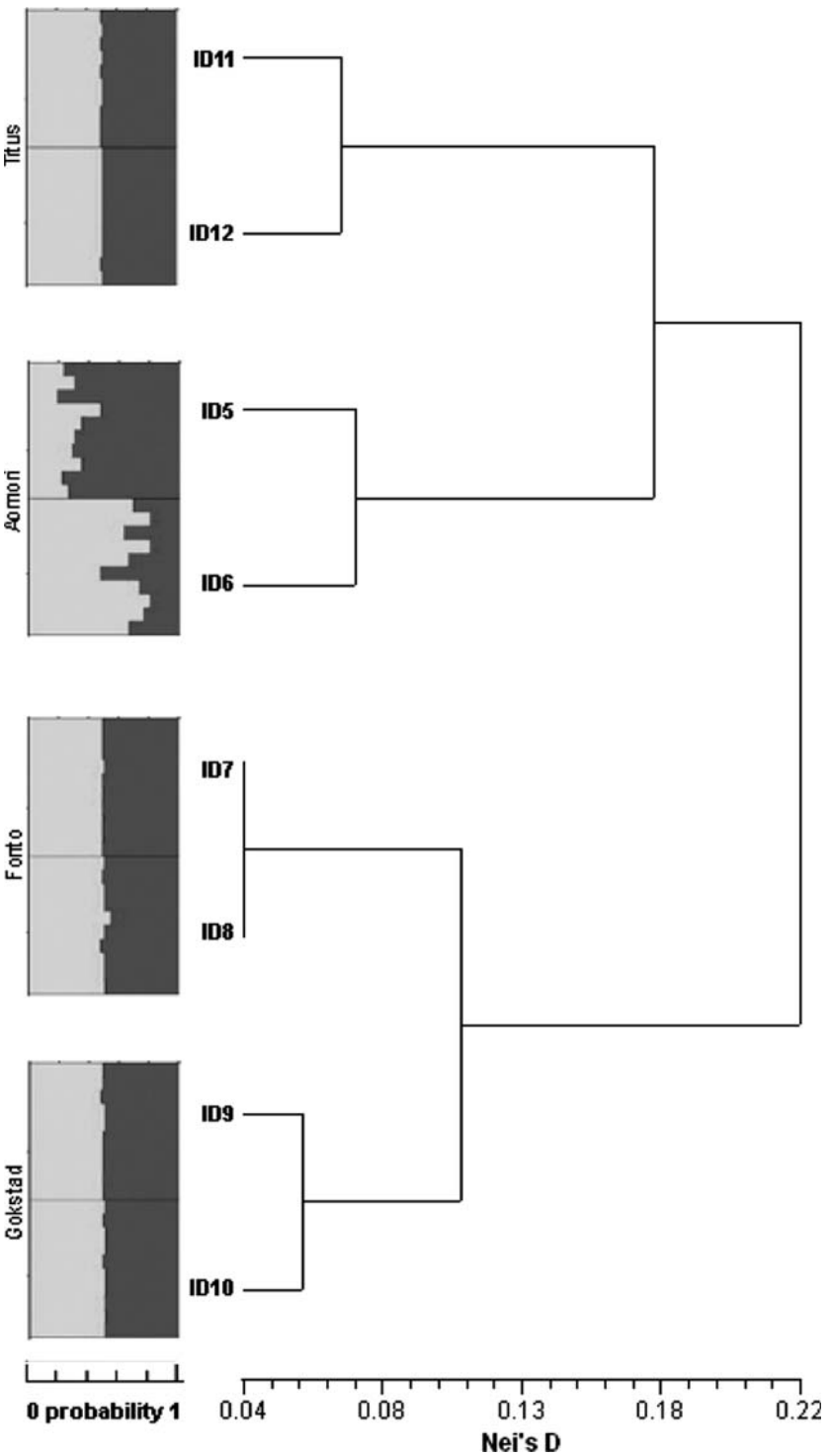

Fig. 5 Dendrogram constructed from Nei's genetic distance showing the relationships among the putative *B. napus* duplicates indicated in Table 1. Probabilities of assignment within an accession are indicated to the left. Probabilities of 0.5 indicate genetic homogeneity of putative duplicates, whereas consistently skewed probabilities (see Aomori) indicate genetic heterogeneity of putative duplicates

In *B. rapa* ssp. *trilocularis*, probabilities of assignment for a few plants of Ames 9698 [ID18] and Ames 9701 [ID20] suggest that these accessions are heterogeneous, or that a low frequency of admixture might have occurred between these accessions and ID19. A similar case

was observed in Ames 9919 [ID29] and Ames 9920 [ID30], with single plants having been assigned to ID30 and ID28, respectively. Original seed lots were used in this study; admixture could have occurred before the germplasm was acquired or during seed processing and incorporation into the collection. If this is the case, identification of probable off-types in populations of these accessions prior to flower during regeneration will help prevent genetic contamination due to gene flow events.

In general, in all *B. rapa* bulks there was concordance between relationships revealed in the dendrograms, the probabilities of assignment derived from Structure, and the pairwise tests of population differentiation. The dendrograms clustered accessions with similar probability profiles, as evident in the Ames 26168, Ames 26170, and Ames 26171 bulks. The unique components of these accessions did not cluster with those that included misassigned plants. Similarly, tests of population differentiation support the relationship among accessions as indicated by the dendrogram. Accessions that are not significantly different are located in the same cluster.

Microsatellite profiles of the *B. napus* accessions differentiated putative duplicates by variety name and, with the exception of the Aomori accessions, confirmed them to be good candidates for bulking. Overall, bulking these probable duplicates is a compromise strategy that will conserve genebank resource and also increase the likelihood that unique alleles (if present) are not lost over cycles of regeneration (Hintum and Visser 1995). As the cost of one cycle of *Brassica* germplasm regeneration in the NCRPIS is more than $500 per accession, bulking the redundant pairs also ensures cost efficiency, reducing the recurrent costs of germplasm maintenance.

It has been suggested that most probable duplicates in genebanks are 'partial' or 'common' duplicates, with only part of the alleles or genotypes in common because of regeneration and management practices. This study did not test for 'common genotypes' in the probable duplicates in *B. napus*. However, if the Aomori accessions originally came from the same source population, factors contributing to their genetic differences may include natural selection during regeneration, contamination of seed lots, or resulting from intentional splitting of accessions into morphologically distinct parts (Hintum and Knüpffer 1995).

Most of the approaches used in identifying duplicates rely on evaluating uniqueness of genetic profiles instead of similarity among the candidate accessions (Dean et al. 1999; Lund et al. 2003). Uniqueness has been favored over similarity because the high resolving power of molecular markers will always allow detection of some level of variation between 'identical' accessions (Treuren and Hintum 2003). Consequently, only resource limitations restrict the discovery of variation between probable duplicates. In our study, we find that information, in the form of misassignment probabilities, genetic distance, and population differentiation, when evaluated concurrently, provides corroborative evidence that can help the crop curator decide which accessions are unique or are more likely candidates for bulking when such situations arise.

Conclusion

This study demonstrates that analyses of a modest number of microsatellites (10 loci) and molecular characterization data using assignment tests, together with more traditional methods, can be useful in verifying the homogeneity of component accessions of bulks in *B. rapa* and addressing the problem of duplicate accessions in *B. napus*. Unique accessions among *B. rapa* bulks and *B. napus* putative duplicates were identified. Use of genetic probability profiles of component accessions corroborated clustering analyses while providing additional information. Tests for population differentiation also added assurance on the uniqueness of component accessions in *B. rapa*.

Acknowledgements This is a journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 1018, and was supported by Hatch Act and the State of Iowa. The authors wish to thank Ms. Jody Hayes and Ms. Lori Lincoln (USDA-ARS Corn Insect and Crop Genetics Research Unit) for help in the laboratory, and Dr. Mark P. Widrlechner for reviewing the draft manuscript. Mention of commercial brand names in this paper does not constitute an endorsement of any product by the U.S. Department of Agriculture or cooperating agencies.

References

Cornuet JM, Luikart G (1996) Description and evaluation of two tests for detecting recent bottlenecks. Genetics 144:2001–2014

Cruz VMV (2006) Molecular genetic analysis of oilseed *Brassica*: determination of life forms and germplasm management strategies by using microsatellite markers and *FLOWERING LOCUS C* gene sequences. PhD dissertation. Iowa State University, Ames, Iowa

Davies N, Villablanco FX, Roderick GK (1999) Determining the source of individuals: multilocus genotyping in nonequilibrium population genetics. TREE 14(1):17–21

Dean RE, Dahlberg JA, Hopkins MS, Mitchell SE, Kresovich S (1999) Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci 39:1215–1221

Eastham K, Sweet J, (2002) Genetically modified organisms (GMOs): the significance of gene flow through pollen transfer. EEA Environmental Issue Report No. 28. European Environment Agency, Copenhagen, Denmark

Hintum TJL, Knüpffer H (1995) Duplication within and between germplasm collections. I. Identifying duplication on the bases of passport data. Genet Res Crop Evol 42:127–133

Hintum TJL, Visser DL (1995) Duplication within and between germplasm collections. II. Duplication in four European barley collections. Genet Res Crop Evol 42:135–145

Hintum TJL, Boukema IW, Visser DL (1996) Reduction of duplication in a *Brassica oleracea* germplasm collection. Genet Res Crop Evol 43:343–349

Houmiel KL, Slater S, Broyles D, Casagrande L, Colburn S, Gonzalez K, Mitsky TA, Reiser SE, Shah D, Taylor NB (1999) Poly(beta-hydroxybutyrate) production in oilseed leukoplasts of *Brassica napus*. Planta 209:547–550

Khachatourians GG, Summer AK, Phillips PWB (2001) An introduction to the history of canola and the scientific basis of innovation. In: Phillips PWB, Khachatourians GG (eds) The biotechnology revolution in global agriculture. CABI, Wallingford, UK, pp 33–47

Kimber DS, McGregor DI (1995) The species and their origin, cultivation and world production. In: Kimber DS, McGregor DI (eds), *Brassica* oilseeds: production and utilization. CABI, Oxon, UK, pp 1–8

Lund B, Ortiz R, Skovgaard IM, Waugh R, Anderson SB (2003) Analysis of potential duplicates in barley gene bank collections using re-sampling of microsatellite data. Theor Appl Genet 106:1129–1138

McNaughton IH (1995) Swedes and rapes – *Brassica napus* (Cruciferae). In: Smartt J, Simmonds NW (eds) Evolution of crop plants, 2nd edn. Longman Scientific & Technical, London, UK pp 68–75

Nei M (1972) Genetic distance between populations. American Naturalist 106:283–392

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Petkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. Mol Ecol 4:347–354

Phippen WB, Kresovich S, Candelas FG, McFerson JR (1997) Molecular characterization can quantify and partition genetic 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

Primmer CR, Koskinen MT, Piironen J (2000) The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. Proc R Soc Lond 267:1699–1701

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rakow G, Woods D (1987) Outcrossing in rape and mustard in Saskatchewan prairie conditions. Can J Plant Sci 67:147–151

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Heredity 86:248–249

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Rohlf FJ (2005) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, ver. 2.2. Exeter Software, Setauket, NY

Snowdon RJ, Friedt W (2004) Molecular markers in *Brassica* oilseed breeding: current status and future possibilities. Plant Breed 123:1–8

Treuren R, Hintum TJL (2003) Marker-assisted reduction of redundancy in germplasm collections: genetic and economic aspects. Acta Hort 623:139–149

Treuren R, Soest LMJ, Hintum TJL (2001) Marker-assisted rationalization of genetic resources collections: a case study in flax using AFLPs. Theor Appl Genet 103:144–152

Waser PM, Strobeck C (1998) Genetic signatures of interpopulation dispersal. TREE 13(2):43–44