Theor Appl Genet (2002) 104:104–110

A.M. Casa · S.E. Mitchell · O.S. Smith
J.C. Register III · S.R. Wessler · S. Kresovich

Evaluation of *Hbr* (MITE) markers for assessment of genetic relationships among maize (*Zea mays* L.) inbred lines

Received: 26 February 2001 / Accepted: 20 April 2001

Abstract Recently, a new type of molecular marker has been developed that is based on the presence or absence of the miniature inverted repeat transposable element (MITE) family *Heartbreaker* (*Hbr*) in the maize genome. These so-called *Hbr* markers have been shown to be stable, highly polymorphic, easily mapped, and evenly distributed throughout the maize genome. In this work, we used *Hbr*-derived markers for genetic characterization of a set of maize inbred lines belonging to Stiff Stalk (SS) and Non-Stiff Stalk (NSS) heterotic groups. In total, 111 markers were evaluated across 62 SS and NSS lines. Seventy six markers (68%) were shared between the two groups, and 25 of the common markers occurred at fairly low frequency (≤0.20). Only two markers (3%) were monomorphic in all samples. Although DNA sequencing indicated that 5.5% of same-sized DNA fragments were non-homologous, this result did not affect the cluster analyses (i.e., relationships obtained from the *Hbr* data were congruent with those derived from pedigree information). Distance matrices generated from *Hbr* markers were significantly correlated (p<0.001) with those obtained from pedigree (r=0.782), RFLPs (r=0.747), and SSRs (r=0.719). Overall, these results indicated that *Hbr* markers could be used in conjunction with other molecular markers for genotyping and relationship studies of related maize inbred lines.

Keywords *Hbr* display · MITEs · Molecular markers · Stiff Stalk · Non-Stiff Stalk

Communicated by D. Hoisington

A.M. Casa
Institute for Genomic Diversity and Department of Plant Breeding, Cornell University, 153 Biotechnology Building, Ithaca, NY 14853, USA

S.E. Mitchell
Institute for Genomic Diversity, Cornell University, 151 Biotechnology Building, Ithaca, NY 14853, USA

O.S. Smith · J.C. Register III
Pioneer Hi-Bred International, Inc., 7300 N.W. 62nd Avenue, Johnston, IA 50131, USA

S.R. Wessler
Departments of Botany and Genetics, The University of Georgia, Plant Sciences Building, Athens, GA 30602, USA

S. Kresovich (✉)
Institute for Genomic Diversity and Department of Plant Breeding, Cornell University, 158 Biotechnology Building, Ithaca, NY 14853, USA
e-mail: sk20@cornell.edu
Tel.: +607-255-1492, Fax: +607-255-6249

Introduction

Breeding strategies for producing modern, high yielding hybrids rely on the use of heterosis, or phenotypic superiority of F_1 progeny relative to parental lines. In maize, crosses between genetically divergent lines generally produce better hybrids than crosses between closely related parents (East 1908; Shull 1908). This observation led breeders to define distinct heterotic groups and assign lines to particular groups based on yield performance in crosses (reviewed by Hallauer 1999).

In the U.S. Corn Belt, maize hybrids are usually produced by crossing inbred lines from Stiff Stalk (SS) and Non-Stiff Stalk (NSS) heterotic groups. Stiff Stalk germplasm originated primarily from the Iowa Stiff Stalk Synthetic (BSSS) population developed by Sprague (1946). Many historically important lines, including B14, B37 and B73, were derived from this synthetic population (Smith et al. 1985). Inbred lines from the Lancaster Surecropper and Iodent populations form the genetic foundation of what is here designated as Non-Stiff Stalk germplasm, and these lines perform extremely well (i.e., produce high yielding hybrids) in crosses with the BSSS-related lines.

Although relationships among inbreds are most easily determined by inspection of pedigrees, molecular markers are valuable tools for establishing relatedness when pedigree data are lacking. Comparison of results from molecular data with pedigree information has shown that RFLPs and SSRs are good predictors of genetic relatedness among maize inbred lines (Smith and Smith 1992; Smith et al. 1997; Pejic et al. 1998). Because RFLPs and

SSRs are co-dominant (multiple alleles are detected at a given locus) and randomly distributed throughout the genome, these markers have also been widely used for assessing maize genetic diversity (Smith et al. 1991, 1997; Smith and Smith 1992; Messmer et al. 1993). Molecular data are useful not only for planning crosses and identifying inbreds for plant variety protection (Melchinger et al. 1991), but they may also provide clues to the underlying genetic basis of the empirical phenomenon known as heterosis.

Recently, a new type of molecular marker based on the presence of the *Heartbreaker* (*Hbr*) family of miniature inverted repeat transposable elements (MITEs) has been described in maize (Casa et al. 2000). Here, a technique similar to AFLP (Vos et al. 1995) is used in which genomic DNA is digested, and AFLP adaptors are ligated to compatible ends. Subsets of MITE-containing fragments are then amplified using one AFLP primer and another primer complementary to an internal sequence of the *Hbr* element. Like AFLP, numerous DNA fragments are then analyzed simultaneously. The *Hbr* family was selected for this application because these elements occur in high numbers (3,000–4,000 copies per genome), have striking within-family sequence identity (>90%), and seem to insert into single- or low-copy genomic regions (Zhang et al. 2000).

The *Hbr*-derived markers were shown to be stable, highly polymorphic, and evenly distributed throughout the maize genome (Casa et al. 2000). Although more than 200 *Hbr* markers were mapped (Casa et al. 2000), they were not used to investigate maize genetic diversity or relationships. In the present study, we used *Hbr*-anchored markers to estimate genetic relationships between and within proprietary and public maize inbred lines widely used in the development of modern, high-yielding maize hybrids. Relationships obtained from the *Hbr* display were compared with results based on other molecular markers (SSRs and RFLPs) and pedigree data. Advantages and limitations of using *Hbr* or other MITE-based markers in genetic analyses are discussed.

Materials and methods

Plant material

The 62 maize accessions analyzed in this study are listed in Table 1. The test population included public SS (A632, B14, B37, B64, and B73) and NSS (F2 and Mo17) lines, and also proprietary inbred lines representing both SS and NSS heterotic groups. Pedigree relationships within the SS and NSS groups are presented in Fig. 1a and b, respectively. DNA was extracted from pooled leaf tissue (5–8 plants) according to Rogers and Bendich (1985).

Hbr *display*

Hbr display was performed as described by Casa et al. (2000). Briefly, 500 ng of maize total genomic DNA was digested to completion with *Mse*I. Adaptors (5´-GACGATGAGTCCTGAG and 5´-TACTCAGGACTCAT) were ligated to the digested DNAs, and aliquots of the restriction/ligation reactions were visualized on 0.8% agarose gels to check the quality of DNA digestion. Pre-

Table 1 Inbreds and their respective parental lines

Accession ID	Parental lines
A632	Mt42×B14
B14	BSSS
B37	BSSS
B64	CBA×B14
B73[a]	BSSS
SS01[a]	B37
SS02[a]	B73×B37B
SS03	DIS-SYN X B37 A
SS04[a]	B37D×B68B
SS05	B64
SS07	B73HT×B37F
SS09[a]	SS01×SS02
SS10	NSS48×SS02
SS11	SS09×NSS52 A
SS19[a]	SS23×SS17
SS20[a]	SS04×B73
SS21	SS04×B37F
SS22	SS04×SS03
SS23	SS09×SS04
SS24	SS07×SS04
SS25	SS20×SS22
SS26	SS24×SS10
SS27	SS20×PESSO
SS30[a]	SS05×SS20
SS31	SS23×SS30
SS32	B73×SS23
SS33	SS23×SS30
SS34	SS11×SS98
F2	Northern European Flint OP
Mo17[a]	187–2×C103
NSS12	NSS96×SS40
NSS14[a]	NSS16×NSS37
NSS15	NSS06×NSS16
NSS16[a]	Eight diverse lines
NSS35	Southern Dent OP
NSS37	ID2×ID1
NSS38	NSS37×Iodent
NSS39	NSS37×B68C
NSS40	ID3×NSS37
NSS41[a]	ID4×NSS37
NSS42[a]	NSS40×NSS39
NSS43	NSS40×NSS54
NSS44	NSS40×NSS39
NSS45	NSS41×NSS42
NSS46	NSS42×NSS43
NSS47[a]	FS1
NSS48	NSS75×Oh43A
NSS49	NSS13×NSS47
NSS50	NSS40×NSS52
NSS51[a]	NSS38×NSS55
NSS52	NSS47×NSS91
NSS53	ID1×NSS47
NSS54	NSS37×NSS47
NSS55	NSS16×NSS47
NSS56	NSS39×NSS53
NSS57[a]	NSS48×NSS53
NSS58	NSS49×NSS15
NSS59	NSS42×NSS50
NSS60	NSS51×NSS50
NSS61	NSS51×NSS58
NSS62	NSS51×NSS58
NSS63	NSS44×NSS51

[a] Inbred lines from which fragments were cloned and sequenced

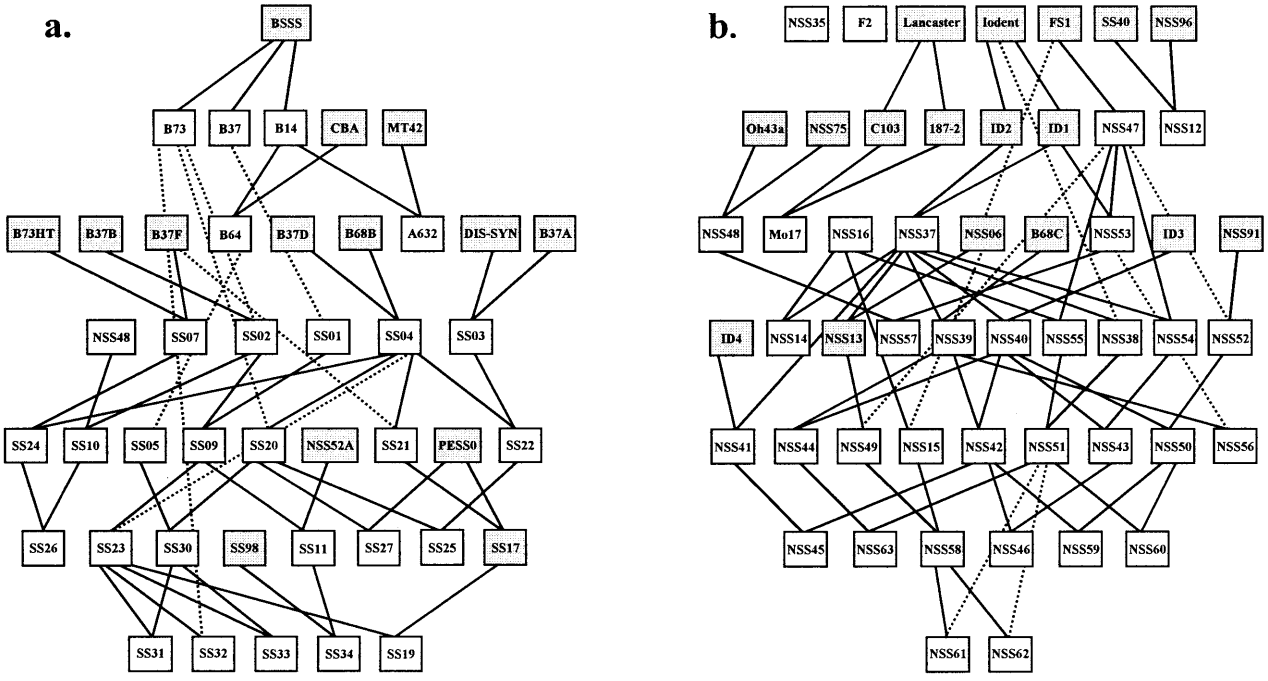

Fig. 1 Pedigree relationships among SS (**a**), and NSS (**b**) inbred lines. *Shaded boxes* identify inbred lines not assayed in this study. *Dashed lines* were used to facilitate the localization of parental lines on upper levels of the network

selective amplifications were first performed with primers *Hbr*Int5-E (5´-GATTCTCCCCACAGCCAGATTC) and *Mse*I+0 (5´-GACGATGAGTCCTGAGTAA). These reaction products served as templates in a second round of selective amplifications with either *Mse*I+G or *Mse*I+T (5´-GACGATGAGTCCTGAGTAAG/T) and an *Hbr* internal primer, *Hbr*Int5-F (5´-6FAM-AGCCAGATTTTCAGAAAAGCTG). Electrophoresis, detection, and sizing of fluorescent fragments were performed following established protocols (Casa et al. 2000).

Data analysis

Hbr display

A file containing all sizing data was created using Genotyper (v.2.5, Applied Biosystems), and a matrix was constructed by scoring fragments as either present (1) or absent (0) in each DNA sample. Genetic distances (GD) between pairs of accessions were calculated using the AFLPdist (Rob Dean, University of Georgia) option in Phylip (v. 3.5, Felsenstein 1993) as follows:

$$GD_{ij} = 1 - 2N_{ij}/[N_i + N_j],$$

where N_{ij} is the total number of fragments common to lines i and j, and N_i and N_j are the total number of fragments present in i and j, respectively (Dice 1945). Phenograms were constructed in NTSYS-pc v. 2.0 (Rohlf 1998) using the Unweighted Pair-Group Method with Arithmetic mean (UPGMA).

Two-sample *t* tests (Devore and Peck 1997) were performed to determine if the average genetic distances between SS and NSS heterotic groups were statistically significant.

Pedigree

The coancestry coefficient, *f* (Malécot 1948), was used to quantify the degree of relatedness of inbred lines based on available pedigree information. The coancestry coefficient between two lines is defined as the probability that a random allele of one line is identical by descent to a random allele of the other line. The genealogical distance between lines i and j was computed as:

$$Dg_{ij} = 1 - f_{ij}$$

For sister lines, *f* was calculated as follows:

$$f = f_{ij} + (1 - f_{ij})(1 - 0.5^{(M-1)}),$$

where i and j are the parents of the sister lines, and M is the number of generations of selfing before sisters were separated. For lines with unknown genetic relationships, we assumed that $f=0$.

Correlation tests

Correspondence between genetic-distance matrices derived from the *Hbr* datasets, and genetic distances based on RFLPs, SSRs and pedigree, was tested with the Mantel Z statistic (Mantel 1967) implemented in NTSYS-pc v. 2.0 (Rohlf 1998). The RFLP, SSR, and pedigree data for the 62 maize accessions included in this study were provided by Pioneer Hi-bred International (Johnston, Iowa). RFLP (n=70), and SSR (n=50) markers were evenly distributed in the maize genome and polymorphism information content (PIC) values derived from a larger set of germplasm were 0.60 and 0.56, respectively.

Marker frequency distribution

For calculating marker frequencies we assumed that same-sized fragments were homologous. Estimation was performed as follows:

$$pi = 2Nii/2\,N,$$

where p*i* is the frequency of marker *i*, N*ii* represents the number of individuals carrying marker *i*, and N the total number of individuals. An exact test for population differentiation was performed according to Raymond and Rousset (1995), implemented in Tools for Population Genetic Analyses (TFPGA) (http://www.public.asu.edu/~mmile8/) (Miller 1996).

Isolation, cloning, and sequencing of amplified fragments

Fifty two radio-labeled DNA fragments, 108 to 410 bp in length, were excised from gels, eluted in buffer (0.5 M ammonium ace-

tate, 10 mM $MgCl_2$, 0.1% SDS, 1 mM EDTA, pH 8.0), precipitated with ethanol, and suspended in 10.0 µl of TE (10 mM Tris-HCl, pH 8.0; 1 mM EDTA). Fragments were amplified in 50.0-µl volumes containing 5.0 µl DNA, 12pmol of each primer (*MseI*+G and *Hbr*Int5-F), 1×PCR buffer (Promega), 0.2 mM dNTPs, 2.5 mM $MgCl_2$ and 1 U of *Taq* polymerase (Promega). Polymerase chain reactions (PCRs) were performed using a "touchdown" cycling protocol as follows: 94°C/5 min; followed by 94°C/30sec, 70°C/30sec, and 72°C/1 min. In subsequent cycles, the annealing temperature was reduced from 69°C to 61°C in 1°C increments each cycle. Twenty seven cycles were performed at the 61°C annealing temperature, followed by a final cycle of 72°C/5 min. Reactions were purified following the QIAquick PCR purification kit protocol (QIAGEN) and fragments were cloned using a commercial kit (TA cloning kit, INVITROGEN). DNA sequencing was done at the BioResource Center (Cornell University) using fluorescent dideoxyterminator chemistry and automated DNA sequencers (Model 3700, Applied Biosystems). DNA sequences were edited with Sequencher (v.3.1.1 Gene Codes Corp.).

Results

Genetic variation among 62 maize inbred lines comprising both SS and NSS germplasm was estimated by *Hbr* display. A total of 111 markers, ranging in size from 70 to 500 bp, were generated using two primer combinations, *Hbr*5-F/*Mse*I+G and *Hbr*5-F/*Mse*I+T. For primer combination *Hbr*5-F/*Mse*I+T, unique fingerprints were obtained in all 62 lines assayed; while *Hbr*5-F/*Mse*I+G generated 61 fragment patterns (lines NSS41 and NSS45 were identical).

Genetic distances and cluster analysis

Among the 28 SS lines, both primer sets combined gave a total of 85 markers. Genetic distances within the SS group ranged from 0.06 (between sister lines SS31 and SS33) to 0.68 (A632 vs SS01; SS09 vs SS21), with an average distance value of 0.40. In comparison, a total of 102 fragments was amplified in the 34 NSS lines. The average genetic distance among NSS germplasm (0.50) was significantly higher ($p<0.001$) than in SS lines, ranging from 0.05 (NSS40 vs NSS44; NSS41 vs NSS45) to 0.79 (F2 vs NSS49; F2 vs NSS58). This result is consistent with other studies that evaluated genetic distances between heterotic groups using RFLP, AFLP and SSR markers (Ajmone-Marsan et al. 1992; Dubreuil et al. 1996; Pejic et al. 1998). Genetic distances between SS and NSS accessions ranged from 0.32 (SS11 vs NSS35) to 0.77 (A632 vs NSS62; SS22 vs NSS36), with an average between-group distance of 0.58.

Results from the UPGMA analysis (Fig. 2) were generally consistent with pedigree data and confirmed two major groupings comprising the SS and NSS germplasm. Compared to the NSS lines, the SS lines formed a more-cohesive grouping. This result probably reflects the common genetic origin of most SS germplasm from the Iowa Stiff Stalk Synthetic (BSSS) population (Fig. 1a).

Overall, clustering of the NSS lines was more variable. One accession, NSS48, grouped with the SS lines (Fig. 2). Pedigree information, however, indicated

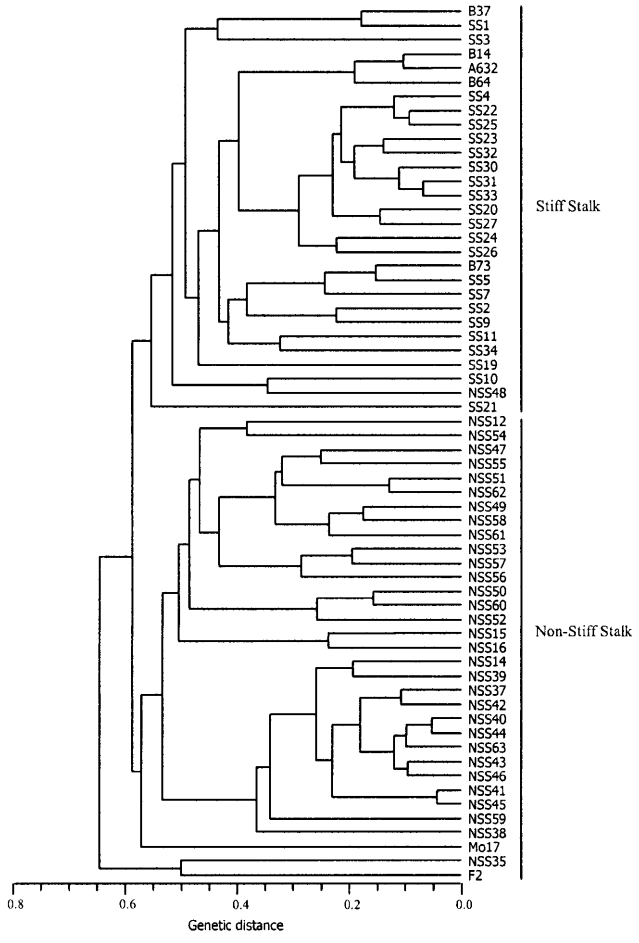

Fig. 2 Phenetic analysis of SS and NSS inbred lines based on *Hbr*-derived markers. Genetic distances were calculated using the Dice coefficient (Dice 1945). The UPGMA dendrogram was obtained using NTSYS-pc v 2.0 (Rohlf 1998). Pedigree information for inbred lines is listed in Table 1

that this inbred was a progenitor of SS10 (Table 1 and Fig. 1a). Within the NSS group, Mo17 did not show a close relationship to any other NSS inbred. This result is consistent with both the line origin (a cross between public inbred lines 187–2×C103; Table 1) and the fact that there were no direct descendants from Mo17 included in the test sample (Fig. 1b). NSS35, a line derived from a Southern Dent open-pollinated (OP) variety, and Flint Public Line F2 clustered outside the major groupings (Fig. 2). This relationship is surprising since NSS35 and F2 have such distinct genetic origins. This association is probably an artifact of the UPGMA clustering algorithm. That is, these lines grouped because they were genetically closer to each other (GD=0.50) than either was to the rest of the lines assayed (average GD=0.64).

For 60 of the 62 lines assayed, phenetic analysis using *Hbr* markers was congruent with expectations based on pedigrees. However, relationships derived for two lines, NSS12 and NSS54, were inconsistent with the pedigree data (see Fig. 1b and 2).

Table 2 Results of correlation tests obtained from comparing genetic distance matrices derived from *Hbr*, RFLP, SSR and pedigree data

Marker	*Hbr*+G	*Hbr*+T	*Hbr*+G/T	RFLP	SSR	Pedigree
Hbr+G	1.0[a]	14.63	29.60	21.97	23.4	25.78
Hbr+T	0.506	1.0	22.61	14.96	15.77	17.08
Hbr+G/T	0.865	0.865	1.0	22.75	23.94	26.05
RFLP	0.667	0.609	0.747	1.0	23.67	25.81
SSR	0.659	0.567	0.719	0.740	1.0	28.20
Pedigree	0.727	0.614	0.782	0.808	0.814	1.0

[a] Numbers above the diagonal represent the approximate Mantel *t*-test. Numbers below the diagonal denote the correlation coefficient between the distance matrices generated by each of the markers indicated and by *f* values based on pedigree data. All values are significant ($p<0.001$)

Correlation tests

Correlation statistics among genetic-distance matrices from molecular marker data (*Hbr*, SSRs, and RFLPs) and coancestry coefficients from pedigrees are shown in Table 2. For the *Hbr* primer sets, *Hbr*Int5-F/*Mse*I+G and *Hbr*Int5-F/*Mse*I+T, correlations were estimated based on data from each set, individually, and from both sets combined. Although all correlation values were statistically significant ($p<0.001$), distance matrices derived from all *Hbr* markers combined were more highly correlated with pedigree, RFLP and SSR data, than were distances from the individual *Hbr* datasets. Interestingly, the lowest correlation value observed was between the two *Hbr* distance matrices. Because we evaluated similar numbers of markers for *Hbr*+G and *Hbr*+T, the lower correlation is not due to sampling a disparate numbers of markers. The result may be due either to differences in genomic distribution or homoplasy levels among the marker sets.

Marker frequency distribution and population differentiation

Seventy six (68%) of 111 markers were shared between the two heterotic groups. Twenty-five (37% of the shared markers) occurred at low frequency (≤0.20) in both groups. Only 3% of the shared *Hbr* markers (2/76) were monomorphic in both populations. Three additional markers were fixed in the SS population, and one other marker was fixed in the NSS group. All four markers, however, were also present at high frequency (>0.50) in the alternate population. Nine markers (8.5%) were unique to the SS lines, while 26 (23.5%) were present only in the NSS material. Remarkably, the range in frequency for these unique markers was similar in both groups (from 0.03 to 0.57).

An exact test for population differentiation showed that SS and NSS populations were differentiated ($p<0.01$) at 48% of the loci examined. These results indicate that both unique and shared alleles with disparate frequencies between the two populations have contributed to population differentiation.

DNA sequence analysis

DNA sequencing was performed to determine if same-sized *Hbr* markers were homologous and if individual bands contained single sequences or populations of co-migrating fragments. In all, 52 *Hbr*-anchored markers comprising 19 fragment sets were cloned and sequenced (Fig. 3). For ten sets, fragments were isolated from inbred lines belonging to both SS and NSS populations, while four sets originated from SS lines and the remaining five sets were from NSS lines.

Three out of 19 fragment sets consisted of a mixture of two different DNA sequences (fragment set 2 in SS lines and fragment sets 1 and 8 in the NSS) (Fig. 3a and b). Among co-migrating fragments that were present in both heterotic groups (fragment sets 1 and 8), at least one of the DNA sequences obtained was common to both populations (i.e., these fragment sets shared

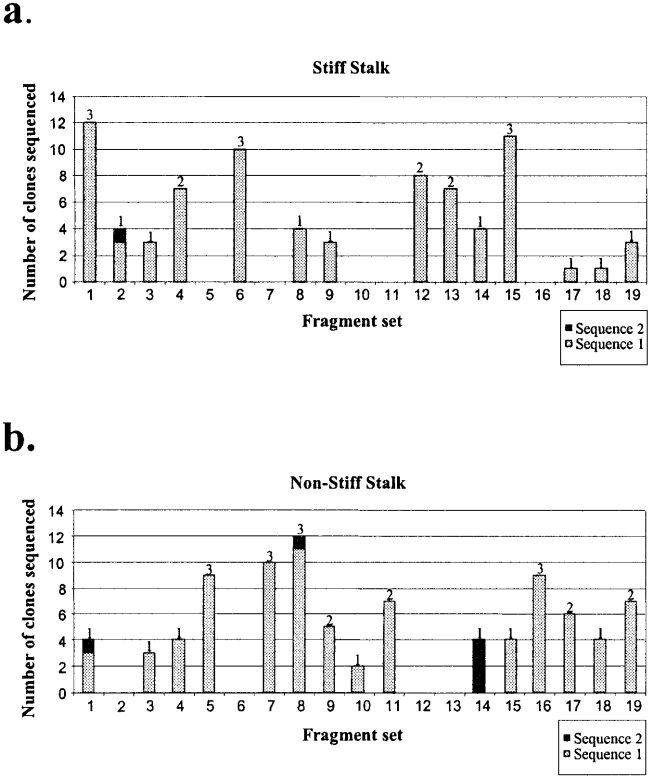

Fig. 3 Summary of sequencing results from SS (**a**) and NSS (**b**)-derived fragments. *Numbers on the X-axis* identify fragment sets. *Numbers on the Y-axis* represent the number of clones sequenced for each fragment set. *Number above bars* indicates the number of inbred lines sampled. Similar shading within each fragment-set denotes clones having identical DNA sequences

homologous sequences). There was one fragment set, however, that was completely non-homologous between groups (fragment set 14) (Fig. 3a and b). Assuming that the overall rate of homoplasy is similar for all *Hbr* markers, these results indicate that 5.5% of same-sized fragments will be non-homologous among U.S. Corn Belt lines.

Discussion

Utility of *Hbr*-derived markers in relationship studies

Molecular characterization of maize inbreds used in the development of elite lines has become an important component of modern plant breeding. In addition, genotyping techniques such as RFLPs and SSRs have allowed the genetic discrimination of very closely related lines for purposes of plant variety protection and pedigree validation. In this study, we have demonstrated that genetic relationships predicted by *Hbr*-anchored markers were highly congruent with those derived from pedigree, SSR and RFLP data. UPGMA analysis revealed that lines were correctly partitioned into two major clusters consistent with the SS and NSS heterotic groups. In addition, the lower mean genetic-distance observed within the SS compared to the NSS group was consistent with both pedigree and breeding strategy. For example, the SS group was largely derived from a limited number of inbred lines (i.e., B14, B37 and B73), whereas NSS lines had a much-broader genetic base (originated from a variety of U.S. landraces). Moreover, hybrids developed using SS accessions as the female parent produce higher yields. Therefore, selection for female traits in SS lines would more likely have led to a reduction in population size.

Although placement of a few lines did not correspond precisely with expectation based on pedigree data, similar inconsistencies have also been observed for RFLP, SSR, and AFLP markers (Smith et al. 1997; Ajmone-Marsan et al. 1998). Minor variations from pedigree information, therefore, are not likely to be caused by some unique characteristic of the *Hbr* markers.

For any molecular marker, scoring errors (Messmer et al. 1993), the number and genomic distribution of the loci assayed, and the amount of linkage disequilibrium affect relationship estimates (Powell et al. 1996). Portrayal of genetic relationships among inbred lines is also compromised if the theoretical assumptions underlying the estimation of the coancestry coefficient (f) are violated. Estimates of relationship based on pedigree assume an equal genetic contribution from each parent, and the absence of selection and mutation. The set of lines assayed, however, represent highly selected genotypes that violate the non-selection assumption and most likely result in pedigree distance inaccuracies. Since molecular markers sample the genome directly, deviations due to selection and/or genetic drift can be tracked throughout the development of inbred lines. Consequently, molecular markers provide a more accurate estimate of genetic relationship than pedigree data (Bernardo et al. 2000).

Advantages and limitations of *Hbr*-derived markers

Hbr display is a highly reproducible PCR-based protocol where multiple fragments are simultaneously detected using only a few primers (universal adaptor and MITE-derived primers) (Casa et al. 2000). Therefore, this technique permits the collection of large amounts of genetic data with minimal effort.

Hbr markers suffer the same constraints as AFLPs and, consequently, may not be well-suited for some analyses. Like their AFLP cousins, *Hbr*-anchored markers are mostly dominant, and their use will be limited in studies where discrimination of multiple alleles at a locus is required. Also, the large number of products generated by *Hbr* display increases the probability of non-homology among same-sized fragments (i.e., homoplasy). Comparative mapping results have shown that approximately 7% of co-migrating *Hbr* markers were non-homologous (Casa et al. 2000). In this study, direct comparison of DNA sequences revealed that the frequency of homoplasy for *Hbr*-anchored markers was slightly lower (5.5%) than the previous estimate and did not influence relationships obtained for the inbred lines analyzed. However, *Hbr* markers should be used with caution for genotyping genetically diverse accessions. Additional DNA sequencing results indicated that lack of fragment homology seems higher in comparisons among more-distantly related inbred lines (A.M.C. and S.K., unpublished). Our DNA sequence data also revealed that homoplasy could be completely eliminated by using *Mse*I selective primers that contain two selective bases. The resulting loss of information per assay, therefore, should be weighed against the impact of homoplasy on specific experiments.

In maize, transposable elements have been exploited as markers for gene cloning, phylogenetic analysis, and mapping (Walbot 1992; Purugganan and Wessler 1995; Casa et al. 2000). Results presented here have indicated that the *Hbr* family of MITEs is useful for both genotyping and predicting genetic relationships among sets of closely related maize inbred lines. Because a growing number of reports confirm the presence of MITE-like elements in a variety of other plants (Bureau and Wessler 1994; Pozueta-Romero et al. 1995; Charrier et al. 1999) and animals (Oosumi et al. 1995; Tu 1997; Izsvak et al. 1999), the potential for using MITE families as molecular markers should extend beyond maize.

Acknowledgements We thank Thomas Brutnell, Joanne Labate, Margaret Smith, and the anonymous reviewers for their insightful comments and suggestions. We also thank Chip Aquadro for making his laboratory available for the radioactive work. The work presented was funded by Pioneer Hi-Bred International (to A.M.C., S.R.W. and S.K.).

References

Ajmone-Marsan P, Livini C, Messmer MM, Melchinger AE, Motto M (1992) Cluster analysis of RFLP data from related maize inbred lines of the BSSS and LSC heterotic groups and comparison with pedigree data. Euphytica 60:139–148

Ajmone-Marsan P, Castiglioni P, Fusari F, Kuiper M, Motto M (1998) Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. Theor Appl Genet 96:219–227

Bernardo R, Romero-Severson J, Ziegle J, Hauser J, Joe L, Hookstra G, Doerge RW (2000) Parental contribution and coefficient of co-ancestry among maize inbreds: pedigree, RFLP, and SSR data. Theor Appl Genet 100:552–556

Bureau TE, Wessler SR (1994) Mobile inverted repeat elements of the *Tourist* family are associated with the genes of many cereal grasses. Proc Natl Acad Sci USA 91:1411–1415

Casa AM, Brouwer C, Nagel A, Wang L, Zhang Q, Kresovich S, Wessler SR (2000) The MITE family *Heartbreaker* (*Hbr*): molecular markers in maize. Proc Natl Acad Sci USA 97: 10083–10090

Charrier B, Foucher F, Kondorosi E, Carafa Yd'A, Thermes C, Kondorosi A, Ratet P (1999) Bigfoot: a new family of MITE elements characterized from the *Medicago* genus. Plant J 18:431–441

Devore J, Peck R (1997) Statistics: The exploration and analysis of data, 3rd edn. Duxbury Press, New York, pp 363–364

Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302

Dubreuil P, Dufour P, Krejci E, Causse M, deVienne D, Gallais A, Charcosset A (1996) Organization of RFLP diversity among inbred lines of maize representing the most significant heterotic groups. Crop Sci 36:790–799

East EM (1908) Inbreeding in corn. Rep Conn Agric Exp Stn 419–428

Felsenstein J (1993) PHYLIP – Phylogeny Inference Package version 3.5c. Department of Genetics, University of Washington, Seattle, Washington

Hallauer AR (1999) Temperate maize and heterosis. In: The genetics and exploitation of heterosis in crops. Proceedings of an International Symposium, CIMMYT, Mexico City, Mexico, 17–22 Aug, 1997

Izsvak Z, Ivics Z, Shimoda N, Mohn D, Okamoto H, Hackett PB (1999) Short inverted repeat transposable elements in teleost fish and implications for a mechanism of their amplification. J Mol Evol 48:13–21

Malécot G (1948) Les mathématiques de l'hérédité. Masson et Cie, Paris

Mantel N (1967) The detection of disease clustering in a generalized regression approach. Cancer Res 27:209–220

Melchinger AE, Messmer MM, Lee M, Woodman WL, Lamkey KR (1991) Diversity and relationships among United States maize inbreds revealed by restriction fragment length polymorphisms. Crop Sci 31:669–678

Messmer MM, Melchinger AE, Herrmann RG, Boppenmaier J (1993) Relationships among early European maize inbreds. II. Comparison of pedigree and RFLP data. Crop Sci 33:944–950

Miller MP (1996) TFPGA: Tools for population genetic analyses version 1.3, Arizona

Oosumi T, Garlick B, Belknap WR (1995) Identification and characterization of putative transposable DNA elements in solanaceous plants and *Caenorhabditis elegans*. Proc Natl Acad Sci USA 92:8886–8890

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G, Motto M (1998) Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor Appl Genet 97: 1248–1255

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2:225–238

Pozueta-Romero J, Klein M, Houlne G, Schantz ML, Meyer B, Schantz R (1995) Characterization of a family of genes encoding a fruit-specific wound-stimulated protein of bell pepper (*Capsicum annuum*): identification of a new family of transposable elements. Plant Mol Biol 28:1011–1025

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

Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution 49:1280–1283

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5:69–73

Rohlf FJ (1998) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System version 2.0. Exeter Software, New York

Shull GH (1908) The composition of a field of maize. Am Breeders Assoc Rep 4:296–301

Smith JSC, Goodman MM, Stuber CW (1985) Genetic variability within U.S. maize germplasm. I. Historically important lines. Crop Sci 25:550–555

Smith JSC, Smith OS, Bowen SL, Tenborg RA, Wall SJ (1991) The description and assessment of distances between inbred lines of maize. III. A revised scheme for the testing of distinctiveness between inbred lines utilizing DNA RFLPs. Maydica 36:213–226

Smith JSC, Chin ECL, Shu H, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997) An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor Appl Genet 95:163–173

Smith OS, Smith JSC (1992) Measurement of genetic diversity among maize hybrids, a comparison of isozymic, RFLP, pedigree, and heterosis data. Maydica 37:53–60

Sprague GF (1946) Early testing of inbred lines of corn. J Am Soc Agron 38: 108–117

Tu Z (1997) Three novel families of miniature inverted repeat transposable elements are associated with genes of the yellow fever mosquito, *Aedes aegypti*. Proc Natl Acad Sci USA 94:7475–7480

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters BA, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Walbot V (1992) Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis Annu Rev Plant Mol Biol 43:527–552

Zhang Q, Arbuckle J, Wessler SR (2000) Recent, extensive, and preferential insertion of members of the family *Heartbreaker* into genic regions of maize. Proc Natl Acad Sci USA 97: 1160–1165