Utility of SSRs for Determining Genetic Similarities and Relationships in Maize Using an Agarose Gel System

M. L. Senior,* J. P. Murphy, M. M. Goodman, and C. W. Stuber

ABSTRACT

Among maize (*Zea maize* L.) breeders, there is a heightened awareness of the necessity for both maintaining genetic diversity for crop improvement and improving the quality of genetic resource management. Restriction fragment length polymorphisms (RFLPs) and isozymes can serve as genetic markers for estimating divergence or diversity; however, the limited number of polymorphic isozyme loci available and the labor intensive and time consuming nature of RFLPs make their use for this purpose prohibitive. Simple sequence repeats (SSRs), when resolved using agarose gels, may be a viable and cost-effective alternative to RFLPs and isozymes. Ninety-four elite maize inbred lines, representative of the genetic diversity among lines derived from the Corn Belt Dent and Southern Dent maize races, were assayed for polymorphism at 70 SSR marker loci using agarose gels. The 365 alleles identified served as raw data for estimating genetic similarities among these lines. The patterns of genetic divergence revealed by the SSR polymorphisms were consistent with known pedigrees. A cluster analysis placed the inbred lines in nine clusters that correspond to major heterotic groups or market classes for North American maize. A unique fingerprint for each inbred line could be obtained from as few as five SSR loci. The utility of polymerase chain reaction (PCR)-based markers such as SSRs for measuring genetic diversity, for assigning lines to heterotic groups and for genetic fingerprinting equals or exceeds that of RFLP markers, a property that may prove a valuable asset for a maize breeding program.

The improvement of any crop species is facilitated by introducing novel genes for agronomic traits such as yield, as well as traits such as pest resistance and abiotic stress tolerance. These new genes are most successfully gleaned from the primary gene pool of the crop species of interest. Many plant breeders are concerned that genetic diversity within such primary gene pools has been decreasing at an alarming rate as a consequence of modern agricultural practices (Lee, 1995). For example, many maize inbreds have been developed from a limited number of elite lines and elite line synthetics, a practice that heightens the risk of decreased genetic diversity in commercial maize production fields (Hallauer et al., 1988). Maize breeders have recently become more aware of the need for both maintaining genetic diversity and improving the management of genetic resources (Goodman, 1994). Information on the genetic similarity among lines selected in a breeding program could help avoid erosion.

A common measure of divergence between two cultivars is Malecot's (1948) coefficient of coancestry (f), which estimates the probability that two random alleles sampled from each of two individuals will be identical by descent. Although widely used, this measure is based on several assumptions that are typically violated by the germplasm used in plant breeding programs, namely (i) absence of selection, (ii) absence of genetic drift, and (iii) accurately recorded pedigree relationships. (Messmer et al., 1993). During the past decade, numerous studies have been conducted with maize to compare estimates of genetic divergence, diversity, and discrimination among lines based on coancestry and DNA or protein marker polymorphisms. Stuber and Goodman (1983) showed that isozyme variation could provide an accurate estimate of genetic distance, consistent with known pedigree, for 31 maize lines. Lee et al. (1989) and Melchinger et al. (1990) reported that patterns of variability in RFLPs provided measures of genetic distance which were consistent with known pedigree and were useful for assigning maize lines to heterotic groups. In an extensive study involving 29 European maize inbreds and RFLP markers from 188 probe-enzyme combinations, Messmer et al. (1993) reported that genetic similarities calculated from the variation in RFLP patterns reflected the true degree of relatedness more accurately than Malecot's f because the measures derived from RFLP markers were not dependent on the assumptions made by Malecot. The same authors also noted that adequate genome coverage with RFLP markers would be essential for accurate similarity estimates.

Although RFLPs and isozymes provide useful estimates of genetic divergence or diversity, there are some concerns as to their discriminatory power. Stuber and Goodman (1983) reported that polymorphism data from 23 isozyme loci would provide unique profiles for 73% of the 406 maize inbred lines they surveyed. Mumm and Dudley (1994) were able to assign 148 maize inbreds to the correct heterotic groups using variability in RFLP markers from 46 probe-enzyme combinations but were concerned that this number of RFLP markers might represent the lower limit for being able to discriminate such lines. Although increasing the number of marker loci may improve discrimination, this might not be feasible because of the few polymorphic isozyme loci available and the labor intensive and time consuming nature of RFLP analyses.

Simple sequence repeats (Jacob et al., 1991) are loci which are comprised of highly variable arrays of tandemly repeated, 2 to 6 base pair (bp) long DNA sequences. The SSR loci can be amplified by the PCR (Saiki et al., 1988) using primers which are complimentary to the regions flanking the repeats. The resulting products, separated electrophoretically, are highly polymorphic and provide codominant genetic markers with

M.L. Senior, USDA-ARS, Molecular Marker Facility, Box 7620, North Carolina State Univ., Raleigh, NC 27695, currently at Novartis Seeds, Inc., 3054 Cornwallis Rd., Research Triangle Park, NC 27709; J.P. Murphy, Dep. of Crop Science, Box 7629, North Carolina State Univ., Raleigh, NC 27695; M.M. Goodman, Dep. of Crop Science, Box 7620, North Carolina State Univ., Raleigh, NC 27695; and C.W. Stuber, USDA-ARS, Dep. of Genetics, Box 7614, North Carolina State Univ., Raleigh, NC 27695. Received 7 Nov. 1997. *Corresponding author (lynn.senior@seeds.novartis.com).

Published in Crop Sci. 38:1088–1098 (1998).

Abbreviations: bp, base pair; PCA, Principal Components Analysis; PCR, polymerase chain reaction; PIC, polymorphism information content; QTL, quantitative trait loci; RFLP, restriction fragment length polymorphism; SSR, simple sequence repeat; UPGMA, Unweighted Paired Group Method Using Arithmetic Averages.

Mendelian inheritance (Beckmann and Soller, 1990). Simple sequence repeats occur in many plant genomes including those of maize (Senior and Heun, 1993; Shattuck-Eidens et al., 1990), soybean [*Glycine max* (L.) Merr.] (Akkaya et al., 1992), *Brassica* spp. (Poulsen et al., 1993; Kresovich et al., 1995), rice (*Oryza sativa* L.) (Wu and Tanksley, 1993; Zhao and Kochert, 1993), and barley (*Hordeum vulgare* L.) (Saghai-Maroof et al., 1994). The level of polymorphism in plant species studied thus far has been greater than that found with RFLPs (Wu and Tanksley, 1993; Saghai Maroof et al., 1994; and Rongwen et al., 1995). Several efforts are underway to utilize SSR markers to study genetic diversity, identify germplasm and characterize population dynamics (Thomas and Scott, 1993; Yang et al., 1994; and Rongwen et al., 1995).

Smith et al. (1997) reported a classification of 58 maize inbred lines based on variability at 131 SSR loci. Primers for PCR amplification were fluorescently labeled so that alleles could be detected in a polyacrylamide gel using an ABI Model 373A automative DNA sequencer equipped with GENESCAN 672 software. This system scores the gels automatically and can distinguish sequences that differ by one bp. Patterns of variability at the SSR loci were compared with those obtained using RFLPs from 80 single locus probe/enzyme combinations. The values for polymorphism information content (PIC), a measure of allele diversity, were similar for both RFLP and SSR loci, although nine SSR loci exceeded the maximum PIC value obtained for RFLPs. Patterns of variability at both the RFLP and the SSR loci closely approximated the genetic divergence estimates among inbreds that were determined by pedigree, with $r = 0.80$ and $r = 0.81$ respectively. Smith et al. (1997) concluded that because SSR variability could be scored more accurately and reproducibly than RFLPs, and because several of the SSR loci had very high PIC values, SSR loci would provide an equal or greater power of discrimination in a more cost effective manner.

Although the SSR system described previously may be a valuable alternative to RFLP and isozymes markers for measuring genetic divergence, an additional evaluation of SSR loci using the less costly and more widely available agarose gel system for detecting polymorphisms (Senior and Heun, 1993) would be desirable. In addition, a larger and more diverse selection of elite maize inbreds should be analyzed to obtain a better sample of the alleles present within U.S. maize breeding germplasm. The goals of the study were (i) to compute allele sizes at 70 SSR loci for each of 94 diverse maize inbreds as resolved via agarose gels, (ii) to estimate genetic similarities and relationships among the 94 inbreds based on the allelic variability detected, and (iii) to identify alleles exclusive to specific germplasm sources or heterotic groups.

MATERIALS AND METHODS

Plant Material

Ninety-four elite maize inbreds were chosen to represent diversity in the U.S. maize germplasm collection. Selections were made from Reid and Lancaster heterotic groups, tropical and temperate types, and from several endosperm types including dent, flint, sweet corn, and popcorn. Information (when available) regarding pedigree, endosperm type, and state or country of origin is listed in Table 1. Seed was obtained from M.M. Goodman's collection at North Carolina State University, Raleigh, NC. DNA was extracted from a bulk of five plants from each line using a modified CTAB procedure (Saghai-Maroof et al., 1984).

Simple Sequence Repeat Primer Selection

All 120 SSR primers from the nc/phi primer set (Research Genetics, Huntsville, AL) were assayed using the sample of 94 inbreds. Primers were excluded from the study if banding patterns were difficult to score accurately on agarose gels or if the primers failed to amplify consistently in all 94 inbred lines. A final set of 70 SSR primers was chosen for further analysis. Information regarding map position and repeat type for each of the SSRs used can be found in Table 2.

Amplification and Detection Conditions

The PCR reactions were performed in 96-well microtiter plates using a PTC-100 Thermal Cycler (MJ Research, Watertown, MA). Each plate contained 94 inbreds amplified with one SSR primer. The amplification consisted of a denaturation step of 1 min at 95°C followed by a "touchdown" profile as described by Mellersch and Sampson (1993). This profile began with 2 cycles of 1 min at 95°C, 1 min at 65°C, and 2 min at 72°C. The annealing temperature was then reduced by 1°C every 2 cycles until a final annealing temperature of 55°C was reached. The last cycle was repeated 20 times and was terminated with a continuous cycle at 4°C. The 15 μL reaction mix consisted of 25 ng of each primer, 1 unit *Taq* DNA polymerase (Boehringer Mannheim, Indianapolis, IN), 100 μ*M* each dNTP, 1× reaction buffer (10 m*M* Tris-HCl, 1.5 m*M* $MgCl_2$, 50 m*M* KCl 100 μg mL^{-1} gelatin; pH 8.3), 10× nonacetylated BSA (New England Biolabs, Beverly, MA), 50 ng template DNA and distilled, deionized H_2O. The reaction mix was overlaid with 50 μL of mineral oil. Reactions were stopped with 10 μL loading-dye (70% glycerol, 20 m*M* EDTA, 0.2% SDS, 0.6 mg mL^{-1} bromphenol blue). After amplification, a total of 24 μL of the mix was loaded into the wells of a 1-mm-wide comb and products were separated by electrophoresis in a Model A3 20 by 40 cm horizontal gel system (Owl Scientific, Cambridge, MA) using 1 × TBE (Sambrook et al., 1989) on a 4% METAPHOR (FMC Bioproducts, Rockland, ME) agarose gel containing 0.15 μg mL^{-1} ethidium bromide. A 50-bp ladder (Gibco/BRL, GIBCO Laboratories, Grand Island, NY) was loaded into lanes 1, 12, 24, and 36. There were three 30 tooth combs per gel, spaced 12 cm apart. All 94 inbreds amplified by a single SSR were run on the same gel, divided evenly among the three combs. The gels ran at ≈150 V for 6 h.

Gel Scoring

Gels were divided into three sections, each section containing the samples loaded into one each of the three combs. The gel sections were scored individually with an Eagle-eye II Still Video System (Stratagene, La Jolla, CA) equipped with RFLPscan Version 2.1 Gel Analysis Software (Scanalytics, Billerica, MA). Standardization lines were drawn by the RFLPscan software using the 50-bp ladders as standards. Allele sizes were determined based on their position relative to the standardization lines. Data were entered directly into an Excel spreadsheet. PCR reactions and gels were then rerun so that inbred lines which shared similar alleles would occur in adjacent lanes. This allowed for confirmation of scoring

Table 1. Pedigree, origin, market class and cluster location for the 94 inbred lines. Pedigree information was obtained from Gerdes et al., 1993.

Inbred	Pedigree/origin	State released by	Endosperm type	Cluster
A	Funk yellow dent	Illinois	Yellow dent	A
A12	Minnesota #13	Minnesota	Yellow dent	B
A188	[(4-29 × 64)4-29 ^ 4]	Minnesota	White dent	G
	4-29 = Silver King, 64 = Northwestern Dent			
A554	(Wf9 × WD)WD ^ 2	Minnesota	Yellow dent	D
	Wf9 = Wilson Farm Reid, WD = Wisconsin #25			
A619	(A171 × Oh43)Oh43	Minnesota	Yellow dent	A
	A171 = Complex cross, Oh43 = Oh40B × W8, W8 = Minnesota#13 × Ill.A48)			
A632	(Mt42 × B14)B14 ^ 3	Minnesota	Yellow dent	E
	Mt42 = Minnesota#13, B14 = BSSS			
A634	(Mt42 × B14)B14@ B14 ^ 2 @ = self	Minnesota	Yellow dent	E
	Mt42 = Minnesota#13, B14 = BSSS			
A635	(ND203 × B14)B14 ^ 2	Minnesota	Yellow dent	E
	ND203 = Haneys Minnesota#13, B14 = BSSS			
A641	(ND203 × B14)	Minnesota	Yellow dent	E
	ND203 = Haneys Minnesota#13, B14 = BSSS			
B2	Reid Yellow Dent	Missouri	Yellow dent	F
B14A	(Cuzco × B14 ^ 8) Rust resistant selection	Iowa	Yellow dent	E
B37	Iowa Stiff Stalk Synthetic	Iowa	Yellow dent	C
B52	Segregating material from a private breeder	Iowa	Yellow dent	D
B68	41.2504B × B14 3	Iowa	Yellow dent	E
B73	Iowa Stiff Stalk Synthetic C5	Iowa	Yellow dent	E
	BSSS = A3G-3-1-3, C.I.187-2, C.I.540, C.I.617, F1B-1-7-1, Hy.1159, 1224, Ill.12E, IN461-5, INAH83, INTr9-1-1-6, LE23, Oh3167B, Os420, WD4576			
B76	(C.I.31A × B37)B37	Iowa	Yellow dent	C
	C.I. 31A = Midland O.P., B37 = BSSS			
B77	BS11 (Pioneer 2-Ear Composite)	Iowa	Yellow dent	F
	Southern prolific germplasm crossed with Corn Belt Lines			
B84	BS13 (S2) CO	Iowa	Yellow dent	E
	BSSS			
B97	BSCB1 (following 9 cycles of RRS w/BSSS)	Iowa	Yellow dent	F
B164	Indiana Reid (developed by Pioneer)	Minnesota	Yellow dent	C
C11	Minn. #13	Minnesota	Yellow dent	A
C49A	Minn. #13	Minnesota	Yellow dent	G
C103	Lancaster Surecrop (from Noah Hershey)	Connecticut	Yellow dent	B
C123	C102 × C103	Connecticut	Yellow dent	B
	C102 = Lancaster Surecrop (from Noah Hershey)			
C.I.187-2	Krug [Reid Yellow Dent (IA and Ill strains) and Iowa Gold Mine]	USDA-ARS	Yellow dent	B
CM37	KE3	Manitoba, CAN	Yellow dent	D
CM105	V3 × B14 ^ 2	Manitoba, CAN	Yellow dent	E
CMV3	A21 × W185, A21 = Golden Gate, W185 = Minnesota#13	Manitoba, CAN	Yellow dent	A
D940Y	[BO60W(A166N × B556Y-B560Y)]B577Y			
DE811	[B68 × [B37Ht × (C103 × Mp3204 double cross)Sel.]]	Delaware	Yellow dent	E
	Mp3204 = (Mp448 × T204)(Mp424 × GT112) = Tropical			
	B68 = 41.2504B × B14 ^ 3, B37 = BSSS, C103 = Lancaster Sure Crop			
E2558W	M162W ^ 3.N6	South Africa	White dent	G
EP1	Spanish population 'Lizargarate'	Mision biologica de Galicia	Yellow flint	D
F2	French OP population 'Lacaune'	INRA-Peronne, France	Yellow flint	H
F44	Smith (Old Florida Variety)	Florida	Yellow dent	H
F2834T	Teko Yellow	South Africa	Yellow dent	
GT112	Multiple cross (Inc. Whatley, Cuban, Garrick, Creole +12% other)	Georgia	Yellow dent	
H95	Ob43 × C.I. 90A	Indiana	Yellow dent	A
	Oh43 = Oh40B × W8, W8 = M#13 × Ill. A48(Funk Yellow Dent)			
	C.I. 90A = L97 × M14, M14 = BR10 × R8), L97 = Tuxpan			
H99	Illinois synthetic 60C	Indiana	Yellow dent	A
	USDA Blight Res. Double Double × B8			
HP301	Supergold	Indiana	Popcorn	I
Hy	Illinois High Yield	Illinois	Yellow dent	A
I29	Unknown (White Kernel)	Iowa	Popcorn	I
1137TN	Variety Cross-Natal Yellow Horsetooth × Teko yellow (from TN)	South Africa	Yellow dent	
I205	Iodent (Reid Yellow dent)	Iowa	Yellow dent	A
IA2132	(TSR × 45) × 4329 (45 is from Bantam Evergreen)	Iowa	Sweet corn	H
IDS28	Yellow Pearl	Iowa	Popcorn	I
IDS91	South American population	Iowa	Popcorn	I
IL677a	[(Bolivia 1035 × IL44b) × IL422a]	Illinois	Sweet corn	H
K55	Pride of Saline	Kansas	White dent	G
K64	Pride of Saline	Kansas	White dent	B
Ki 14	Suwan1(S) C4, Suwan1 = Yellow Flint from Thailand	Hawaii	Yellow Flint	G
Ky21	Boone County White = White Mastodon	Kentucky	White dent	C
Ky226	NC.I.aDDC × Coahuila 8	Kentucky	White dent	A
Ky228	Pride of Saline	Kentucky	White dent	B
L317	Lancaster Surecrop	Iowa	Yellow dent	H
M14	BR10 × R8	Illinois	Yellow dent	A

continued on next page

Table 1. continued.

Inbred	Pedigree/origin	State released by	Endosperm type	Cluster
	BR10 (Funk Yellow Dent), R8 (Tx Surecropper)			
M37W	21A ^ 2 Jellicorse	South Africa	White dent	G
Mo17	C.I.187-2 × C103	Missouri	Yellow dent	B
Mt42	Minn.#13 (Owen's)	Minnesota	Yellow dent	D
N28(Ht)	Stiff Stalk Synthetic 1	Nebraska	Yellow dent	
NC258	TZ ^ 2 × [(NC248 × NC246) × C103] (TZ = Coker811A × C103 ^ 4)	North Carolina	Yellow flint	B
NC268	(B73 × NC250) × B73	North Carolina	Yellow flint	E
NC296	Pioneer X105A × H5	North Carolina	White dent	G
NC298	Line from Agroceres 155 × Line from (Pioneer X105A × H5)	North Carolina	Yellow flint	G
NC300	Line from Pioneer X105A × Line from (Pioneer X306B × H5)	North Carolina		G
	Line from ear to row selfing from F2 of (Pioneer X105A × H5)			
NC304	× Line from H101	North Carolina	Yellow flint	G
Oh07B	(Oh07 × 38-11)Oh07	Ohio	Yellow dent	C
Oh40B	8-line composite of Lancaster Surecrop lines	Ohio	Yellow dent	A
Oh43	Oh40B × W8	Ohio	Yellow dent	A
	W8 = M#13 × Ill.A48(Funk Yellow Dent)			
Os420	Osterland Yellow Dent, a strain of Reid Yellow Dent	Iowa	Yellow dent	B
P39	Purdue Bantam; Golden Bantam	Indiana	Sweet corn	H
Pa91	(Wf9 × Oh40B)S4 × [(38-11 × L317)38-11]S4	Pennsylvania	Yellow dent	C
R4	Funk Yellow Dent	Illinois	Yellow dent	A
R177	Snelling Corn Borer Synthetic	Illinois	Yellow dent	A
SA24	South American	Indiana	Popcorn	I
SC76	Hastings Prolific × Yellow Tuxpan	South Carolina	Yellow dent	G
SC213	(GT112 × NC33)GT112	South Carolina	Yellow dent	G
SG1533	Supergold	Indiana	Popcorn	I
T8	Jarvis Golden Prolific (developed by NC farmer, J.M. Jarvis)	Tennessee	Yellow dent	B
T232	Jellicorse × Teko Yellow, Jellicorse developed by R. Jellicorse	Tennessee	Yellow dent	G
Tx303	Yellow Surecropper	Texas	Yellow dent	H
Tx601	Yellow Tuxpan	Texas	Yellow dent	G
U267Y	Mex.155 ^ 3.Wf9r	South Africa	Yellow dent	G
Va22	Va17 × C103 ^ 2	Virginia	Yellow dent	B
Va35	(C103 × T8)T8	Virginia	Yellow dent	B
Va59	(C103 × T8 ^ 2) × (K4 × C103 ^ 2), K4 = Kansas sunflower	Virginia	Yellow dent	B
Va99	Oh07B × Pa91	Virginia	Yellow dent	C
Va102	Va59 × Va60	Virginia	Yellow dent	B
W64A	Wf9 × C.I.187-2	Wisconsin	Yellow dent	C
W117Ht	Ht1 conversion of W117 [W117 = 643 × Minn.#13]	Wisconsin	Yellow dent	D
W153R	(Ia153 × W8)Ia153	Wisconsin	Yellow dent	A
	Ia153 = US selection #133. W8 = Minn #13 × Ill.A48			
W182B	WD × W22	Wisconsin	Yellow dent	C
	WD = Wisc#25, W22 = Ill. B10(Funk Yellow Dent) × Wisc #25			
Wf9	Wilson Farm Reid	Indiana	Yellow dent	C
38-11	Outcross in line from 176A	Indiana	Yellow dent	C

and for standardization among inbreds scored on separate sections of the gel.

Data Analysis

Polymorphism information content (PIC) for each SSR marker was determined as described in Smith et al. (1997). PIC is a measure of allele diversity at a locus and is equal to

$$1 - \sum_{i=1,n}^{n} f_i^2 \quad [1]$$

where f_i is the frequency of the ith allele. When calculated in this manner, PIC is synonymous with the term "gene diversity" as described by Weir (1996). The PIC values provide an estimate of the discriminatory power of a marker by taking into account not only the number of alleles at a locus, but also the relative frequencies of those alleles in the population under study. Marker loci with a large number of alleles occurring at equal frequencies will have the highest PIC values.

Genetic similarities (GS) were estimated from the allele size data using a simple matching coefficient, such that GS = $m/(m + n)$, where m = the number of matches and n = the number of mismatches (Sneath and Sokal, 1973). The 94 inbreds were clustered based on the matrix of genetic similarities using the Unweighted Pair Group Method Using Arithmetic Averages (UPGMA) clustering algorithm. As a second means of cluster analysis, data were converted into columns of "1's" and "0's", where a "1" indicated the presence of a specific allele (column) in an inbred line (row) and a "0" indicated its absence. A Principal Components Analysis (PCA) was then performed based on the variance–covariance matrix calculated from this transformed data. The principal component scores were not standardized and thus had variance equal to the corresponding eigenvalue. The UPGMA results were used to draw the dendrogram. Nine clusters were then delineated for further analysis based on results from both the PCA and the UPGMA algorithm. There were five or more inbreds per cluster. Allelic frequencies were obtained for each of the nine clusters and a second PCA was performed on the variance–covariance matrix of the allelic frequencies among clusters. The distance matrix and dendrogram were constructed with NTSYS-pc version 1.8 (Rohlf, 1992). The principal component analyses were performed with PROC PRINCOMP with the COVARIANCE option (SAS Institute, 1985).

RESULTS AND DISCUSSION

The 70 SSR primers have been mapped to regions that were well dispersed throughout the maize genome (Table 2) and produced 365 alleles among the 94 maize inbreds. The 365 alleles collectively yielded unique genotypes for each of the 94 inbreds. Simple sequence repeat genotypes for each inbred line are listed in the Maize Genome Database and can be accessed via the

Table 2. Allele numbers, size range (in base pairs), and PIC[1] value for SSR loci found in 94 inbred lines of maize. Bin numbers were determined from the 1995 UMC Maize Map (Coe et al., 1995).

SSR locus	Repeat type	Bin no.	No. alleles	Size range (bp)	PIC value†
phi056	GCC	1.01	4	84–93	0.69
phi097	TAG	1.01	3	94–100	0.36
phi001	AG	1.04	23	66–154	0.92
phi037‡	CT	1.08	9	130–230	0.83
phi011	GCT	1.1	2	110–122	0.66
phi055	GAA	1.1	6	103–118	0.75
phi094	GGA	1.1	2	90–126	0.28
phi120	AAG	1.11	5	63–88	0.56
phi083	CTAG	2.04	5	122–138	0.67
nc131‡	AC	2.04	5	131–147	0.61
nc132‡	AG	2.04	3	214–230	0.19
nc133	GTGTC	2.04	3	110–120	0.31
phi127	GTCT	2.07	5	112–138	0.71
phi090	ATATC	2.08	3	141–266	0.33
phi036	AG	3.02	14	63–119	0.83
nc030	CT	3.04	5	104–114	0.61
phi029	CCCT-CT	3.04	4	146–164	0.61
phi073	CAG	3.05	4	90–99	0.67
phi053	GTAT	3.05	5	170–194	0.67
phi046	GCAC	3.08	2	62–66	0.48
phi047	TAC	3.09	3	140–152	0.66
nc135	AG	4.01	4	112–124	0.58
phi072	AAAC	4.01	5	142–162	0.52
nc004‡	AG	4.02	10	140–188	0.68
phi021	AG	4.02	8	90–128	0.67
phi026	CT	4.04	11	76–126	0.94
phi074	CAA	4.04	3	89–95	0.59
phi079	CATCTG	4.04	4	180–195	0.55
phi096	GAGGT	4.05	2	102–112	0.5
phi006	CCT	4.1	5	84–96	0.85
phi019	ATT	4.1	4	93–102	0.73
phi076	GAGCGG	4.1	3	156–168	0.45
nc130	AGC	5	2	140–143	0.45
phi024	CCT	5	4	159–165	0.69
phi113	GTCT	5.02	3	120–284	0.64
phi085	GCGTT	5.06	5	70–95	0.57
phi101	TAC	5.06	3	92–98	0.63
phi058	GCC	5.07	2	148–151	0.35
phi128	AAGCG	5.07	4	100–120	0.51
phi075‡	CT	6	3	217–233	0.61
phi077	AG	6.01	9	122–154	0.76
phi126‡	AG	6.01	17	128–200	0.87
phi031	GTAC	6.03	4	187–227	0.57
nc013	CT	6.04	8	102–132	0.83
phi070	GAGCT	6.06	3	78–88	0.57
phi078	AAAG	6.07	5	122–166	0.6
phi081	GAT-TAC	6.07	3	160–171	0.39
phi112‡	AG	7.01	5	136–160	0.45
phi034	CCT	7.02	5	120–141	0.71
phi114	GCCT	7.02	4	135–166	0.69
phi082	AG	7.05	7	120–134	0.62
phi069	CCT	7.05	3	197–206	0.62
phi116	TGAC-GAC	7.06	7	154–173	0.77
phi115	TA-ATAC	8.03	3	93–113	0.53
phi119‡	CT	8.03	6	162–174	0.77
phi014	GGC	8.04	3	148–163	0.29
phi121	CCG	8.04	2	99–102	0.22
phi015	TTTG	8.08	5	78–102	0.59
phi080	AGGAG	8.08	7	140–170	0.69
phi028	GAA	9.01	4	63–78	0.63
phi068	AT	9.01	5	66–76	0.55
nc134‡	CT	9.03	2	166–172	0.5
phi027	GCGCT	9.03	4	142–157	0.54
phi061	TTCT-GTAT	9.03	3	80–88	0.52
phi065	CACTT	9.03	5	132–157	0.68
phi059	CCA	10.02	3	147–159	0.56
phi063	TATC	10.02	5	152–220	0.67
phi050	CTTG	10.03	2	80–84	0.17
phi084	GAA	10.04	4	148–157	0.49
phi035‡	AC	10.06	5	131–146	0.53

† PIC (Polymorphism Information Content) = $1 - \sum_{i=1,n}^{n} f_i^2$, where f_i is the frequency of the ith allele.
‡ SSR loci whose PIC values may be underestimated due to limitations in the resolution capacity of agarose gels. See text for explanation.

World Wide Web at http://www.agron.missouri.edu. The average number of alleles per locus was five, with a range from two to twenty-three. The PIC values for the SSR loci ranged from 0.17 to 0.92, with an average of 0.59 (Table 2). Simple sequence repeats comprised of dinucleotide repeats had the highest average PIC value of 0.67. These results were consistent with the results of Smith et al. (1997) who found an average PIC value for all SSR loci of 0.62 and an average PIC value for dinucleotide repeats of 0.70. The slightly higher val-

ues in Smith et al. (1997) probably resulted from their use of acrylamide gels for allele detection. Acrylamide gels can resolve nucleotide differences of one base pair, whereas 4% Metaphor gels can detect size differences (in bp) of ≈2% (FMC Bioproducts, Rockland, ME). The increased resolution of acrylamide gels over agarose gels could result in the detection of a larger number of alleles per locus on acrylamide. This is of particular concern for SSR loci containing dinucleotide repeats whose amplification products are in the 130 to 200 bp range, since sequences differing by 2 bp could not be resolved using agarose gels. Ten SSR loci in this study fall into this category (Table 2), and the number of alleles detected for these loci may be underestimated.

Several of the SSR primers amplified more than one band per inbred line, perhaps because of residual heterozygosity, contamination, or the amplification of similar sequences in two separate genomic regions. Duplication in maize is well documented (Helentjaris, 1995). The first two possibilities are likely where only one or two of the 94 inbreds were heterozygous for a particular primer pair. The last possibility is probable for three loci, phi011, phi055, and phi096 where 34, 40, and 11 of the lines consistently had two banded phenotypes. Primers for phi011 and phi055 were derived from sequences of *Globulin-1*, and phi096 was derived from the sequence of a 19 kDa zein protein (Senior et al., 1996). These storage proteins are encoded by members of large gene families which may contain duplicate sequences (Kriz et al., 1987; Kriz, 1989). Further genetic studies are needed to confirm whether the SSR primers were, in fact, amplifying more than one locus.

Within the sample of 94 inbred lines, 70.4% of the 365 alleles occurred at a frequency of 0.25 or less, indicating high allelic diversity. Only 2.5% of the alleles occurred at a frequency of 0.75 or greater. This allelic diversity was likely due to the high level of polymorphism of the markers, the large range of phenotypic (endosperm) classes present in the panel of inbreds, and the inclusion of several exotic inbreds (Table 1). Seventy alleles occurred in no more than two inbreds (Table 3). Fifty-four of these were exclusive to only one inbred. Most of the unique alleles occurred among the popcorn and tropical inbreds. Eight alleles were specific to either one or two popcorn lines and three alleles were specific to groups of three popcorn lines. Eighteen alleles were specific to either one or two tropical lines, three alleles were specific to groups of three tropical lines and one allele was specific to a group of four tropical lines. The presence of many unique alleles may be an indication of the relatively high rate of mutation in SSR loci (Henderson and Petes, 1992) and the inclusion of exotic inbreds. Such alleles are important as they may be diagnostic for particular inbred lines or for regions of the genome specific to a particular type of maize (i.e., popcorn).

The high rate of polymorphism of several of the SSR loci allowed for the selection of a group of five markers, whose alleles in combination, provided unique genotypes for all 94 inbreds. These loci were phi001, phi126, phi113, phi036, and phi021. If feasible, combinations of two or more primer pairs in a single PCR reaction mix could enable rapid fingerprinting of many inbreds.

Table 3. SSR alleles which occurred in no more than 2 of 94 maize inbreds.

Locus	Bin no.	Allele size	Inbreds allele is specific for
		(bp)	
phi097	1.01	94	EP1
phi001	1.04	74	I29
		100	Ky21
		104	Hy
		106	SG1533
		108	EP1, IDS28
		112	IDS91
		120	B52
		124	C103, Mo17
		136	I137TN
phi037	1.08	138	DE811, HP301
		230	I29
phi055	1.1	103	IL677A
		106	I29, NC304
phi120	1.11	74	SC213
		88	I29, NC304
nc131	2.04	131	P39
nc132	2.04	230	NC304
nc133	2.04	110	NC304
phi127	2.07	116	C103, L317
phi090	2.08	266	GT112, SC213
phi036	3.02	83	NC304, W117Ht
		95	NC300
		99	F2834T
		105	I29
		107	SG1533
		119	H95, NC296
nc030	3.04	106	CMV3
phi053	3.05	186	EP1
nc135	4.01	124	Tx303, Va102
phi072	4.01	158	NC296
		162	Tx303
nc004	4.02	166	L317
		172	B77
		184	F44, P39
		188	M37W
phi021	4.02	110	L317
		120	B77
phi085	5.06	80	M37W
		95	NC268
phi128	5.07	120	NC304
phi077	6.01	122	F44
		142	F2834T
		154	B2
phi126	6.01	128	I137TN, NC268
		158	F2834T
		162	SG1533
		172	Ky226
		176	U267Y
		188	P39
		192	Ia2132
		200	NC300
phi031	6.03	227	IDS28
nc013	6.04	132	B164, B52
phi078	6.07	150	Os420
		166	B84
phi034	7.02	132	I29
phi082	7.05	132	Tx601
phi116	7.06	158	Tx303
phi014	8.04	148	KI14
phi015	8.08	78	W64A
phi080	8.08	140	F2834T
		170	NC304
phi080	8.08	220	NC296
phi084	10.04	148	F2834T
		157	KI14
phi035	10.06	131	Hy, NC300
		142	R177

Genetic similarities (depicted in the Maize Genome Database) among the 94 lines ranged from a low of 0.21 between the pairs Hy vs. A188 and H95 vs. I29 up to 0.90 between B14A vs. A635. A635 is a cross between ND203 and B14, which was then backcrossed twice to B14. The UPGMA clustering algorithm grouped the inbreds into nine clusters, designated A through I (Fig.

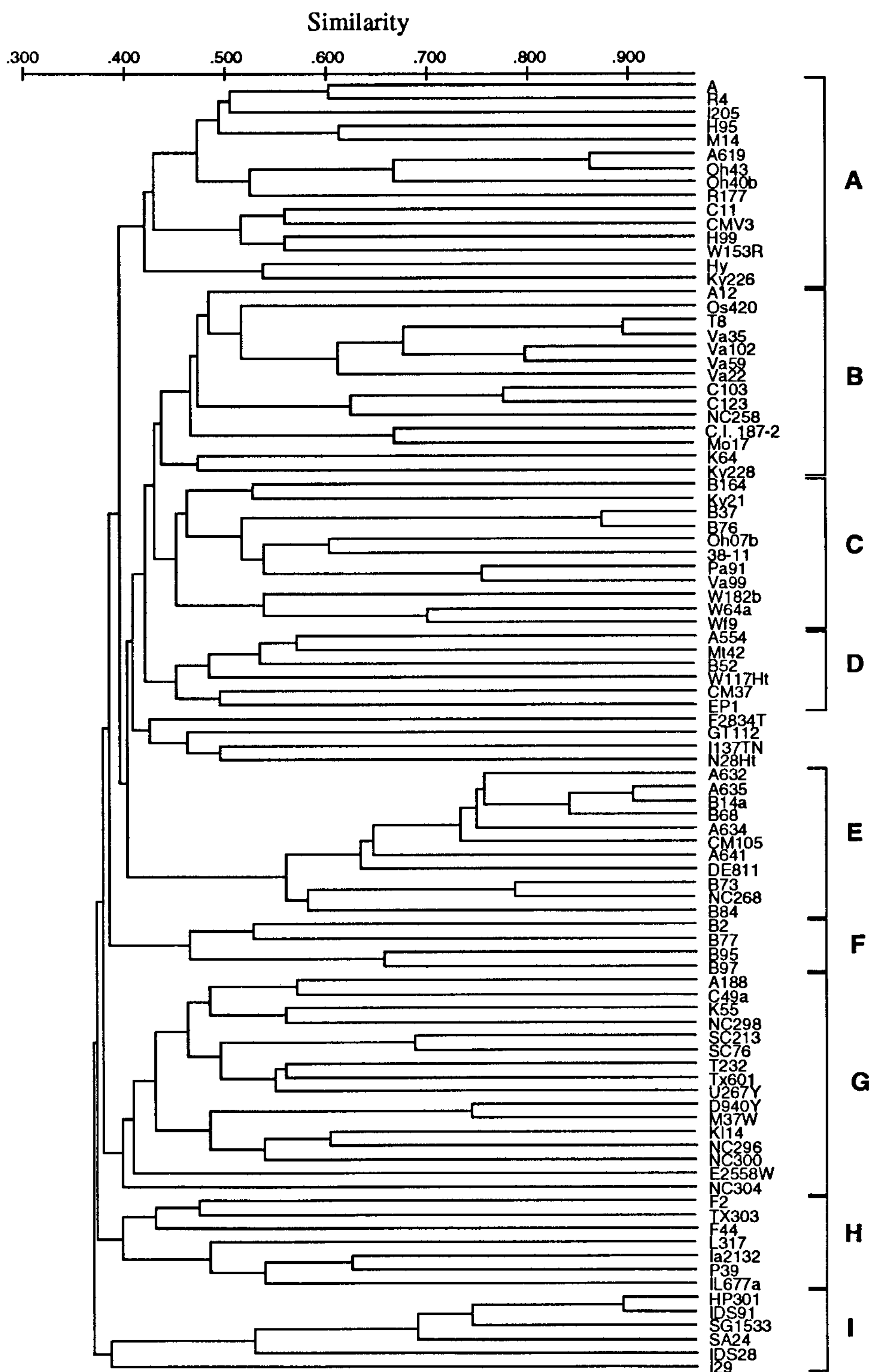

Fig. 1. Dendrogram constructed with a Unweighted Paired Group Method Using Arithmetic Averages (UPGMA) clustering algorithm from the pairwise matrix of genetic similarity among 94 maize inbred lines. Brackets to the right of the dendrogram indicate clusters used in a Principle Component Analysis (PCA).

1). The PCA revealed similar groupings (data not shown). Cluster A consisted mainly of the Lancaster lines Oh43 and Oh40B, both of which contain some Reid and Minnesota #13 germplasm, along with associated Reid and Minnesota #13 lines. Cluster B contained predominantly the Lancaster lines containing C103 germplasm. An additional grouping of Reid lines containing Oh07B and B37 comprised Cluster C. Pa91, a Lancaster line, also fell within this cluster, associated with Va99, which contains both Pa91 and Oh07B germplasm. Cluster D was a mixed group of mainly Northwestern Dent lines. A tight grouping of Iowa Stiff Stalk Synthetic (BSSS) lines, including B73 and B14, were found in Cluster E. Four additional lines, B2, B77, B95,

and B97 were grouped separately in Cluster F. B2 is a Reid line, B95 and B97 are both derived from the BCBS1 synthetic and are nonstiff stalk lines and B77 is derived from the BS11 synthetic. B77 is not related to either the BSSS or Lancaster lines, but combines well with BSSS lines (Lee et al., 1989). Cluster G comprised a large grouping of Tropical and Southern lines containing exotic germplasm. Cluster H contained, in part, the only sweet corn lines in the study, whereas all popcorn lines were exclusive to Cluster I.

Groupings of inbreds revealed by the present analysis generally agreed with the pedigrees of these lines and the clusters were representative of heterotic groups. Several discrepancies were present, such as the grouping of Reid and Lancaster lines in Cluster A and the grouping of sweet corn and Lancaster lines in cluster H. Such incongruities are not uncommon when comparing the results of molecular analyses with classifications based on pedigree. The incongruities may be the result of one or more of several factors outlined in Mumm and Dudley (1994). DNA markers, when identical in size, may represent alleles that are only *identical in state*, and may not always be *identical by descent*. In addition, DNA markers may be affected by selection, drift, and mutation (factors which are often ignored when estimating associations based on pedigree). Finally, incongruities can result from the clustering process whenever clusters are nonoverlapping. Because of the latter, an inbred line that is related to two other inbred lines from separate clusters will only be grouped with the one to which it is most closely related.

A PCA was performed on the covariance matrix calculated from the mean cluster allele frequencies of each major cluster. The first five principal components accounted for 21.1, 19.7, 15.8, 11.8, and 9.4% of the total variation, in that order. These five principal components partitioned Cluster E from Cluster F and I, Cluster F from Cluster I, Cluster H and D from Cluster E, F and I, Clusters A and H from Cluster D, and Clusters B and G from Cluster H, respectively. Heavily weighted characters were considered as those with eigenvector coefficients of greater than 0.125 or less than −0.125. Such characters and their frequencies within each cluster are shown in Table 4. Of the heavily weighted characters, those which were found to be predominant in only one cluster (e.g., phi065, allele 142 in Cluster E) occurred mainly in either Cluster E, H, or I, suggesting that these three clusters were formed in part because of the occurrence of many rare alleles. In general, the other six clusters resulted from significantly different frequencies of common, rather than rare alleles, an observation also noted by Mumm and Dudley (1994).

As expected, average allelic diversity within the clusters was slightly less than that in the overall population, with about twice as many alleles falling into the high frequency class (those which occur at frequencies of 0.75 or higher). Cluster I (containing only popcorn lines) had the least diversity with 8.5% of the high frequency alleles, compared with an average of 4.2% for the remaining clusters. This was expected because of the few lines and narrow range of germplasm in that cluster. The Lancaster lines in Cluster B appeared more diverse than the Stiff Stalk lines in Cluster E, as indicated by a slightly higher percentage of alleles at a low frequency (78% in Cluster B vs. 73% in Cluster E). Similar patterns of allelic variability were reported by Lee et al. (1989) and Godshalk et al. (1990).

Closer examination of two of the clusters revealed some noteworthy genetic relationships. Cluster B contained inbred lines predominantly from the Lancaster heterotic group, including Mo17, while Cluster E contained inbred lines predominantly from the Reid or Stiff Stalk heterotic group, including B73. The B73 and Mo17 inbred lines provide a historically important source of germplasm for U.S. maize hybrids because of the significant heterotic response that occurs whenever these two inbreds are crossed. Heterosis occurs, in part, when loci are in the heterozygous condition; therefore, alleles whose frequencies are significantly different between the two clusters may be more likely to be involved in the heterotic response (Falconer and MacKay, 1996). Forty-three of 365 alleles, from 25 loci, had frequencies that differed significantly (by frequency ≥ 0.50) between Cluster B and E (Table 5). In contrast with Mumm and Dudley's (1994) findings, alleles with pronounced differences between the BSSS and Lancaster clusters occurred throughout the genome, rather than being restricted to chromosomal regions 3S, 5S, 6L, 8S, 9S, and 10L. It is interesting to note that of the loci whose frequencies differed significantly between Clusters B and E, many of the loci that were polymorphic between B73 and Mo17 were located in regions where large quantitative trait loci (QTL) have been detected for yield heterosis in a study using B73 and Mo17 (Stuber et al., 1992; C.W. Stuber and J.R. LeDeaux, 1997, unpublished data). While many studies have failed to show a significant relationship between marker heterozygosity and the prediction of a heterotic response (Melchinger et al., 1990, Godshalk et al. 1990, Dudley et al., 1992), most authors suggest that selection of markers for this purpose should be restricted to those located in regions of potential effect. Dudley et al. (1992) note that although the lines evaluated for heterotic response may contain a number of unique alleles, only a few may be linked to favorable QTLs. Predicting heterotic response through genetic polymorphisms would be more effective if such markers were associated with loci that positively affect heterosis. Perhaps those markers responsible for distinguishing heterotic groups are linked to loci that are responsible, in part, for heterosis. This hypothesis merits further testing.

CONCLUSIONS

Ninety-four elite inbred lines of maize were analyzed via agarose gel electrophoresis for variability at 70 SSR loci. 365 alleles were detected. The allelic frequency data were used to estimate genetic similarities among the lines and the genetic diversity in the sample. A UPGMA clustering algorithm grouped the lines into nine clusters that corresponded to major maize heterotic groups or endosperm types. A PCA of the covariance

Table 4. SSR alleles with high eigenvector coefficients in a principal component analysis performed on the matrix of within cluster allele frequencies. A threshold eigenvector coefficient of >0.125 or <-0.125 was used to select the alleles.

			Principal component		Allele frequency in clusters								
Locus	Allele	Bin no.	No.	Eigenvector coefficient	A	B	C	D	E	F	G	H	I
phi056*	84	1.01	3	−0.173	0.333	0.071	0.000	0.000	0.636	0.250	0.000	0.143	0.833
	97	1.01	3	0.198	0.667	1.000	1.000	0.833	0.364	0.500	0.938	1.000	0.167
	100	1.01	3	−0.217	0.333	0.000	0.000	0000	0.636	0.500	0.063	0.000	0.833
phi001*	68	1.04	4	0.151	0.267	0.071	0.000	0.000	0.000	0.000	0.188	0.714	0.000
phi037*	134	1.08	1	0.134	0.333	0.071	0.000	0.000	0.727	0.000	0.063	0.143	0.000
	150	1.08	4	−0.181	0.200	0.714	0.091	0.833	0.000	0.250	0.375	0.000	0.000
phi011*	110	1.1	1	−0.130	0.200	0.143	0.455	0.333	0.000	0.750	0.250	0.429	0.833
	122	1.1	4	−0.173	0.067	0.714	0.364	0.667	0.727	0.250	0.125	0.143	0.167
	122,110	1.1	4	0.156	0.733	0.143	0.182	0.000	0.273	0.000	0.625	0.429	0.000
nc131*	139	2.04	1,3	0.146	0.600	0.143	0.727	0.833	0.909	0.000	0.750	0.714	0.167
nc133*	115	2.04	4	0.173	0.333	0.000	0.000	0.000	0.000	0.000	0.375	0.714	0.167
	120	2.04	4	−0.176	0.667	1.000	1.000	1.000	1.000	1.000	0.563	0.286	0.833
phi083*	122	2.04	1	0.133	0.000	0.143	0.273	0.000	0.727	0.000	0.063	0.000	0.000
	130	2.04	5	0.146	0.467	0.286	0.364	0.833	0.000	0.500	0.000	0.571	0.667
phi127	128	2.07	4	−0.198	0.000	0.429	0.182	1.000	0.818	0.000	0.125	0.286	0.500
nc030*	104	3.04	2	0.141	0.533	0.714	0.455	0.500	0.818	0.750	0.688	0.000	0.000
	108	3.04	2	−0.164	0.333	0.286	0.091	0.000	0.182	0.000	0.125	0.429	1.000
	114	3.04	3,5	0.136	0.067	0.000	0.364	0.500	0.000	0.000	0.063	0.571	0.000
phi029*	146	3.04	2	0.136	0.533	0.643	0.364	0.500	0.818	0.750	0.688	0.000	0.000
	152	3.04	2	−0.160	0.400	0.286	0.182	0.000	0.182	0.000	0.125	0.429	1.000
	164	3.04	3	0.143	0.067	0.071	0.455	0.500	0.000	0.000	0.063	0.571	0.000
phi053*	170	3.05	5	−0.138	0.600	0.500	0.182	0.500	0.091	0.000	0.500	0.143	0.667
	194	3.05	5	0.212	0.200	0.143	0.818	0.333	0.273	0.750	0.188	0.857	0.333
phi046*	62	3.08	3	0.138	0.333	0.786	0.364	1.000	0.273	0.000	0.250	0.571	0.500
	66	3.08	3	−0.138	0.667	0.214	0.636	0.000	0.727	1.000	0.750	0.429	0.500
phi047*	143	3.09	5	−0.198	0.067	0.786	0.273	0.500	0.364	0.000	0.688	0.000	0.167
	152	3.09	5	0.132	0.467	0.143	0.364	0.167	0.545	0.250	0.000	0.714	0.167
nc135*	112	4.01	4	0.138	0.533	0.357	0.909	0.167	0.818	0.250	0.625	0.857	0.333
nc004*	150	4.02	1	0.132	0.267	0.571	1.000	0.833	0.818	0.000	0.563	0.143	0.000
	150	4.02	4	−0.156	0.267	0.571	1.000	0.833	0.818	0.000	0.563	0.143	0.000
phi021*	96	4.02	1	0.132	0.267	0.571	1.000	0.833	0.818	0.000	0.563	0.143	0.000
	96	4.02	4	−0.156	0.267	0.571	1.000	0.833	0.818	0.000	0.563	0.143	0.000
phi074*	89	4.04	4	0.196	1.000	0.286	0.364	0.167	0.636	0.000	0.375	0.857	0.500
	92	4.04	4	−0.167	0.000	0.500	0.636	0.833	0.364	0.250	0.438	0.143	0.500
phi096*	102	4.05	4,5	0.168	0.867	0.500	0.545	0.167	0.818	0.000	0.938	0.714	0.667
	112	4.05	5	0.181	0.133	0.071	0.364	0.667	0.091	1.000	0.063	0.286	0.333
phi024*	156	5	5	0.207	0.133	0.143	0.545	0.667	0.818	0.750	0.063	0.429	0.667
phi024*	162	5	5	−0.161	0.467	0.500	0.455	0.000	0.000	0.250	0.500	0.000	0.000
phi101*	98	5.06	2	0.126	0.467	0.429	0.455	0.500	0.455	0.750	0.813	0.286	0.000
phi075*	217	6	2	0.132	0.333	0.071	0.727	0.000	0.091	0.750	0.125	0.286	0.000
	233	6	2,4,5	−0.182	0.333	0.643	0.273	1.000	0.909	0.000	0.188	0.429	1.000
nc013*	102	6.04	2	−0.152	0.067	0.071	0.000	0.167	0.000	0.000	0.000	0.000	0.833
phi070*	83	6.06	2	−0.126	0.533	0.786	0.091	0.167	0.364	0.000	0.313	0.429	0.833
phi081*	160	6.07	3	−0.165	1.000	0.571	0.818	0.167	0.727	1.000	0.875	0.286	1.000
	166	6.07	3	0.139	0.000	0.429	0.091	0.667	0.273	0.000	0.063	0.714	0.000
phi034*	120	7.02	1	0.171	0.133	0.143	0.182	0.000	0.909	0.000	0.250	0.571	0.000
phi114*	135	7.02	3	−0.157	0.600	0.786	0.182	0.333	1.000	0.500	0.125	0.000	0.667
phi069*	200	7.05	3	−0.164	0.133	0.000	0.091	0.167	0.909	0.250	0.375	0.000	0.667
	206	7.05	3	0.132	0.267	0.000	0.364	0.667	0.091	0.000	0.188	0.571	0.167
phi082*	124	7.05	2	0.136	0.600	0.714	0.636	0.333	0.818	0.750	0.438	0.857	0.000
phi082*	128	7.05	2	−0.151	0.000	0.000	0.000	0.167	0.000	0.000	0.125	0.000	0.833
phi116*	162	7.06	2	−0.132	0.133	0.000	0.000	0.333	0.000	0.000	0.313	0.429	0.667
phi119*	168	8.03	2	−0.154	0.067	0.000	0.000	0.500	0.182	0.250	0.313	0.000	1.000
phi080*	160	8.08	5	0.136	0.267	0.071	0.182	0.167	0.818	0.250	0.250	0.857	0.667
phi028*	72	9.01	2	0.144	0.333	0.786	0.545	0.333	1.000	1.000	0.188	0.143	0.167
	78	9.01	2	−0.140	0.200	0.143	0.364	0.333	0.000	0.000	0.625	0.571	0.833
phi027*	142	9.03	1	−0.147	0.867	0.786	0.636	0.667	0.000	0.750	0.688	0.286	0.833
	147	9.03	1	0.187	0.067	0.214	0.364	0.167	1.000	0.000	0.063	0.571	0.000
phi061*	80	9.03	2	−0.165	0.400	0.500	0.364	0.833	1.000	0.000	0.375	0.429	1.000
	88	9.03	2	−0.158	0.600	0.500	0.636	0.167	0.000	1.000	0.375	0.714	0.000
phi065*	132	9.03	1	−0.157	0.267	0.571	0.182	0.667	0.000	0.750	0.813	0.143	1.000
	142	9.03	1	0.151	0.067	0.071	0.000	0.000	0.818	0.000	0.000	0.000	0.000
phi059*	147	10.02	1	0.127	0.533	0.429	0.636	0.333	0.727	0.000	0.563	0.571	0.000
phi063*	168	10.02	1	0.140	0.267	0.429	0.545	0.000	0.909	0.250	0.438	0.571	0.000
	160	10.02	2	−0.180	0.400	0.429	0.091	0.500	0.091	0.000	0.313	0.286	1.000

matrix calculated from the mean allelic frequencies among clusters revealed alleles with large eigenvector coefficient values for the first five principal components, allowing for the possibility that a subset of the 365 alleles might suffice for assigning lines to specific clusters. Diagnostic profiles for each of the 94 inbreds could be obtained from the genotypes of five of the most polymorphic SSRs. The utility of PCR-based markers such as SSRs for measuring diversity, for assigning lines to heterotic groups, and for genetic fingerprinting equals or

Table 5. SSR loci and the alleles whose frequencies differed by greater than 0.50 between the BSSS heterotic group (E) and the Lancaster heterotic group (E). A "1" indicates the presence of an allele, a "0" indicates its absence.

Locus	Bin no.	Allele	Allele frequency Cluster B	Allele frequency Cluster E	C† 103 (B)	C 123 (B)	CI 187-2 (B)	Mo 17 (B)	Nc 258 (B)	T8 (B)	Va 102 (B)	Va 22 (B)	Va 35 (B)	Va 59 (B)	A 12 (B)	K 64 (B)	Ky 228 (B)	Os 420 (B)	A 632 (E)	A 634 (E)	A 635 (E)	A 641 (E)	B 14A (E)	B 68 (E)	B 73 (E)	B 84 (E)	CM 105 (E)	DE 811 (E)	Nc 268 (E)
phi056	1.01	84	0.071	0.636	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	1	1	1	0	0	0	1
		87	0.714	0.091	1	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
phi097	1.01	97	1.000	0.364	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	0
		100	0.000	0.636	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	0	0	0	1
phi037	1.08	134	0.071	0.727	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	1	0	0
		150	0.714	0.000	1	0	0	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
phi055	1.1	112,103	0.000	0.636	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	0	1	1	0	0
phi120	1.11	67	0.071	0.818	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	1
		70	0.857	0.182	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	1	1	0
phi083	2.04	122	0.143	0.727	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0	0	1	1	0
nc131	2.04	139	0.143	0.909	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1
		147	0.571	0.000	1	1	0	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
phi073	3.05	90	0.500	0.000	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		93	0.143	0.727	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	1	0	1
phi046	3.08	62	0.786	0.273	1	1	1	1	0	1	1	1	1	1	1	0	0	1	0	1	0	0	0	0	1	0	0	0	1
		66	0.214	0.727	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	0	1	1	1	1	0	1	1	1	0
nc135	4.01	112	0.357	0.818	1	1	0	0	0	1	0	1	0	0	0	0	1	0	1	1	1	1	1	0	1	1	1	0	1
phi026	4.04	80	0.071	0.636	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1	0	1	1	0
phi079	4.04	180	0.071	0.727	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0	1	1	0
		195	0.857	0.182	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1
phi024	5	156	0.143	0.818	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1	1	1	1	1	1	1
		162	0.500	0.000	1	1	0	1	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
phi126	6.01	164	0.000	0.636	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1	0	1	0	0
phi031	6.03	223	0.214	0.727	0	0	0	0	0	1	0	1	1	0	0	0	0	0	1	0	1	1	1	0	1	1	1	0	1
phi034	7.02	120	0.143	0.909	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	1
		138	0.714	0.091	0	0	1	1	1	1	0	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0
phi069	7.05	197	1.000	0.000	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
		200	0.000	0.909	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1
phi121	8.04	99	1.000	0.364	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	1
		102	0.000	0.636	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0
phi080	8.08	160	0.071	0.818	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	1	1	0	1	1	1	0
nc134	9.03	166	0.143	1.000	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1
		172	0.857	0.000	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
phi027	9.03	142	0.786	0.000	1	1	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
		147	0.214	1.000	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
phi061	9.03	80	0.500	1.000	0	0	0	0	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
		88	0.500	0.000	1	1	1	1	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
phi065	9.03	132	0.571	0.000	0	0	0	0	0	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0
		142	0.071	0.818	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	1	1	0
phi050	10.03	80	0.500	0.000	1	0	0	1	1	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
		84	0.500	1.000	0	1	1	0	0	1	0	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1
phi035	10.06	135	0.071	0.636	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	0	1	0	1
		139	0.929	0.273	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1	0	1	0	1	0

† Letter in parentheses denotes cluster.

exceeds that of the more time and labor-intensive RFLP markers and should prove valuable for a maize breeding program.

ACKNOWLEDGMENTS

The research was in part funded by the NC Biotechnology Center through Institutional Development Grant no. 9110-IDG-1022 and the USDA through NRICGP Grant no. 94-37300-0296 to C.W. Stuber and M.L. Senior. We wish to acknowledge D.C. Rhyne and D.Z. Beattie for their meticulous and patient help with generation of data, and to A.W. Dillard for preparation of leaf tissue samples.

REFERENCES

Akkaya, M.S., A.A. Bhagwat, and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139.

Beckmann, J.S., and M. Soller. 1990. Toward a unified approach to genetic mapping of eukaryotes based on sequence tagged microsatellite sites. Bio/Technology 8:930–932.

Coe, E., D. Hancock, S. Kowalewski, and M. Polacco. 1995. Gene list and working maps. Maize Genet. Coop. Newsl. 69:191–267.

Dudley, J.W., M.A. Saghai-Maroof, and G.K. Rufener. 1992. Molecular marker information and selection of parents in corn breeding programs. Crop Sci. 32:301–304.

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman Group, Ltd., Essex, England.

Gerdes, J.T., C.F. Behr, J.G. Coors, and W.F. Tracy. 1993. Compilation of North American maize breeding germplasm. CSSA, Madison, WI.

Godshalk, E.B., M. Lee, and K.R. Lamkey. 1990. Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. Theor. Appl. Genet. 80:273–280.

Goodman, M.M. 1994. Maize. p. 192–202. *In* N.W. Simmonds (ed.) Evolution of crop plants. Longman Group, Ltd., Essex, England.

Hallauer, A.R., W.A. Russell, and K.R. Lamkey. 1988. Corn breeding. p. 463–564. *In* G.F. Sprague and J.W. Dudley (ed.) Corn and corn improvement. 3rd ed. Agron. Monogr. 18. ASA, CSSA, and SSSA, Madison, WI.

Helentjaris, T. 1995. Atlas of duplicated sequences. Maize Genet. Coop. Newsl. 69:67–81.

Henderson, S.T., and T.D. Petes. 1992. Instability of simple sequence DNA in *Saccharomyces cerevisae*. Mol. Cell. Biol. 12:2749–2757.

Jacob, H.J., K. Lindpainter, S.E. Lincoln, K. Kusumi, R.K. Bunker, Yi-Pei Mao, D. Ganten, V.J. Dzau, and E.S. Lander. 1991. Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat. Cell l67:213–224.

Kresovich, S., A.K. Szewc-McFadden, S.M. Bliek, and J.R. McFerson. 1995. Abundance and characterization of simple-sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. Theor. Appl. Genet. 91:206–211.

Kriz, A.L. 1989. Characterization of embryo globulins encoded by maize GLB genes. Biochem. Genet. 27:239–252.

Kriz, A.L., R.S. Boston, and B.A. Larkins. 1987. Structural and transcriptional analysis of DNA sequences flanking genes that encode 19kD zeins. Molec. Gen. Genet. 207:90–98.

Lee, M. 1995. DNA markers and plant breeding programs. Adv. Agron. 35:265–344.

Lee, M., E.B. Godshalk, K.R. Lamkey, and W.W. Woodman. 1989. Association of restriction fragment length polymorphisms among maize inbreds with the agronomic performance of their crosses. Crop Sci. 29:1067–1071.

Malecot, G. 1948. "Les Mathematiques de l'Heredite." Maason et Cie, Paris.

Melchinger, A.E., M. Lee, K.R. Lamkey, A.R. Hallauer, and W.L. Woodman. 1990. Genetic diversity for restriction length polymorphisms and heterosis for two diallele sets of maize inbreds. Theor. Appl. Genet. 80:488–496.

Mellersch, C., and J. Sampson. 1993. Simplifying detection of microsatellite length polymorphisms. BioTechniques 15:582–584.

Messmer, M.M., A.E. Melchinger, R. Herrmann, and J. Boppenmaier. 1993. Relationships among early European maize inbreds: II. Comparison of pedigree and RFLP data. Crop Sci. 33:944–950.

Mumm, R.H., and J.W. Dudley. 1994. A classification of 148 U.S. maize inbreds: I. Cluster analysis based on RFLPs. Crop Sci. 34: 842–851.

Poulsen, G.B., G. Kahl, and K. Weissing. 1993. Abundance and polymorphism of simple repetitive DNA sequences in *Brassica napus* L. Theor. Appl. Genet. 85:994–1000.

Rohlf, J.F. 1992. NTSYS-pc: Numerical taxonomy and multivariate analysis system. Version 1.7 Exeter Software, Setauket, NY.

Rongwen, J., M.S. Akkaya, A.A. Bhagwat, U. Lavi, and P.B. Cregan. 1995. The use of microsatellite DNA markers for soybean genotype identification. Theor. Appl. Genet. 90:43–48.

Saghai-Maroof, M.A., K.M. Soliman, R. Jorgenson, and R.W. Allard. 1984. Ribosomal DNA spacer length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc. Natl. Acad. Sci. (USA) 81:8014–8018.

Saghai-Maroof, M.A., R.M. Biyashev, G.P. Yang, Q. Zhang, and R.W. Allard. 1994. Extraordinarily polymorphic microsatellite DNA in barley: Species diversity, chromosomal locations, and population dynamics. Proc. Natl. Acad. Sci. (USA) 91:5466–5470.

Saiki, R.K., D.H. Gelfand, S. Stoffel, S. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning: A laboratory manual. Cold Spring Harbor Lab. Press, NY.

SAS Institute. 1985. SAS user's guide: Statistics. 5th ed. SAS Inst., Cary, NC.

Senior, M.L., and M. Heun. 1993. Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. Genome 36:884–889.

Senior, M.L., E.C.L. Chin, M. Lee, J.S.C. Smith, and C.W. Stuber. 1996. Simple sequence repeat markers developed from maize sequences found in the Genbank database: Map construction. Crop Sci. 36:1676–1683.

Shattuck-Eidens, D.M., R.N. Bell, S. Neuhausen, and T. Helentjaris. 1990. DNA sequence variation within maize and melon: Observations from polymerase chain reaction amplification and direct sequencing. Genetics 126:207–217.

Smith, J.S.C., E.C.L. Chin, H. Shu, O.S. Smith, S.J. Wall, M.L. Senior, S.E. Mitchell, S. Kresovitch, and J. Ziegle. 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.

Sneath, P.H.A., and R.R. Sokal. 1973. Numerical taxonomy. W.H. Freeman & Co., San Francisco.

Stuber, C.W., and M.M. Goodman. 1983. Allozyme genotypes for popular and historically important inbred lines of corn, *Zea mays* L. USDA Agr. Res. Results, Southern. Ser., no. 16.

Stuber, C.W., S.E. Lincoln, S.W. Wolff, T. Helentjaris, and E.S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823–839.

Thomas, M.R., and W.S. Scott. 1993. Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites. Theor. Appl. Genet. 86:985–990.

Weir, B.S. 1996. Genetic data analysis II. 2nd ed. Sinauer Associates, Inc., Sunderland, MA.

Wu, K-S., and S.D. Tanksley. 1993. Abundance, polymorphism and genetic mapping of microsatellites in rice. Molec. Gen. Genet. 241:225–235.

Yang, G.P., M.A. Saghai-Maroof, C.G. Xu, Q. Zhang, and R.M. Biyashev. 1994. Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Molec. Gen. Genet. 245: 187–194.

Zhao, X., and G. Kochert. 1993. Phylogenetic distribution and genetic mapping of a (GCG)n microsatellite from rice (*Oryza sativa* L.). Plant Molec. Biol. 21:607–614.