Maydica 50 (2005): 163-170

USE OF SSR DATA TO DETERMINE RELATIONSHIPS AMONG EARLY MATURING IRANIAN MAIZE INBRED LINES

R. Choukan[1], M.L. Warburton[2,*]

[1] *Seed and Plant Improvement Institute, Shahid Fahmideh Blvd., 4119-31585, Karaj, Iran*
[2] *International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600 Mexico D.F., Mexico*

Received March 29, 2005

ABSTRACT - Information on germplasm diversity and relationships among elite materials is fundamentally important in hybrid crop improvement. Maize (*Zea mays* L.) is a crop of major economic importance in Iran, but most hybrids are developed in Iran by using inbreds extracted from hybrids introduced from other countries. Pedigree and genetic information is often lacking for these inbreds. The objective of this study was to determine the level of genetic diversity within and relationships among the most commonly used Iranian maize inbred lines using simple sequence repeat (SSR) markers. Thirty-six Iranian inbred lines and two lines from the International Maize and Wheat Improvement Center (CIMMYT) were fingerprinted with 43 SSR markers. A total of 194 alleles, moderate levels of diversity, and a high number of unique alleles were found in these lines. UPGMA clustering grouped the Iranian inbreds into four clusters and the two CIMMYT lines formed a separate and more distant cluster. Clustering was consistent with the known information about source materials. The highest distance was found between the CIMMYT lines and the cluster containing the Lancaster Sure Crop related lines. The genetic distance information may be used by breeders when planning future crosses among these inbred lines.

KEY WORDS: Maize (*Zea mays* L.); Iran; Simple Sequence Repeats.

INTRODUCTION

Despite its recent introduction, maize (*Zea mays* L.) is currently one of the major crops in Iran. Although both the area planted and yields of maize have dramatically increased in the last decade, it is anticipated that to meet future needs, a two-fold increase in area as well as a 30% increase in yield will be required by 2011. To achieve these gains will require an efficient breeding program with the ability to exploit heterosis in the germplasm; this in turn requires well-characterized germplasm with clearly defined heterotic groups. Maize breeding programs have been active in Iran since 1970, at which time the main activity was the introduction and evaluation of different hybrids from other countries. Inbred lines were extracted from hybrids from both known and unknown sources, resulting in new parents to produce hybrid combinations for Iran. These lines have been characterized in the field for per se and hybrid performance, and hybrid performance has been used to extrapolate information about relationships among the lines; however, limited pedigree and no marker data exist for these lines and there have been no systematic efforts to establish heterotic groups.

Information on germplasm diversity and relationships among elite materials is fundamentally important in crop improvement (HALLAUER and MIRANDA, 1988). In hybrid maize breeding programs, efficiency of procedures used to identify the lines to be crossed to develop outstanding single crosses as quickly as possible strongly affect the success of the program (HALLAUER and LOPEZ-PEREZ, 1979; HALLAUER, 1990). The best hybrid combinations can be identified using information from a variety of methods, including: field crosses (mainly using diallel or topcrosses to testers (HAN *et al.*, 1991; GONZALEZ *et al.*, 1997; TERRON *et al.*, 1997) but this method is slow and laborious); pedigree information (when available, which is often not the case with Iranian breeding materials); morphological traits (which may be greatly influenced by the environment); and molecular markers that detect variation at the DNA sequence level (SMITH and SMITH, 1992; BETRAN *et al.*, 2003; REIF *et al.*, 2003a).

* For correspondence (fax: +52 555 804 7567; e.mail: m.warburton@cgiar.org).

Molecular markers have been shown to be useful for assigning lines to heterotic groups in theory as well as practice (LEE *et al.*, 1989; LIVINI *et al.*, 1992; DUBREUIL *et al.*, 1996; MELCHINGER and GUMBER, 1998; YUAN *et al.*, 2001; REIF *et al.*, 2003b). PEJIC *et al.* (1998) compared different markers for their effectiveness in estimating genetic similarity among maize inbreds and showed that SSR markers revealed the highest level of polymorphism per single marker locus, due to their codominant nature and high number of alleles per locus, and may thus be used to discriminate even highly related germplasm. SSRs are easy to use and automate, are mapped to specific genomic location, and inexpensive (SENIOR and HEUN, 1993; DRINIC and KONSTANTINOV, 1998; KONSTANTINOV and DRINIC, 2000).

The objective of this study was to determine the level of genetic diversity within and relationships among the most commonly used Iranian maize inbred lines using simple sequence repeat (SSR) markers.

MATERIALS AND METHODS

Plant material and DNA extraction

A total of 36 early maturing Iranian maize inbred lines and two lines from CIMMYT, Mexico (CML051 and CML292, used as controls) were included in this study (Table 1). In a companion paper, the important medium- to late-maturing inbred lines from Iran are examined. The two sets of germplasm are analyzed separately, because they were bred separately in different breeding programs, and it was assumed they will neither be related, nor crossed in future breeding programs, and thus information on how they are related to each other is of limited use. The inbreds consisted of fourteen lines that were extracted from CIMMYT breeding materials that were found to be suitable for Iranian growing conditions, which are largely temperate; seven lines extracted from Yugoslavian germplasm; one line extracted from a hybrid from Belgium; two lines extracted from Hungarian hybrids; two sister lines from Polish source germplasm; one Iranian inbred of unknown origin; two lines extracted from a mixture of source germplasm; three sister lines related to Lancaster Sure Crop from the United States, and two introduced lines of unknown origin. Some of the inbred lines are closely related by pedigree, as can be seen in Table 1.

Equal quantities of leaf tissue (approximately one young, fully mature leaf) were collected from 10 seedlings per line at 4 weeks of age and bulked for DNA extraction. Extraction was done using a modified CTAB method (CLARK *et al.*, 1989) and a sap extractor (MEDV Enrich Pollahne, Am Weingarten Germany) according to the Applied Biotechnology Center's Manual of Laboratory Protocols (CIMMYT, 2001).

SSR assays

Forty-three SSR primers were assayed on the 38 inbred lines (Table 2). SSR markers were chosen from the MaizeGDB database (http://www.maizegdb.org/ssr.php) on the basis of bin loca-

TABLE 1 - *List and pedigree or origin (when known) of the early maturing Iranian inbred lines assayed in this study with 43 SSR markers. Lines with names containing the same letter and number combination before the backslash are closely related by pedigree (usually sister lines from the same parents).*

Inbred lines	Pedigree sources/origin*
Lines received or extracted from CIMMYT originated source materials	
K1264	Unknown CIMMYT source germplasm
K1264/4	Unknown CIMMYT source germplasm
K1264/5	Unknown CIMMYT source germplasm
K1263/1	Unknown CIMMYT source germplasm
K1263/2-1	Unknown CIMMYT source germplasm
K1263/17	Unknown CIMMYT source germplasm
K1271/6	Unknown CIMMYT source germplasm
K1396/4	Unknown CIMMYT source germplasm
K1249/3	Unknown CIMMYT source germplasm
K1298/1	Unknown CIMMYT source germplasm
K1254/8	Unknown CIMMYT source germplasm
K1604/5	Unknown CIMMYT source germplasm
KE72012/1	TL88B- 6233H Pool93
KE75039/3-1-1-1	[G29 MDRS]-13-23/Pool30
Lines received or extracted from Yugoslavian source materials	
K33	Unknown
TVA	Unknown
TVA926/1-22	Unknown
K1728/7	Unknown
K1728/8	Unknown
KE72011/1	ZPLB103
K58	Unknown
Lines extracted from hybrids	
KE75006/2-1-2-1	Spelender (Belgium)
KE76005/1-1-1-1-1-1-1	MV 484 (Hungary)
KE76009/1-7-1-1	Gazda (Hungary)
Other lines	
S61	O.P.Wigor (Poland)
S61/4-2	O.P.Wigor (Poland)
K615/1	Unknown Iranian local source
K2817/1	Unknown
K2816	Unknown
KE72008/1	Combination of 16 inbred lines
KE72009/1	Combination of 16 inbred lines
Lancaster Sure Crop related lines	
Oh43/1-411	Oh40 X WB
Oh43/1-42	Oh40 X WB
Oh43/1-42	Oh40 X WB
Lines from CIMMYT (Checks):	
CML051	POB79STAROSA8079-1-2-3-ff-#-#-#
CML292	(POB28XTSR)-33-2-7-1-2-BB-f-#-#

* The use of the number sign (#) in the pedigree of some CIMMYT germplasm indicates that seed was bulked in that generation of the pedigree.

tion (to maximize genomic coverage) while avoiding dinucleotide repeats (except phi112) because of the difficulty in accurately sizing alleles. Oligonucleotide primers were bought from commercial companies (Operon technologies, Inc., Alameda, Ca. or Sigma Aldrich, Mexico City) and forward primers were labeled at the 5′ end with either 6-carboxyflorescein (6-FAM) tetrachloro-6-carboxyflourescein (TET) or hexachro-6-carboxy flourescein (HEX).

Multiplexed PCR reactions were performed in 10 µl volumes containing 2 µl of template DNA (output of the sap extraction method diluted 5X with distilled, deionized H_2O), 0.4 to 4 pmols each of 1 to 4 primer pairs, 1X PCR buffer, 0.25 mM dNTPs, 1.5 to 2.5 mM $MgCl_2$ and 0.75 U Taq polymerase. The reactions were performed with a Peltier Thermal Cycler (MJ Research, Inc., Watertown, MA) using the amplification conditions of 94°C for 2 min, followed by 30 cycles of 94°C for 30s, X°C for 1 min, and 72°C for 1 min, followed by extension at 72°C for 5 min. X°C refers to the annealing temperature, which is specified for each primer in Table 2.

TABLE 2 - *Information on the 43 SSR loci used in this study, including name, repeat unit, bin location (Bin No.), annealing temperature (Ann. Temp.) in degrees Celsius, number of alleles, observed size range (measured in base pairs), gene diversity values, and PIC values in the 38 inbred lines measured.*

SSR locus	Repeat unit	Bin no.	Ann. Temp.	Allele No.	Size Range (bp)	Gene diversity	PIC value
phi339017	AGG	1.03	52	4	148-169	0.39	0.36
phi308707	AGC	1.1	56	6	119-134	0.76	0.72
phi064	ATCC	1.11	56	10	69-105	0.85	0.84
phi083	AGCT	2.04	52	3	125-133	0.6	0.52
nc133	GTGTC	2.05	54	8	105-160	0.7	0.66
phi374118	ACC	3.02	54	6	214-232	0.68	0.63
phi102228	AAGC	3.06	54	4	119-131	0.53	0.49
umc1399	CTAG	3.07	54	4	115-127	0.64	0.59
phi046	ACGC	3.08	54	2	62-66	0.36	0.3
umc1109	ACG	4.1	54	3	110-116	0.22	0.2
phi076	AGCGGG	4.11	60	3	161-173	0.66	0.59
nc130	AGC	5	54	3	139-145	0.48	0.39
phi109188	AAAG	5.03	54	7	156-204	0.67	0.64
phi331888	AAG	5.04	58	5	124-136	0.69	0.65
umc1153	TCA	5.09	54	4	105-114	0.67	0.62
umc1143	AAAAT	6	54	6	64-89	0.77	0.73
phi423796	AGATG	6.01	54	3	126-141	0.16	0.15
phi031	GTAG	6.04	56	4	187-227	0.63	0.59
phi452693	AGCC	6.04	52	4	125-137	0.59	0.54
phi123	AAAG	6.07	54	2	146-150	0.5	0.38
phi299852	AGC	6.07	58	8	111-144	0.79	0.76
phi089	ATGC	6.08	54	2	87-95	0.4	0.32
umc1545	AAGA	7	54	4	70-98	0.57	0.47
phi112	AG	7.01	56	3	153-159	0.17	0.16
phi328175	AGG	7.04	54	9	100-133	0.81	0.78
phi069	GAC	7.05	58	6	193-208	0.69	0.65
umc1304	TCGA	8.02	54	3	129-137	0.47	0.42
phi100175	AAGC	8.03	54	4	129-141	0.71	0.66
phi121	CCG	8.03	56	3	97-106	0.17	0.16
phi233376	CCG	8.09	54	6	136-154	0.8	0.77
umc1279	CCT	9	54	4	98-101	0.6	0.54
phi065	CACTT	9.03	54	5	126-151	0.62	0.56
phi032	AAAG	9.04	56	3	233-245	0.53	0.41
phi059	ACC	10.02	60	5	147-162	0.61	0.53
phi96342	ATCC	10.02	54	4	238-250	0.59	0.52
umc1152	ATAG	10.02	54	5	155-175	0.57	0.53
phi050	AAGC	10.03	56	5	72-104	0.48	0.43
phi062	ACG	10.04	56	4	155-164	0.59	0.51
phi084	GAA	10.04	54	5	153-165	0.72	0.68
umc1061	TCG	10.06	52	3	101-107	0.52	0.41
umc1196	CACACG	10.07	52	5	143-167	0.76	0.72
phi109642	ACGG	2.03-2.04	54	3	136-144	0.61	0.53
umc1277	AATA	9.07-9.08	54	4	130-142	0.68	0.61

Samples containing up to two PCR reactions (0.5ml of each), 0.3 ml GeneScan350 size standard (Applied Biosystems, Foster City California) labeled with N′,N′,N′,N′- teteramethyl-6-carboxyrhodamine (TAMRA), and 30% (v/v) formamide were heated at 95°C for 5 min, placed on ice, then loaded on 4.5% (w/v) denaturing (6M Urea) acrylamide:bisacrylamide (29:1) gels (30cm well-to- read). DNA samples were electrophoresed in 1X TBE buffer (pH 8.3) at constant voltage (3kV) for up to 2 hours on an ABI Prism 377 DNA sequencer (Applied Biosystems, Foster City, CA).

Statistical analysis

Fragment sizes were automatically calculated with GeneScan3.1 (Perkin Elmer/Applied Biosystems) using the local southern sizing method. The GeneScan data were assigned to alleles using the Category function of Genotyper 2.1 (Perkin Elmer/Applied Biosystems), manually checked for confirmation of correct identification of alleles, and exported to an excel file recording allele identity for each locus for each individual. Data were stored in an Access database and peak sizes were converted to the proper configuration for subsequent analysis (0,1 binary matrix for each locus).

The number of alleles per locus was determined for the entire set of 38 inbred lines. Polymorphic Information Content (PIC) of each SSR marker was determined as described by SMITH *et al.* (1997):

$$1 - \sum_{i=1}^{n} P_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2\, P_i^2\, P_j^2$$

where P_i and P_j are the frequencies of the ith and j[th] allele of a given marker, respectively.

Genetic similarities (GS) were estimated using the Dice coefficient (SNEATH and SOKAL, 1973). Cluster analysis was carried out on the matrix of genetic similarities using the Unweighted Pair Group Method using Arithmetic Averages (UPGMA) clustering algorithm. The similarity matrix and dendrogram were constructed with NTSYS-pc version 2.02 (ROLF, 1997). In order to determine the statistically correct number of clusters, the pseudo t^2 statistic was estimated using the SAS software package, and Ward's method, according to FRANCO *et al.* (2001) and WARBURTON and CROSSA (2000). The number of alleles per locus, PIC values, genetic distance, and average gene diversity in cluster groups were all calculated using the PowerMarker software package (v.3.06, LIU, 2001-2004) and Private Alleles (alleles specific to one group of lines or heterotic group) determined for these groups with the GDA software version 1.1 (LEWIS and ZAYKIN, 2002).

RESULTS AND DISCUSSION

SSR polymorphism

All 43 SSR loci analyzed were polymorphic with a total of 194 alleles detected (Table 2). There was an average of 4.5 alleles per locus, with a range from 2 (phi046, phi089 and phi123) to 10 (phi064). These numbers are similar those previously report-

TABLE 3 - *Unique SSR alleles (occurred in no more than 1 inbred line) and private alleles (occurred in only one cluster; private alleles from cluster E was not included) as identified in the 38 inbred lines measured in this study.*

Unique alleles			Private alleles			
Locus	Allele size	Inbred allele is specific for	Locus	Allele Size	Inbred allele is specific for	Cluster
nc130	139	KE72012/1	phi112	155	K1249/3	A
nc133	105	H43/1-42(2)	phi121	106	K1228/7	A
nc133	125	K1264/4	phi299852	117	K1264/4	A
nc133	145	H43/1-42(1)	phi299852	129	K1604/5	B
nc133	150	TVA926/1-22	phi299852	144	K2816	A
nc133	160	K1263/17	phi308707	125	K1604/5	B
phi032	245	K1264/4	phi308707	128	K2816	A
phi050	104	K1249/3	phi328175	103	CML051	E
phi050	72	K1728/7	phi328175	115	TVA962/1-22	A
phi050	80	KE72008/1	phi328175	133	OH43/1-411	D
phi059	153	K2816	phi339017	163	K615/1	C
phi059	159	OH43/1-411	phi374118	214	OH43/1-42(2)	D
phi062	155	S61/1-42	phi374118	226	KE76005/1-1-1-1-1-1	A
phi064	101	K58	phi423796	141	KE72012/1	B
phi064	69	H43/1-42(2)	umc1061	107	OH43/1-42(1)	D
phi064	81	K615/1	umc1109	113	K3304/1-1	A
phi064	89	TVA	umc1143	64	K1264	A
phi065	126	CML292	umc1152	167	K1298/1	C
phi065	141	K1249/3	umc1196	167	K1264/4	A
phi069	208	KE72009/1	umc1277	130	KE72009/1	B
phi102228	119	CML292	umc1279	95	CML292	E
phi109188	156	KE76005-1-1-1-1-	umc1445	78	S61/4-2	A

ed for CIMMYT (WARBURTON *et al.*, 2002; REIF *et al.*, 2003a) and US maize inbred lines (SENIOR *et al.*, 1998; LU and BERNARDO, 2001). Other studies showed either lower (DRINIC and KONSTANTINOV, 2002; BANTTE and PRASANNA, 2003) or higher numbers of alleles (PEJIC *et al.*, 1998; LI *et al.*, 2002) depending on the SSR repeat type and number of lines used. SSR loci amplifying a dinucleotide repeat unit show higher numbers of alleles than other SSR loci; however, dinucleotide SSRs are not used at CIMMYT in general because of the difficulty in accurately sizing alleles. The number of lines used in the current study was lower than most of the previously mentioned studies, which would tend to display fewer alleles due to the decreased probability of finding rare alleles in smaller samples.

The PIC values of SSR loci ranged from 0.15 (phi423796) to 0.78 (phi328175) with an average of 0.53 (Table 2), which also were within the range of previous studies when repeat type was taken into account (SMITH *et al.*, 1997; SENIOR *et al.*, 1998; ENOKI *et al.*, 2002; LI *et al.*, 2002; BANNTE and PRASANNA, 2003). Forty-four unique alleles (those present in only one inbred) were detected at 27 SSR loci across the inbreds analyzed, the highest numbers coming from locus nc133 (5 out of 8 alleles), and phi064 (4 out of 10 alleles) (Table 3). Of the 44 unique alleles, 4 were specific to the CIMMYT lines (CML051 and CML292). Although the exact reason why so many unique alleles were detected is unknown, the fact that many of the inbreds were derived from CIMMYT materials may be partially the reason, as many unique or rare (present in low frequency) alleles are found in CIMMYT germplasm (WARBURTON *et al.*, 2002; REIF *et al.*, 2003a,b). The presence of so many unique alleles does suggest that SSR markers will be useful in germplasm source identification.

Gene diversity ranged from 0.16 (phi423796) to 0.80 (phi064) with an average of 0.58 (Table 2). Gene diversity is equivalent to the expected heterozygosity for diploid data, and it is defined as the probability that two randomly chosen alleles are different in the sample (i.e., came from different ancestors). Summary statistics measuring diversity were slightly lower for the early maturing lines than those in the medium and late maturing maize inbred lines studied in a companion paper. This was possibly due to the slightly smaller number of lines in the early maturing sample, but may also reflect a narrower genetic base in the early maturing lines as compared to the medium and late maturing lines.

Cluster analysis

Using the pseudo t^2 calculations, inbreds were grouped into five clusters, designated A through E (Fig. 1). Cluster A was composed of three subgroups, labeled a1 – a3 in the dendrogram. This cluster contained predominantly the lines extracted from CIMMYT germplasm (10 lines), along with some lines selected from germplasm originating from Yugoslavia (3 lines), Hungary (1 line), and Poland (2 lines), and two lines from unknown sources. Subgroup a1 was a mixed group of Yugoslavian, Polish and CIMMYT materials, while subgroup a2 and a3 consisted of 8 lines extracted from CIMMYT germplasm, plus K33 in subgroup a2 (from Yugoslavian source material) and KE76005/1-1-1-1-1-1-1 in subgroup a3 (from Hungarian source material). Cluster B consisted of two lines extracted from CIMMYT source materials and two lines from the combined source material, while Cluster C was a mixed group of 4 Yugoslavian, 2 CIMMYT, 1 Belgian, 1 Hungarian and 1 Iranian source material extracted lines. Clusters B and C were also very heterogenous and grouped at a low level of similarity (~0.40), indicating a low level of cohesiveness within the cluster (similar to what is seen in Cluster A). Cluster D was comprised of only Lancaster related sister lines of Oh43 pedigree. Cluster E, which was the most distant cluster in the dendrogram, included only the 2 CIMMYT control lines (CML051 and CML292), showing the power of SSR markers to discriminate temperate materials from non-temperate.

The grouping pattern of inbred lines in this study generally agreed with the known pedigree or geographic information of the source materials (hybrids from which the inbred lines were drawn). In addition, lines with no previous information about pedigree or source can now be classified with the lines most closely related to them. REIF *et al.* (2003b) suggest that when a large number of lines exist but established heterotic groups are unavailable, genetically similar germplasm can be identified with molecular markers to form potential heterotic groups, which then must be verified via field crosses. Supporting this suggestion is the fact that sister lines S61 and S61/4-2, as well as sister lines TVA and TVA962/1-22 grouped in Cluster A, subgroup a1. All sister lines K1264, K1264/4, K1264/5, K1263/1, K1263/2-1 and K1264/17 clustered together in Cluster A, subgroup a2. On the other hand, contrary to what their close genetic similarities suggest, S61 and TVA are the parents of a successful early maturing hybrid in Iran.

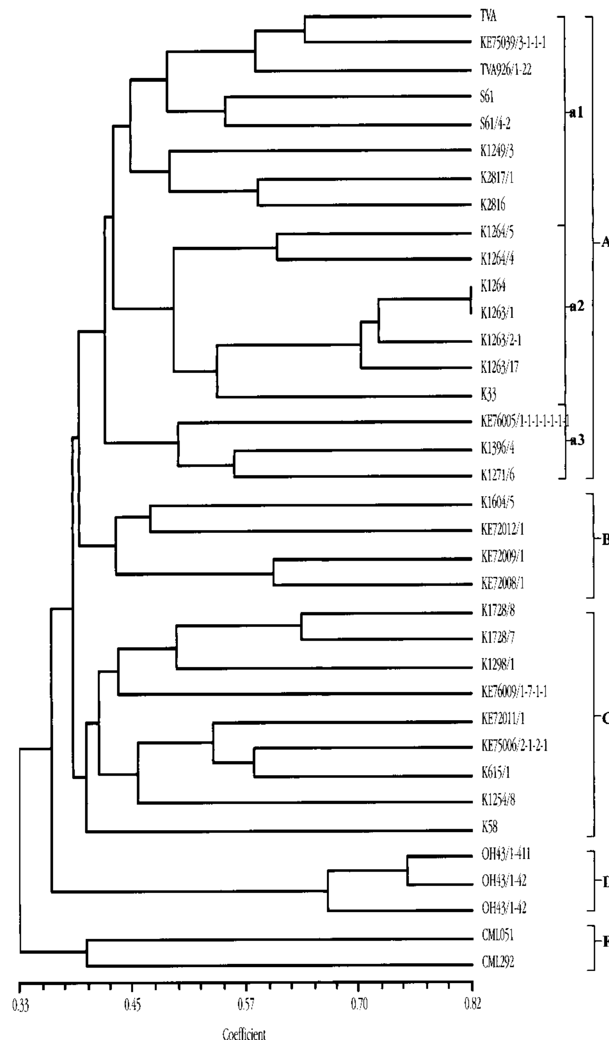

FIGURE 1 - UPGMA dendrogram of 38 maize inbred lines determined on the basis of Dice genetic similarity coefficients calculated from 43 SSR markers.

Incongruities are not uncommon when comparing the results of molecular analyses with classifications based on known pedigree or field performance. MO17 and B73 clustered together in a marker study by MESSMER *et al.* (1992), although they belong to different heterotic groups. LI *et al.* (2002) found that B73 and B75 clustered into separated groups, although B73 is from BSSS and B75 from a related population, BSCB#3. The incongruities can be caused by one or more factors. DNA markers, when identified by size, may represent alleles that are only identical in state, and may not always be identical by descent, which can be true in our case (MUMM and DUDLEY, 1994). In addition, DNA sequences are affected by selection, drift and mutation (SENIOR *et al.*, 1998; MUMM and DUDLEY, 1994); these are recorded in the markers, but are not accounted for in pedigrees. Furthermore, it must also be considered that these lines could have had pedigrees recorded incorrectly or were subject to seed handling errors in the past; markers may be more accurate measurements of relationship in these cases.

Genetic diversity

Within the sample of 38 inbred lines, 68% of the 194 alleles occurred at a frequency of 0.25 or less, and 3% of the alleles occurred at a frequency of 0.75 or greater, indicating moderate allelic diversity, as alleles occurring at very low or very high frequency did not contribute as much to the overall diversity within the set of lines in this study. Average allelic diversity within clusters was slightly less than overall allelic diversity, while it was greatly diminished in clusters D and E. Total numbers of alleles and number of alleles per locus in Clusters A and D were highest and lowest, respectively. This was not unexpected due to the higher number of lines in Clusters A and C and fewer lines and narrower range of germplasm in Clusters D and E (Table 4). When the tropical checks (Cluster E) were excluded, the highest genetic distance based on marker data was found between Clusters B and D (Table 5) and the lowest distance between Clusters A and B. Similar and moderate distances were found between the other groups. The higher distance between Cluster D and other groups was most likely due to the fact that the germplasm in Group D derived from U.S. hybrids, which come from a very different breeding program, and tend not to be mixed with germplasm from other breeding programs. Crosses should be made between Clusters A and B; and B and C to confirm what appears to be potential heterotic patterns based on the marker study, and previously reported successful hybrid combinations (data not shown).

Clusters A and C appear to be highly diverse within the groups considering gene diversity, number of alleles and alleles per locus (Table 4). This suggests that these two groups have a broad genetic base and may be used as a source for re-extraction of new inbred lines. By using Reciprocal Recurrent Selection (RRS) within each group, lines may be

TABLE 4 - *Summary statistics for all inbreds and clusters identified in this study including sample size, number of alleles, number of alleles per locus, number of private alleles, number of unique alleles, gene diversity, and mean PIC values.*

Statistic	Overall	Cluster groups				
		A	B	C	D	E
Sample size	36	18	4	9	3	2
Allele no.	194	152	89	125	64	69
Allele/Locus	4.5	3.5	2.1	2.9	1.5	1.6
# of private alleles	54	7	7	16	24	9
# of unique alleles	45	5	4	11	16	9
Gene diversity	0.58	0.51	0.39	0.49	0.20	0.30
PIC	0.53	0.46	0.32	0.44	0.16	0.23

TABLE 5 - *Dice Genetic Distance coefficient matrix between clusters as calculated using allele frequencies.*

	B	C	D
A	0.49	0.41	0.61
B		0.50	0.77
C			0.71

better adapted to Iranian growing conditions without mixing up the potential heterotic groups. After development of inbred lines from extraction from other sources or RRS, breeders need to make thousands of crosses and evaluate grain yield in resulting F_1 plants in replicated field experiments. Heterotic groups allow breeders to make fewer crosses, as they can maximize heterosis by only crossing between heterotic groups, and not within. However, in Iran, no established heterotic groups currently exist, and much of the breeding material has unknown pedigrees and sources. Therefore, we propose that the SSR classification information presented in this study could be used to establish preliminary heterotic groups. Once confirmed by field crossing, these heterotic groups would allow fewer hybrids to be tested each year, and hybrid maize breeding efforts may be greatly enhanced.

ACKNOWLEDGEMENTS - Financial support provided to R. Choukan by the Agricultural Research and Education Organization (AREO) and the Seed and Plant Improvement Institute, (SPII), Iran is gratefully acknowledged. We gratefully acknowledge the laboratory facilities provided by CIMMYT (The International Maize and Wheat Improvement Center), Mexico. The authors wish to thank the excellent laboratory assistance of Ana Lidia Gomez and Leticia Diaz. We gratefully acknowledge Mrs. Z. Dehghanpoor (SPII, Iran) for supplying us with inbred line seeds.

REFERENCES

BANTTE K., B.M. PRASANNA, 2003 Simple sequence repeat polymorphism in Quality Protein Maize (QPM) lines. Euphytica **129:** 337-344.

BETRAN F.J., J.M. RIBAUT, D. BECK, D. GONZALEZ DE LEON, 2003 Genetic diversity, specific combining ability, and heterosis in tropical maize under stress and nonstress environments. Crop Sci. **43:** 797-806.

CIMMYT, 2001 The applied biotechnology center's manual of laboratory protocols. First edition. CIMMYT, Mexico D.F.

CLARK B.C., L.B. MORAN, R. APPELS, 1989 DNA analysis in wheat breeding. Genome **32:** 334-339.

DRINIC M.S., K. KONSTANTINOV, 1998 Genetic diversity and its relationship to heterosis in maize as revealed by PCR-based markers. Genetika **30:** 227-237.

DUBREUIL P., P. DUFOUR, E. KREJCI, M. CAUSSE, D. VIENNA, A. GALLAIS, A. CHARCOSSET, 1996 Organization of RFLP diversity among inbred lines of maize representing the most significant heterotic groups. Crop Sci. **36:** 790-799.

ENOKI H., H. SATO, K. KOINUMA, 2002 SSR analysis of genetic diversity among maize in bredlines adapted to cold regions of Japan. Theor. Appl. Genet. **104:** 1270-1277.

FRANCO J., J. CROSSA, J.M. RIBAUT, J. BETRAN, M.L. WARBURTON, M. KHAIRALLAH, 2001 A method for combining molecular markers and phenotypic attributes for classifying plant genotypes. Theor. Appl. Genet. **103:** 944-952.

GONZALEZ S., H. CORDOVA, S. RODRIGUE Z., H. DE LEON, V.M. SERRATO, 1997 Determinación de un patrón heterótico a partir de la evaluación de un dialelo de diez líneas de maíz subtropical. Agron. Mesoamericana **8:** 1-7.

HALLAUER A.R., 1990 Methods used in developing maize inbred. Maydica **35:** 1-16.

HALLAUER A.R., E. LOPEZ-PEREZ, 1979 Comparisons among testers for evaluating lines of corn. Proc. Ann. Corn and Sorghum Res. Copn. **34:** 57-75.

HALLAUER A.R., J.B. MIRANDA, 1988 Quantitative genetics in maize breeding. 2nd ed. Iowa State Univ. Press. Ames.

HAN G.C., S.K. VASAL, D.L. BECK, E. ELIS, 1991 Combining ability of inbred lines derived from CIMMYT maize (*Zea mays* L.) germplasm. Maydica **36:** 57-64.

KONSTANTINOV K., S.M. DRINIC, 2000 Heterosis in maize molecular markers polymorphism and probability of prediction. Quantitative genetics and breeding methods: The way ahead. pp.310-312. Paris. France, August 30-september 1.2000.

LEE M., E.B. GODSHALK, K.R. LAMKEY, W.W. WOODMAN, 1989 Association of restriction fragment length polymorphisms among maize inbreds with agronomic performance of their crosses. Crop Sci. **29:** 1067-1071.

LEWIS P.O., D. ZAYKIN, 2002 Genetic data analysis software. Release 1.1. Available online at http://lewis.eeb.unconn.edu/lewishome/.

LI Y., J. DU., T. WANG, Y. SHI, Y. SONG, J. JIA, 2002 Genetic diversity and relationships among Chinese maize inbred lines revealed by SSR markers. Maydica **47:** 93-101.

LIU J., 2002-2004 Power Marker V3.06 User's manual.

LIVINI C., P. AJMONE-MARSAN, A.E. MELCHINGER, M.M. MESSMER, M. MOTTO, 1992 Genetic diversity of maize inbred lines within and among heterotic groups revealed by RFLPs. Theor. Appl. Genet. **84:** 17-25.

LU H., R. BERNARDO, 2001 Molecular marker diversity among current and historical maize inbreds. Theor. Appl. Genet. **103:** 613-617.

MELCHINGER A.E., R.K. GUMBER, 1998 Overview of heterosis and heterotic groups in agronomic crops. pp. 29-44. *In:* K.R. Lamkey, J.E. Staub (Eds.), Concepts and breeding of heterosis in crop plants. CSSA, Madison, Wisconsin.

MESSMER M.M., A.E. MELCHINGER, J. BOPPENMAIR, R.G. HERMANN, E. BRUNKLAUS-JUNG, 1992 AFLP analyses of early-maturing European maize germplasm. I. Genetic diversity among flint and dent inbreds. Theor. Appl. Genet. **83:** 1003-1012.

MUMM R.H., J.W. DUDLEY, 1994 A classification of 148 U.S. maize inbreds: I. Cluster analysis based on RFLPs. Crop Sci. **34:** 842-851.

PEJIC I., P. AJMONE-MARSAN, M. MORGANTE, V. KOZUMPLICK, P. CASTIGLIONI, G. TARMINO, M. MOTTO, 1998 Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor. Appl. Genet. **97:** 1248-125.

REIF J.C., A.E. MELCHINGER, X.C. XIA, M.L. WARBURTON, D.A. HOISINGTON, S.K. VASAL, G. SRRINIVASAN, M. BOHN, M. FRISCH, 2003a Genetic distance based on the simple sequence repeats and heterosis in the tropical maize populations. Crop Sci. **43:** 1275-1282.

REIF J.C., A.E. MELCHINGER, X.C. XIA, M.L. WARBURTON, D.A. HOISINGTON, S.K. VASAL, D. BECK, M. BOHN, M. FRISCH, 2003b Use of SSRs for establishing heterotic groups in subtropical maize. Theor. Appl. Genet. **107:** 947-957.

ROLF F.J., 1997 NTSYpc. Numerical taxonomy and multivariate analysis system, version 2.02. Applied Biostatistics Inc., New York.

SENIOR M.L., M. HEUN, 1993 Mapping maize microsatellite and polymerase chain reaction confirmation of the targeted repeats using a CT primer. Genome **36:** 884-889.

SENIOR M.L., J.P. MURPHY, M.M. GOODMAN, C.W. STUBER, 1998 Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. Crop Sci. **38:** 1088-1098.

SMITH J.S.C., O.S. SMITH, 1992 Fingerprinting crop varieties. Adv. Agron. **47:** 85-129.

SMITH J.S.C., E.C.L. CHIN, H. SHU, O.S. SMITH, S.J. WALL, M.L. SENIOR, S.E. MITCHELL, S. KRESOVICH, 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., R.R. SOKAL, 1973 Numerical taxonomy. Freeman, San Francisco.

TERRON A., E. PRECIADO, H. CORDOVA, H. MICKELSON, R. LOPEZ, 1997 Determinacion del patron heterotico de 30 lineas de maiz derivadas del la poblacion 43 SR del CIMMYT. Agron. Mesoamericana **8:** 26-34.

WARBURTON M.L., J. CROSSA, 2000 Data analysis in the CIMMYT Applied Biotechnology Center for fingerprinting and genetic diversity studies. First edition. Mexico, D.F., CIMMYT.

WARBURTON M.L., X. XIANCHUN, J. CROSSA, J. FRANCO, A.E. MELCHINGER, M. FRISCH, M. BOHN, D. HOSINGTON, 2002 Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. Crop Sci. **42:** 1832-1840.

YUAN L.X., J.H. FU, S.H. ZHANG, X. LIU, Z. PENG, X.H. LI, M.L. WARBURTON, M. KHAIRALLAH, 2001 Heterotic grouping of maize inbred lines using RFLP and SSR markers. Acta Agron. Sinica **27:** 149-156.