Crop Breeding and Applied Biotechnology 5:1-8, 2005
Brazilian Society of Plant Breeding. Printed in Brazil

Genetic variation among and within sweet corn populations detected by RAPD and SSR markers

Fernanda Bered, Tatiana de Freitas Terra, Marciel Spellmeier, and José Fernandes Barbosa Neto

Received 9 September 2005

Accepted 29 October 2005

ABSTRACT - *The main objective of the present study was to characterize the genetic structure of three Brazilian sweet corn populations. In addition, we aimed to investigate the molecular genetic diversity within and among the populations and evaluate the ability of RAPD and SSR markers to detect inter and intra-population variation in these populations. PCR were performed using 11 RAPD and six microsatellite primer pairs. Genetic similarity was estimated within and among populations considering the markers independently and jointly. Variability for each locus was measured using the Polymorphism Information Content and the correlation between pairwise distance matrices was obtained. The genetic similarity and analysis of the molecular variance results indicated a higher similarity among than within populations, and RAPD showed lower PIC values than SSR. Knowledge on the genetic diversity distribution in this germplasm is important to determine strategies for future conservation and selection for breeding programs.*

Key words: Genetic diversity, germplasm, molecular markers, *Zea mays*.

INTRODUCTION

In the United States sweet corn is economically one of the most important vegetables (Tracy 1997). In Brazil the economical potential of the crop consists in fresh consumption, although its production is restricted to Southern Brazil. Several factors have contributed to this small expression, including the lack of adapted varieties and the presence of undesirable agronomic traits such as tall plants, lodging susceptibility and poor ear placement (Silva 1983, Tosello 1997). The development of superior sweet corn genotypes is crucial to overcome these problems.

Information about germplasm diversity in crops is relevant for plant breeding since the genetic improvement of any organism depends on the availability of genetic variability (Allard 2000). Moreover, the partitioning of variability between and among populations determines the adopted breeding strategy (Chalmers et al. 1992). Genotypes are often identified through morphological and physiological traits, although these estimators are limited in number and can be affected by environmental factors (Manifesto et al. 2001).

Molecular methods are adequate for genetic diversity studies, particularly markers generated by applying the polymerase chain reaction (PCR). (Lynch and Milligan 1994, Chan and Sun 1997). RAPD (Random Amplified Polymorphic DNA) is a quick and low-cost non-radioactive analysis and has been widely used in plant population studies, including maize (Chalmers et al.1992, Lynch and Milligan 1994, Bartish et al. 1999, Mengoni et al. 2000, Popi et al. 2000). On the other hand, the SSR (Simple Sequence Repeat or Microsatellite) technique has quickly gained acceptance because of its codominant nature, reproducibility and high information content (Manifesto et al. 2001). Recent studies have documented microsatellite

loci for a wide range of crops, including maize, and have been used in genetic population and cultivar studies as well as genetic mapping and fingerprinting (Senior and Heun 1993, Chin et al. 1996, Warburton et al. 2002, Reif et al. 2004). The simultaneous use of RAPD and SSR markers may be relevant to assess diversity in genotypes or populations (Mengoni et al. 2000). The main objective of the present study was to characterize the genetic structure of three Brazilian sweet corn populations as a basis for breeding programs. In addition, we aimed to i) investigate the molecular genetic diversity within and among populations, and ii) evaluate the ability of RAPD and SSR markers to detect inter and intra-population variation in these populations.

MATERIAL AND METHODS

Plant material and DNA isolation

This study focused on three sweet corn populations which represent the main type of germplasm available, including two mutant types, the sweet and supersweet series (Table 1). Seeds of 30 plants per population were germinated and the total DNA isolated from the leaves of each plant according to Edwards et al. (1991). The same DNA samples were used in the RAPD and SSR analyses. Crossa et al. (1993) showed that the sample size (n) required to retain, at probability P, at least one copy of each of K allelic classes in each one of m loci should be larger than

$$n > \frac{\log\left[1-(P)^{\frac{1}{m}}\right]-\log(k-1)}{\log(1-P)}$$

Considering six SSR loci with four alleles each, at 95% probability, we need 22 plants to obtain a reliable estimate.

PCR markers

RAPD amplification was performed in a total volume of 25 ml according to the protocol described by Mailer et al. (1997). Initially, 20 RAPD decamer primers were screened in order to test the amplification profiles for readability and reproducibility. After this procedure, 11 primers were chosen for analysis (Table 2). Six microsatellite loci were selected from the Missouri Maize Project using MaizeDB (nucleus.agron.Missouri.edu/index.html). Information about SSR loci is included in Table 3. PCR were performed in a total volume of 20 ml according to Liu et al. (1996).

RAPD amplification products were resolved on a 1.4% agarose gel and visualized after ethidium bromide staining, while the SSR products were resolved in a 6% polyacrylamide gel stained with silver nitrate.

Data analysis

The bands visualized on the gels were scored using a binary code for their presence (1) or absence (0) for both RAPD and SSR. This way we determined a molecular binary phenotype for each individual plant by linearly combining the presence/absence of each marker.

In view of the respective dominant and codominant nature of RAPD and SSR markers, it was assumed that each RAPD band represented a distinct locus with two alleles. Unbiased frequencies of null alleles were computed according to Lynch and Milligan (1994). Allelic frequencies were estimated from SSR data and each locus was characterized and tested for the Hardy-Weinberg Equilibrium using TFPGA software (Miller 1997).

Genetic similarity was estimated within populations using the Jaccard similarity index $J = \frac{a}{(a+b+c)}$, where *a* is the number of positive matches, and *b* and *c* refer to the number of bands present in genotype 1 and 2 only, respectively (Jaccard 1908). The distance between the populations was estimated using distance $D = \sqrt{\frac{\text{Roger's W}}{n}}$, where n = number of observations (Barbosa-Neto et al. 1997) considering RAPD and SSR independently and jointly. The similarity analysis was performed using NTSYS-pc 2.10m (Rohlf 2000).To determine how many fragments have to be evaluated to estimate genetic distances the expression $N = \frac{(Z_{\alpha/2})^2 f(1-f)}{E^2}$ was used where N = optimal number of fragments to achieve a desired level of precision for the estimated genetic distance, f = frequency of polymorphic fragments and E = confidence interval (Barbosa-Neto et al. 1997).

The variability for each locus was measured using the Polymorphism Information Content (PIC) by the expression $PIC = 1 - \sum p_u^2$ where p_u was the frequency of the *u* allele. The correlation between pairwise distance matrices derived from RAPD and SSR data was obtained by the method of Mantel (1967). The PIC calculation and

Table 1. Type, germplasm, mutant gene, and developer of the three studied sweet corn populations

Population	Population	Germplasm	Mutant	Developer
BR 400	Open pollinated	Hawaiian supersweet series	*bt1*	EMBRAPA
BR 401	Open pollinated	Hawaiian sweet series	*su1*	EMBRAPA
BR 402	Open pollinated	Doce de Cuba	*su1*	EMBRAPA

Table 2. Number and sequence of primers employed in the study, number and size of RAPD amplified fragments, Polymorphism Information Content (PIC), corrected PIC and number of fragments exclusive to or absent in determined populations only for the primers employed in the study

RAPD Primer	Primer sequences	Number of amplified fragments	Fragment size (bp)	PIC (0-0.5)	corrected PIC (0-1)	Nr of exclusive fragments			Nr of absent fragments		
						400	401	402	400	401	402
20	TGCGCCCTTC	03	700-900	0.21	0.41						3
21	GTCCACACGG	03	450-1000	0.13	0.26		1	1		1	
22	ACGGATCCCT	01	700	0	0						
23	GAGGATCCCT	04	550-1700	0.23	0.45						
24	CCTGATCACC	10	400-1400	0.18	0.37		1	2		1	
25	GGTGATCAGG	04	600-1000	0.24	0.48						1
26	CCGAATTCCC	12	350-1400	0.24	0.49	1					1
27	GGGAATTCGG	09	300-1550	0.21	0.41	2		1	1	2	
28	CCGATATCCC	08	350-1500	0.18	0.35	2		1	1		1
29	GGGATATCGG	08	550-1550	0.09	0.19	3	2				1
30	GTAGCACTCC	10	450-1150	0.21	0.42	2			2	2	
Mean		6.5	300-1700	0.17	0.35	11	3	5	4	6	7

Table 3. Maize SSR, chromosomal locations, motif, primer sequences, number of alleles, Polymorphism Information Content (PIC), and number of fragments exclusive to or absent in determined populations only for the loci employed in the study

SSR loci	Motif	Sequence (5`- 3`) R/F	Nr. of alleles	PIC (0-1)	Nr of exclusive fragments			Nr of absent fragments		
					400	401	402	400	401	402
Umc 1620	TTC(4)	CCACCGAGTGACTAGTTGTGAGAG// CCTTTCAATGTTCATGTTCTCTTCC	02	0.46						
Umc 1023	AT(11)	CTTGTGCCACCACATGCAGTA// CAGTTTGGAACAGGGAAAAGTACG	06	0.73	1			1		
Umc 1097	CA(8)	CTCGTCAACGTCAACCCAAGTAAG// CTGTTAGATGTGCGACAACAGAGC	05	0.62						
Umc 1034	GA(12)	GTGTTTCCGTTTCGCTGATTTTAC// TCATCCATGTGACAGAGACGACTT	06	0.62	1					
Umc 1331	GGT(10)	TTATGAACGTGGTCGTGACTATGG// ATATCTGTCCCTCTCCCACCATC	03	0.39						
Umc 1648	TC(8)	CTGCAGTACGTGAGCCTGTACG// GCTTGAGCTGTGAGGAAGTTTTG	03	0.63		1		2	2	
Mean			4.2	0.57	1	1	1	2	3	

Mantel test analysis were performed using Microsoft Excel and NTSYS-pc respectively.

Analysis of Molecular Variance (AMOVA) (Excoffier et al. 1992) was used to analyze the partition of the total genetic variation among and within populations.

RESULTS AND DISCUSSION

Only one of the eleven RAPD primers under study was monomorphic. A total of 72 loci were scored (Table 2), and only bands that showed clear reproducible patterns were chosen for analysis. The mean number of bands per primer was 6.5, ranging from one to twelve. Sun et al. (2001) reported 6.7 fragments per RAPD primer in maize hybrids.

The results of PCR amplification of the SSR loci are summarized in Table 3. All six SSR loci displayed polymorphism and a total of 25 alleles distributed among the analyzed genotypes were assessed, with a mean number of 4.2 and ranging from two to six. The amplified fragments varied in size from 137 to 143 bp for all loci included in the analysis. An average of 4.6 and 4.9 alleles per SSR loci in maize were described by Sun et al. (2001) and Warburton et al. (2002), respectively.

As expected for dominant markers, RAPD presented lower PIC values than SSR. In the present study, the mean PIC estimated for SSR loci was 0.57 (Table 3) and varied from 0.39 (umc 1331) to 0.73 (umc 1023), while for RAPD it varied from 0 (primer 22) to 0.24 (primers 6 and 7) with a mean of 0.17 (Table 2). Absolute PIC values of RAPD and SSR markers are not comparable because the maximum PIC values are 0.5 and 1.0 for RAPD and SSR loci, respectively. A correction column was inserted in Table 2 to make values comparable. Fahima et al. (2002) studied SSR polymorphism of 15 *Triticum dicoccoides* populations and the results showed that SSR loci are more polymorphic and revealed a higher genetic diversity than allozyme and RAPD loci. Loci with high PIC values can be useful to discriminate germplasm in future studies. Warburton et al. (2002) reported that the most discriminatory SSRs were not always those with the highest PIC. Some SSR loci can be identified as being discriminatory in the evaluated germplasm and still have a low PIC value. In the present study, fragments identified as specific for a determined population did not necessarily have high PIC value (Tables 2 and 3). It is therefore not possible to predict which SSRs will be the most discriminatory on the sole basis of PIC (Warburton et al. 2002).

Table 4 shows the variability estimates among and within populations, which were calculated for a total of 95 fragments. According to Barbosa-Neto et al (1997), 45 fragments are sufficient to obtain a reliable genetic distance when $\alpha = 0.10$ and $Z = 1.64$. The similarities observed in the RAPD and SSR data were higher among than within populations, which was as expected for open-pollinated populations. The similarity index was calculated among populations considering the expression 1-D, and results showed that the mean similarity was very high. The RAPD and SSR data sets

Table 4. Genetic diversity among (D distance) and within populations (Jaccard coefficient) considering each marker type independently and jointly

Marker Type	Mean distance among populations (D distance)	Mean similarity among populations (1-D)	BR400 (Jaccard)	BR401 (Jaccard)	BR4012 (Jaccard)
RAPD	0.053	0.947	0.52	0.40	0.50
SSR	0.072	0.928	0.47	0.28	0.36
RAPD+SSR	0.044	0.956	0.50	0.32	0.41

were combined and 97 markers were analyzed. The genetic similarity results indicated the same tendency of higher similarity among than within populations. The results of the AMOVA partitioning of genetic variation (Table 5) agree with the genetic similarity results, because most of the total genetic variation was found within populations (88%). According to Bartish et al. (1999) the genetic variability within plant populations depends on the plant breeding system, distribution area and other minor factors. Outcrossing and wind-pollinated species such as maize tend to have a high within-population variability. Amongst all data sets, population BR 401 presented the highest variability (Table 4). Populations BR 400 and BR 401 were very heterogeneous at the molecular marker level and poorly distinguished from one another in spite of having the same origin. BR 402, although from a different origin, could not be distinguished from the other two populations by D distance. The genetic variability assessed within populations seems to be population-specific since there were exclusive fragments in each of them (Table 2 and 3). Knowledge on the genetic diversity distribution in this germplasm is important to determine strategies for future conservation and for selection for breeding programs.

Table 5. AMOVA for the three sweet corn populations under study

Source of variation	df	SS	Variance component	Variation (%)
Among populations	2	2.34	0.04	12
Within populations	69	19.23	0.28	88
Total	71	21.57	0.32	100

The Mantel test (Mantel 1967) established low and non-significant correlations between RAPD and SSR pairwise similarity matrices, considering each population (r=0.024 to BR400, r=0.019 to BR401 and r=0.06 to BR402). The different molecular basis of RAPD and SSR markers can explain the differences in the results. SSR polymorphism is created by DNA polymerase slippage, while RAPD variation resides in many different types of mutational events that occur in the annealing site of the primer and in/in-between? the sites responsible for the amplification (Mengoni et al. 2000). The standard scoring procedure used for both markers can affect the results and result in a low correlation between RAPD and SSR similarity matrices. Mengoni et al. (2000) used RAPD and SSR variation to assess genetic relationships among alfalfa populations and obtained low correlation between RAPD and SSR data through the Mantel test. Manifesto et al. (2001) evaluated wheat germplasm using SSR and AFLP markers, and the correlation between similarity matrices was very low.

The Hardy-Weinberg Equilibrium tests considering all three populations and the six SSR loci studied here are presented in Table 6. Populations BR 400 and BR 401 showed five of the six loci analyzed in Hardy-Weinberg equilibrium, however, the locus that presented deviation from equilibrium in BR 400 showed an excess of heterozygosity, while in BR 401 there was a deficit of heterozygosity. In the BR 402 population only two of six loci were in equilibrium. All loci that presented deviation from the Hardy-Weinberg equilibrium presented excess homozygosity. The departures from the Hardy-Weinberg equilibrium observed at some loci with excess homozygosity and low heterozygosity values can be related to various causes: (i) loci with smaller numbers of alleles or with a skewed frequency distribution; (ii) positive assortative matings between individuals (homogamy); (iii) selection favoring homozygotes (Collevatti et al. 2001, Reif et al. 2004).

The analysis of the three sweet corn populations revealed that most of the genetic diversity is within populations, and that this variation is not necessarily the same, since there are fragments exclusive to particular populations. This result can be very useful for breeding programs, since different genotypes can be discovered within the populations. Moreover, the populations are in Hardy-Weinberg equilibrium in most loci, and this germplasm can be useful for the improvement of open-pollinated varieties. This germplasm is however less suitable for hybrid breeding, where clearly distinct heterotic groups are advantageous (Reif 2004). The results obtained with the RAPD and SSR analyses were similar to each other, and although the correlation (the Mantel test) between genetic distances calculated from the RAPD and SSR data was poor, molecular markers provide a valuable tool for germplasm grouping and are a good complementation to field trials for identifying groups of genetically similar germplasm (Reif et al. 2003).

ACKNOWLEDGMENTS

The authors would like to thank CNPq and FAPERGS for the support of this research.

Table 6. Number of loci with and without deviation from Hardy-Weinberg Equilibrium in the sweet corn populations

Population	Loci in Hardy-Weinberg equilibrium	Loci with deviation from Hardy-Weinberg equilibrium		Total
		Excess homozygosity	Excess heterozygosity	
BR 400	5	0	1	6
BR 401	5	1	0	6
BR 402	2	4	0	6
Total	12	5	1	

Variação genética em populações de milho doce detectada através de RAPD e SSR

RESUMO - *Os objetivos deste trabalho foram caracterizar a estrutura genética de três populações brasileiras de milho doce; investigar a variabilidade genética, em nível molecular, entre e dentro de populações; avaliar a capacidade dos marcadores RAPD e SSR em detectar a variação intra e inter-populacional. Foram utilizados 11 primers de RAPD e seis pares de primers de microssatélites. A similaridade genética foi estimada dentro e entre as populações considerando os marcadores independentemente e combinados. A variabilidade para cada loco foi estimada através do índice de conteúdo de polimorfismo e foi realizada a correlação entre as matrizes de distância. Os índices de similaridade genética obtidos e os resultados da análise da variância molecular indicaram uma similaridade maior entre as populações do que dentro das mesmas. O conhecimento da distribuição da diversidade genética neste germoplasma é extremamente importante para determinar estratégias que serão utilizadas no programa de melhoramento de milho doce.*

Palavras-chave: Diversidade genética, germoplasma, marcadores moleculares, *Zea mays*.

REFERENCES

Barbosa-Neto JF, Hernández, CM, O'Donoughue, LS and Sorrells ME (1997) Precision of genetic relationship estimates based on molecular markers. **Euphytica 98**: 59-67.

Bartish IV, Jeppsson N and Nybom H (1999) Population genetic structure in the dioecious pioneer plant species *Hippophae rhamnoides* investigated by random amplified polymorphic DNA (RAPD) markers. **Molecular Ecology 8**: 791-802.

Chalmers KJ, Waugh R, Sprent JI, Simons AJ and Powell W (1992) Detection of genetic variation between and among populations of *Gliricidia sepium* and *G. maculata* using RAPD markers. **Heredity 69**: 465-472.

Chan KF and Sun M (1997) Genetic diversity and relationships detected by isozyme and RAPD analysis of crop and wild species of *Amaranthus*. **Theoretical and Applied Genetics 95**: 865-873.

Chin ECL, Senior ML, Shu H and Smith JSC (1996) Maize simple repetitive DNA sequences: abundance and allele variation. **Genome 39**: 866-873.

Chtourou-Ghorbel N, Lauga B, Combes D and Marrakchi M (2001) Comparative genetic diversity studies in the genus *Lathyrus* using RFLP and RAPD markers. **Lathyrus Lathyrism Newsletter 2**: 62-68.

Crossa J, Hernandez CM, Bretting P, Eberhart SA and Taba S (1993) Statistical genetic considerations for maintaining germplasm collections. **Theoretical and Applied Genetics 86**: 673-678.

Edwards K, Johnstone C and Thompson C (1991) A simple and RAPD method for the preparation of plant genomic DNA for PCR analysis. **Nucleic Acids Research 19**: 1349.

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

Fahima T, Röder MS, Wendehake K, Kirzhner VM and Nevo E (2002) Microsatellite polymorphism in natural populations of wild emmer wheat, *Triticum dicoccoides*, in Israel. **Theoretical and Applied Genetics 104**: 17-29.

Jaccard P (1908) Nouvelles recherches sur la distribution florale. **Bulletin de la Société vaudoise des sciences naturelles 44**: 223-270.

Juaréz-Munõz J, Carrillo-Castañeda G, Arreguin R and Rubluo A (2002) Inter- and intra-genetic variation of four wild populations of Prosopis using rapd-pcr fingerprints. **Biodiversity and Conservation 11**: 921-930.

Liu ZW, Biyashev RM and Saghai-Maroof MA (1996) Development of simple sequence repeat DNA markers and their integration into a barley linkage map. **Theoretical and Applied Genetics 93**: 869-876.

Lynch M and Milligan BG (1994) Analysis of population genetic structure with RAPD markers. **Molecular Ecology 3**: 91-99.

Mailer RJ, Wratten N and Vonarx M (1997) Genetic diversity amongst Australian canola cultivars determined by randomly amplified polymorphic DNA. **Australian Journal of Experimental Agriculture 37**: 793-800.

Manifesto MM, Schlatter AR, Hopp HE, Suárez EY and Dubcovsky J (2001) Quantitative evaluation of genetic diversity in wheat germplasm using molecular markers**. Crop Science 41**: 682-690.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. **Cancer Research 27**: 209-220.

Mengoni A, Gori A and Bazzicalupo M (2000) Use of RAPD and microsatellite (SSR) variation to assess genetic relationships among populations of tetraploid alfalfa, *Medicago sativa*. **Plant Breeding 119**: 311-317.

Miller MP (1997) **Tools for population genetic analyses** (TFPGA)1.3: A windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by author.

Moretzsohn MC, Ferreira MA, Amaral ZPS, Coelho PJA, Grattapaglia D and Ferreira ME (2002) Genetic diversity of Brazilian oil palm (*Elaeis oleifera* H.B.K.) germplasm collected in the Amazon Forest. **Euphytica 124**: 35-45.

Oiki S, Kawahara T, Inoue K, Ohara M and Maki M (2001) Random Amplified Polymorphic DNA (RAPD) variation among populations of the insular endemic plant *Campanula microdonta* (Campanulaceae). **Annals of Botany 87**: 661-667.

Popi J, Rajnpreht L, Kannenberg LW and Pauls P (2000) Random amplified polymorphic DNA-based evaluation of diversity in the hierarchical, open-ended population enrichment maize breeding system. **Crop Science 40**: 619-625.

Reif JC, Melchinger AE, Xia XC, Warburton ML, Hoisington DA, Vasal SK, Srinivasan G, Bohn M and Frisch M (2003) Genetic distance based on simple sequence repeats and heterosis in tropical maize population. **Crop Science 43**: 1275-1282.

Reif JC, Xia XC, Melchinger AE, Warburton ML, Hoisington DA, Beck D, Bohn M and Frisch M (2004) Genetic diversity determined within and among CIMMYT maize populations of tropical, subtropical, and temperate germplasm by SSR markers. **Crop Science 44**: 326-334.

Rohlf FJ (2000) **NTSYS-pc: Numerical taxonomy and multivariate analysis system**. Exeter Software, New York.

Sanz-Cortés F, Badenes ML, Paz S, Íñiguez A and Llácer G (2001) Molecular characterization of olive cultivars using RAPD markers. **Journal of American Society of Horticultural Science 126**: 7-12.

Senior ML and Heun M (1993) Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. **Genome 36**: 884-889.

Silva W J (1983) Melhoramento do milho para maior uniformidade de resposta. In: **Simpósio sobre produtividade do milho**. SPM, Londrina, p. 211-213.

Sun GL, William M, Liu J, Kasha KJ and Pauls KP (2001) Microsatellite and RAPD polymorphism in Ontario corn hybrids are related to the commercial sources and maturity ratings. **Molecular Breeding 7**: 13-24.

Tosello GA (1987) Milhos especiais e seu valor nutritivo. In: Paterniani E and Viégas GP (eds.) **Melhoramento e produção do milho**. Fundação Cargill, Campinas, p. 375-409.

Tracy WF (1997) History, genetics and breeding of supersweet (shrunken 2) Sweet corn. **Plant Breeding Reviews 14**: 189-236.

Warburton ML, Xianchun X, Crossa J, Franco J, Melchinger AE, Frisch M, Bohn M and Hoisington D (2002) Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. **Crop Science 42**: 1832-1840.