Crop Breeding and Applied Biotechnology 9: 31-36, 2009
Brazilian Society of Plant Breeding. Printed in Brazil

Genetic divergence in popcorn genotypes using microsatellites in bulk genomic DNA

Tereza Aparecida da Silva[1], Ronald José Barth Pinto[2*], Carlos Alberto Scapim[2], Claudete Aparecida Mangolin[3], Maria de Fátima Pires da Silva Machado[3], and Misael S N Carvalho[4]

Received 25 June 2008

Accepted 04 November 2008

ABSTRACT – *The genetic diversity of 25 popcorn genotypes was estimated based on DNA bulks from 78 plants of each variety. The procedure involved 23 microsatellite loci distributed on 9 maize chromosomes. Clustering analysis according to the Tocher method and the hierarchical clustering procedures (nearest neighbor, furthest neighbor and Unweighted Pair-Group Method Using Arithmetic Averages - UPGMA) were performed. The cophenetic correlation coefficients indicated the UPGMA method as adequate to distinguish the varieties. The clusters suggested by the molecular analysis generally grouped genotypes with the same genealogy together. The genetic dissimilarity of the varieties Argentina, Chile, PA-091 and PR-023 was higher than of the others. Therefore, higher heterozygosity is expected in progenies from crosses with the other genotypes.*

Key words: *Zea mays* L., molecular markers, clustering methods.

INTRODUCTION

Popcorn breeding in Brazil is still incipient, and there are few national varieties and hybrids with high yield and quality. In view of the continental dimensions of Brazil, the selection of a relatively large number of genotypes is essential to supply varieties and hybrids adapted to the different regions of the country. This is a prerequisite to ensure high yield and quality for farmers who use high, medium or low levels of technology.

The strategy of heterosis exploitation is often used in maize breeding to produce superior hybrids derived from adequately combining inbred lines. Considering the large number of different hybrids derived from a relatively small number of lines, breeder must carefully examine the evaluation method of the genotypes, since the performance of multiple hybridation is laborious, time-consuming and expensive (Serafini et al. 2001, Paterniani and Campos 2005).

The use of molecular markers in maize breeding may be advantageous for previous analyses of the genetic distance between the parents, to reduce the number of crosses needed. Estimates of genetic distance can indirectly help predict the hybrid performance, based on the assumption that a higher molecular dissimilarity in the lines may represent greater heterozygosity and consequently a greater probability

[1]Programa de Pós-Graduação em Genética e Melhoramento (PGM), Universidade Estadual de Maringá (UEM), Maringá, PR, Brazil
[2]Departamento de Agronomia, Centro de Ciências Agrárias, Universidade Estadual de Maringá (UEM), Av. Colombo 5790, bloco J-45, 87020-900, Maringá-PR, Brazil. *E-mail: rjbpinto@uem.br
[3]Departamento de Biologia Celular e Genética (DBC), Universidade Estadual de Maringá (UEM), Maringá, PR, Brazil
[4]Programa de Pós-Graduação em Agronomia (PGA), Universidade Estadual de Maringá (UEM), Maringá, PR, Brazil

of generating superior hybrids (Serafini et al. 2001). The molecular characterization of the genetic diversity in germplasm can provide breeders with useful information for the identification and selection of parents (Pinheiro et al. 2003). The efforts can be focused to prioritize the most promising crosses.

This study is part of a popcorn breeding program of the Universidade Estadual de Maringá - PR, to evaluate the genetic diversity in popcorn varieties based on microsatellite markers (SSR), using genetic dissimilarity estimates of the genotype of interest. The study aimed to identify consistent genotype groups, to confirm heterotic groups for a practical exploration of heterosis of important traits for popcorn cultivation.

MATERIAL AND METHODS

Twenty-five popcorn genotypes were analyzed, most of which have a relatively broad genetic base, that is, are open-pollinated varieties (Table 1).

Table 1. Popcorn genotypes evaluated

Bulk	Genotype	Observation
1	Jade	hybrid
2	Angela	open-pollinated variety
3	Zélia	hybrid
4	Composite V1	open-pollinated variety
5	Iguatemi 2	open-pollinated variety
6	Composite V2	open-pollinated variety
7	Composite Barreto	open-pollinated variety
8	Composite Matheus	open-pollinated variety
9	Synthetic P1	open-pollinated variety
10	Composite Fracaro	open-pollinated variety
11	Iguatemi I	open-pollinated variety
12	Composite Aelton	open-pollinated variety
13	PR-023	open-pollinated variety
14	IAC-112	hybrid
15	IAC-112 F2	open-pollinated variety
16	Argentina	-
17	UNB-2	open-pollinated variety
18	PA-091	open-pollinated variety
19	Chile	-
20	Viçosa	open-pollinated variety
21	CMS-42	open-pollinated variety
22	SE-013	open-pollinated variety
23	Laranjeiras do Sul	open-pollinated variety
24	UEM – M2	open-pollinated variety
25	IAC-125	hybrid

For the extraction of DNA samples and establishment of bulks, 0.15g leaves were collected from each of the 78 plants of each genotype, about 20 days after sowing. The leaves of bulk constituents were sprayed with liquid nitrogen, and 0.30g of this genotype was used for DNA extraction. The genomic DNA was isolated according to the methodology described by Hoisington et al. (1994), with minor modifications. The amount of extracted DNA analysis was quantified by gel analysis in 0.8% agarose, prepared with TAE buffer at a concentration of 1X pH 8.0 (0.04 M Tris-acetate and 0.001 M EDTA) (Hoisington et al., 1994), at 80 V. Standard phage l DNA solutions, at concentrations of 50, 100 and 150 ng, were used for quantification. For visualization, the gel was stained in ethidium bromide prepared with 0.5 g mL^{-1} and photographed with an EDAS 290 UV transilluminator using a Digital Imaging System Kodak 1D 3.5.

To choose the polymorphic markers, 100 previously mapped maize microsatellites were evaluated, from the site http://www.maizegdb.org/ssr.php of the Maize Genetics and Genomic Database. The PCR was performed in 0.2 mL microplates, using a thermal cycler (Techne TC-512). For DNA amplification and primer selection, we used 25 ng of two DNA samples, with 2.0 mL of 10X reaction buffer, plus 2.0 mM MgCl2, 0.1 mM of each dNTP, 1 unit Taq-DNA polymerase, and 0.2 mM F and R primers in a final volume of 20 ìL. Based on the results with the two DNA samples, 23 primers were selected (UMC 2293, UMC 2196, UMC 2262, UMC 2245, UMC 2281, UMC 1336, UMC 2292, UMC 1071, UMC 1736, UMC 1636, UMC 2343, BNLG 1063, UMC 1241, UMC 1292, BNLG 2295, UMC 2080, UMC 1664, UMC 2116, BNLG 1083, UMC 1077, UMC 1363, UMC 1590, and MMC 0501). The microsatellites were initially amplified by Touchdown PCR (Don et al. 1991). However, a program with a specific annealing temperature of 50 ºC was used for the primers UMC 1664, UMC 2116 and UMC 2245. The amplification products were separated on 4% agarose gel (MS-8) and 10% denaturing polyacrylamide gel. The agarose gel was prepared with 50% agarose (MS-8) and 50% standard agarose using 0.5 X TBE buffer (44.5 mM Tris, 44.5 mM boric acid and 1 mM EDTA). Electrophoresis was run at 60 V for 4 hours. The gels were stained in 0.5 ìg mL^{-1} ethidium bromide. The amplified fragment size was determined by the molecular marker weight of a 100 pb DNA Ladder (Invitrogen) and the gels photographed

as above. The 10% denaturing polyacrylamide gel (29 acrylamide: 1 bisacrylamide) was prepared with 6M urea, 1X TBE buffer. Electrophoresis was performed at 325 V, for about 3.5 hours. The gels were silver-stained, according to the methodology described by Sanguinetti et al. (1994), at a silver nitrate concentration of 0.2 instead of 0.1%.

The genetic distance between the two populations (I and J) was estimated by the complement of the genetic similarity index used by Yao (2007). The dissimilarity between pairs of items was estimated by:

$$D_{ij} = 1 - GS = 1 - \left\{ 2\left[\frac{N_{ij}}{N_i + N_j} \right] \right\}$$

where N_{ij} is the number of alleles common to both populations for a particular primer, N_i is the total number of SSR alleles in population i and N_j is the total number of alleles in population j.

The matrix of D_{ij} values, underlying the dendrograms for cluster analysis was calculated using software Genes, version 4.1 (Cruz 2008). The same program Genes (Cruz 2008) was used to run three hierarchical grouping methods: method of simple linkage or nearest neighbor; method of complete linkage or furthest neighbor and the method of mean distances (Unweighted Pair-Group Method Using Arithmetic Averages - UPGMA). Dias (1998) described the calculations required for the methods mentioned above in detail. The efficiency of the grouping methods was compared based on the estimate of cophenetic correlation coefficients (CCC) (Sokal and Rohlf 1981). In parallel, for a greater efficiency of the analysis, the matrix of D_{ij} values was also used to group the genotypes by Tocher optimization (Rao 1952, Cruz and Carneiro 2006), as an alternative to the cluster analysis, often used in the study of genetic divergence in accessions.

RESULTS AND DISCUSSION

The quantification of genomic DNA indicated that the amount of DNA ranged from 25 to 100 ng ìL^{-1}, a sufficient amount for the analysis.

The selection of microsatellites was based on the complementarity and reproducibility of the primers, verified by annealing, and on the definition degree of the gel bands and the presence of different alleles. Of the 100 primers tested, 41 were polymorphic and 59 monomorphic. From among the 41 polymorphic primers, 23 were selected to investigate the genetic variability in the 25 bulks of the populations studied. A total of 100 alleles was detected among the varieties by the 23 microsatellite primers distributed on 9 chromosomes. The number of alleles per locus ranged from 2 to 7; the highest number of alleles was detected by primer BNLG 1083.

The use of genotype grouping methods is one of the best options for the analysis and interpretation of data contained in a genetic distance matrix (Cruz and Carneiro 2006), which represents an important stage of germplasm evaluation (Vasconcelos et al. 2007). Since no hierarchical method is fully applicable to all data types (Everitt 1993), it would be unwise to use a single grouping method.

Based on a literature review, Dudley (1994) stated that for applications related to genetic improvement, the method of mean distances (UPGMA) can be considered superior to the methods of nearest and furthest neighbor. This method has been recommended by some authors (Sneath and Sokal 1973, Romesburg 1984, Mohammadi and Prasanna 2003). In our study, three hierarchical grouping methodologies were considered (the nearest neighbor, furthest neighbor and UPGMA method). A cophenetic analysis was performed to identify the most suitable method, by which the cophenetic value of the UPGMA method was determined at 0.80. Dias (1998) found that correlation cophenetic values of over 0.80 indicate a good fitting of the dendrogram and the genetic distance matrix. The cophenetic values detected for the nearest-neighbor (0.78) and the furthest-neighbor clustering (0.61) did not meet this criterion.

The interpretation of the dendrogram constructed by the UPGMA method (Figure 1) depends on the cut-off point. If the cut-off line is drawn at a distance of 0.34, four groups are formed, the first holding the vast majority of the genotypes and the others with one or two genotypes. For the more distant genotypes, results indicate a differentiated integration of the varieties 16 (Argentina), 19 (Chile), 13 (PR-023) and 18 (PA-091).

The application of the Tocher method resulted in the partitioning of genotypes into four groups, where the first included most of the genotype (Table 2). The varieties 16 (Argentina) and 19 (Chile) formed a separate group, while 18 (PA-091) and 13 (PR-023) are isolated in their own groups. The results were similar to those

observed with the UPGMA method assuming a cut-off point in the position mentioned above.

The way the dissimilarity measures were obtained may have influenced the distance matrix. DNA bulks are being used systematically (Kongkiatngam et al. 1996, Ornella et al. 2001, Parentoni et al. 2001, Yao et al. 2007). The savings in time, labor and research resources justify the use of bulks, as opposed to the efforts invested in studies that require the genotyping of hundreds or thousands of individual plants. However, it is worth emphasizing that, when in this study the strategy of DNA bulks was used, only the situations of the presence of alleles in one population and lack of the same alleles in other populations, using the same primer, contributed effectively to calculate the distances and the subsequent formation of groups, with a tendency to underestimate the genetic distances.

The implementation of both methods (UPGMA and Tocher) indicated a low number of clusters, because many genotypes were grouped together, demonstrating that the genetic similarity in most varieties is relatively great, which explains the inclusion of so many genotypes in the same group.

The Tocher optimization was reapplied to investigate the possibility of a great genetic similarity of the majority of varieties, this time including only the varieties of the first group (composites V1, V2, Aelton, Barreto, Fracaro and Matheus, besides Iguatemi 1, Iguatemi 2, IAC-125, Viçosa, synthetic cultivar P1, UNB-2, UEM-M2, Zélia, CMS-42, Laranjeiras do Sul, Angela,

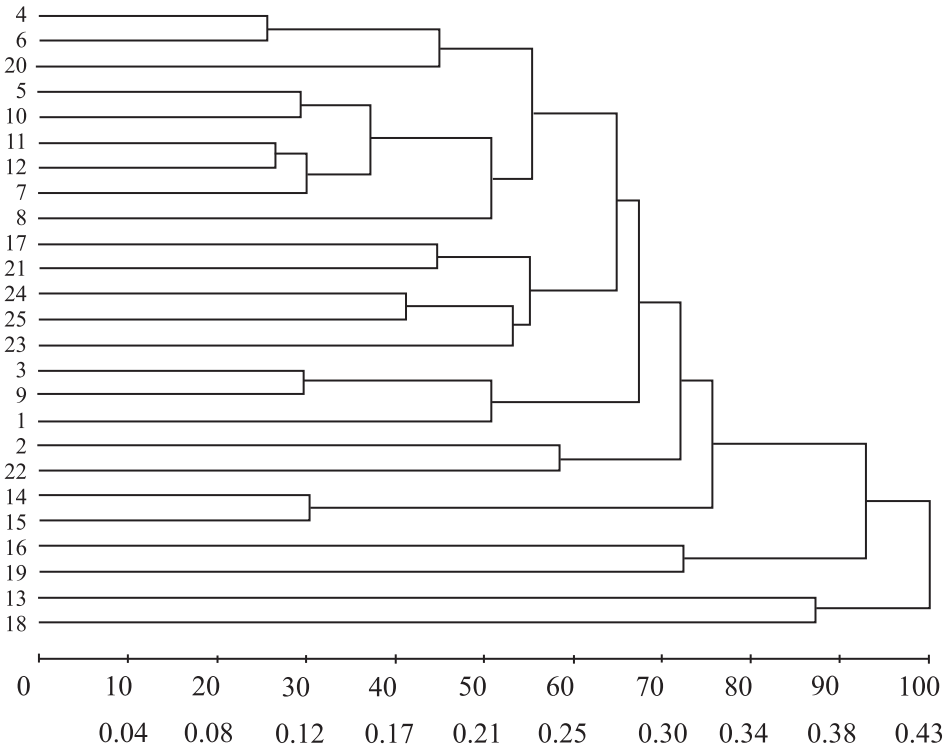

Figure 1. Clustering of 25 popcorn varieties by the UPGMA method

Table 2. Grouping of 25 popcorn varieties by the Tocher method

Group	Genotypes
1	4 6 12 7 11 5 10 25 8 20 9 17 24 3 21 23 2 15 1 14 22
2	16 19
3	18
4	13

IAC-112 F2, Jade, IAC-112 and SE-013). In this analysis without the most dissimilar varieties, the large group was split in six subgroups, the first compound with the varieties V1, Composite V2, Composite Aelton, Composite Barreto, Iguatemi 1, Iguatemi 2, Composite Fracaro, IAC-125 compound and Matheus. The second group contained the varieties Zélia, synthetic P1, EMU-M2 and Jade. The third consisted of the varieties IAC-112 and IAC-112 F2 and the fourth of UNB-2, CMS-42 and Viçosa. The fifth comprised Angela and SE-013 and the sixth the variety Laranjeiras do Sul.

The similarities between the group composition obtained by the Tocher reapplication and the groups constituted by UPGMA are interesting, indicating a new cut-off point (at about 0.26). The groups identified by the different methods are therefore consistent.

The information consistency and level of the variety groups can also be evaluated in relation to the correspondence between molecular data and pedigree of the genotype under study. For example, except for accessions 8 and 20, the first group contains only accessions with variety Zélia in the genealogy. Four accessions of this group have both Zélia and IAC112 in their pedigree (accessions 4, 5, 6 and 10). All other groups are formed by individuals with unknown pedigree.

The relationship between the varieties 16 (Argentina) and 19 (Chile) can not be accurately evaluated, except on a molecular basis, since the pedigree of both is unknown. The cluster among the genotypes 13 (PR-023) and 18 (PA-091) can be justified because both are open-pollinated population derived from American hybrids, somewhat distant from the other groups.

The correlation between the groups revealed by the UPGMA dendrogram and genealogy of the varieties under study showed that several clusters indicated by molecular analysis grouped genotypes of the same parental source together. The same is true for the Tocher method, especially after the subgrouping of the first cluster.

A detailed analysis of the results indicated that the genetic basis of most varieties evaluated consists of few genotypes, with predominance of Zélia, IAC-112, Angela and CMS-42.

The use of molecular markers can provide insights on the inter-population heterosis in the cases where the genotypes that are genetically more distant from each other can be used in programs of reciprocal recurrent selection aimed at the synthesis of contrasting lines that are capable of generating superior hybrids.

The consistency of results involving the different varieties (13 (PR-023), 16 (Argentina), 18 (PA-091) and 19 (Chile)) suggests that the performance of inter crosses between these and the other genotypes is a promising alternative to obtain heterozygosity as a strategy for the exploitation of heterosis.

ACKNOWLEDGEMENTS

The authors thank the Coordenação e Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the scholarship.

Divergência genética entre genótipos de milho-pipoca utilizando microssatélites em bulk de DNA genômico

RESUMO - *A divergência genética entre 25 genótipos de milho-pipoca foi estimada utilizando-se bulks de DNA genômico formados por 78 plantas de cada variedade através da utilização de 23 primers microssatélites distribuídos em 9 cromossomos. Foram realizadas análises de agrupamento pelo método de Tocher e pelos métodos hierárquicos do vizinho mais próximo, vizinho mais distante e UPGMA. Os coeficientes de correlação cofenética indicaram que o método UPGMA foi adequado para discriminar as variedades. Os clusters indicados pela análise molecular geralmente agruparam materiais de mesma origem parental. As variedades Argentina, Chile, PA-091 e PR-023 manifestam grande dissimilaridade genética em relação às integrantes de outros grupos, sendo esperada maior heterozigose nas progênies que possam formar por cruzamentos com as demais variedades.*

Palavras-chave: *Zea mays* L., marcadores moleculares, métodos de agrupamento.

REFERENCES

Cruz CD and Carneiro PCS (2006) **Modelos biométricos aplicados ao melhoramento genético**. 2.ed. Editora UFV, Viçosa, 585p.

Cruz CD (2008) **Programa genes: diversidade genética**. Editora UFV, Viçosa, 278p.

Dias LAS (1998) Análises multidimensionais. In: Alfenas AC **Eletroforese de isoenzimas e proteínas afins**. Editora UFV, Viçosa, p. 405-475.

Don RH, Cox PT, Wainwright BJ, Baker K and Mattick JS (1991) Touchdown PCR to circumvent spurious priming during gene amplification. **Nucleic Acids Res 19:** 4008-4008.

Dudley JW (1994) Comparison of genetic distance estimators using molecular marker data. In: Symposium analysis of molecular data. **Crop Science Society of America**. American Society for Horticultural Sciences, p.3-7.

Everitt BS (1993) **Cluster analysis**. University Press, Cambridge, 170p.

Hoisington D, Khairallah M and González Léon D (1994) Laboratory Protocols: **CIMMYT Applied Molecular Genetics Laboratory**. D.F.: CIMMYT, Mexico, 50p.

Kongkiatngam P, Waterway MJ, Coulman BE and Fortin MG (1996) Genetic variation among cultivars of red clover (*Trifolium pratense* L.) detected by RAPD markers amplified from bulk genomic DNA. **Euphytica 89:** 355-361.

Mohammadi SA and Prasanna BM (2003) Analysis of genetic diversity in crop plant – Salient statistical tools and considerations. **Crop Science 43:** 1235-1248.

Ornella L, Schlatter AR, Vonhaniel Niethammer F, Manifesto MM, Eyherabide G, Suárez EY and Acevedo A (2001) **Empleo de marcadores moleculares de ADN para identificar genótipos de maíz (*Zea mays* L.).** Available at: http://www.redbio.org/portal/encuentros/enc_2001/ on 08 Fev. 2008.

Parentoni SN, Magalhães JV, Pacheco CA, Santos MX, Abadie T, Gama EEG, Guimarães PEO, Meirelles WF, Lopes MA, Vasconcelos MJV and Paiv, E (2001) Heterotic groups based on yield-specific combining ability data and phylogenetic relationship determined by RAPD markers for 28 tropical maize open pollinated varieties. **Euphytica 121:** 197-208.

Paterniani E and Campos MS (2005) Melhoramento do milho. In: Borém A **Melhoramento de Espécies Cultivadas**. Editora UFV, Viçosa, p.491-552.

Pinheiro JB, Zucchi MI, Teles FL and Ázara NA (2003) Diversidade genética molecular em acessos de açafrão utilizando marcadores RAPD. **Acta Scientiarum 25:** 195-199.

Rao CR (1952) **Advanced statistical methods in biometric research**. John Wiley & Sons, New York, 390p.

Rohlf FJ and Sokal RR (1981) Comparing numerical taxonomic studies. **Systematic Zoology 30:** 459-490.

Romesburg HC (1984) **Cluster analysis for researchers**. Lifetime Learning, California, 334p.

Sanguinetti CJ, Dias Neto E and Simpson AJG (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. **Biotechniques 17:** 915-919.

Serafini LA, Barros MN and Azevedo JL (2001) **Biotecnologia na agricultura e na agroindústria**. Livraria e Editora Agropecuária Ltda., Guaíba, 463p.

Sneath PHA and Sokal RR (1973) **Numerical taxonomy**. W.R. Freeman, San Francisco, 573p.

Vasconcelos ES, Cruz CD, Bhering LL and Resende Júnior MFR (2007) Método alternativo para análise de agrupamento. **Pesquisa Agropecuária Brasileira 42:** 1421-1428.

Yao Q, Yang K, Pan G and Rong T (2007) Genetic diversity of maize (*Zea mays* L.) landraces from Southwest China based on SSR data. **Journal of Genetics and Genomics 34:** 851-860.