Inference of genetic diversity in popcorn S_3 progenies

**G.F. Pena, A.T. do Amaral Jr, R.M. Ribeiro, H.C.C. Ramos,
M.S.B. Boechat, J.S. Santos, G.S. Mafra, S.H. Kamphorst, V.J. de Lima,
M. Vivas and G.A. de Souza Filho**

Universidade Estadual do Norte Fluminense Darcy Ribeiro,
Campos dos Goytacazes, RJ, Brasil

Corresponding author: A.T. do Amaral Jr
E-mail: amaraljr@pq.cnpq.br

ABSTRACT. Molecular markers are a useful tool for identification of complementary heterotic groups in breeding programs aimed at the production of superior hybrids, particularly for crops such as popcorn in which heterotic groups are not well-defined. The objective of the present study was to analyze the genetic diversity of 47 genotypes of tropical popcorn to identify possible heterotic groups for the development of superior hybrids. Four genotypes of high genetic value were studied: hybrid IAC 125, strain P2, and varieties UENF 14 and BRS Angela. In addition, 43 endogamous S_3 progenies obtained from variety UENF 14 were used. Twenty-five polymorphic SSR-EST markers were analyzed. A genetic distance matrix was obtained and the following molecular diversity parameters were estimated: number of alleles, number of effective alleles, polymorphism information content (PIC), observed and expected heterozygosities, Shannon diversity index, and coefficient of inbreeding. We found a moderate PIC and high diversity index, indicating that the studied population presents both good discriminatory ability and high informativeness for the utilized markers. The dendrogram built based on the dissimilarity matrix indicated six

distinct groups. Our findings demonstrate the genetic diversity among the evaluated genotypes and provide evidence for heterotic groups in popcorn. Furthermore, the functional genetic diversity indicates that there are informative genetic markers for popcorn.

Key words: EST-SSR; Topcross; *Zea mays* L. var. Everta

INTRODUCTION

For successful genetic breeding of popcorn, it is important that the end product meets the demands of farmers, grain-producing companies, and consumers (Silva et al., 2010; Ribeiro et al., 2012). To meet the requirements of all the three groups, it is fundamental to understand the crop characteristics that directly or indirectly influence its commercial acceptance; these characteristics include high productivity indices for cultivars, expansive grain capacities, and optimal organoleptic properties such as tenderness, flavor, aroma, and popcorn color (Alexander and Creech, 1977; Moterle et al., 2012).

By the end of the 20th century, little genetic breeding had been conducted on Brazilian popcorn populations, making it difficult to obtain strains due to high inbreeding depression, high susceptibility to spot blotch and wilt, and low popping expansion rates (Andrade, 1996; Coimbra, 2000). Until December 2015, only 74 records of popcorn cultivars, including hybrids, strains, and varieties, were available in the records of the National System of Plant Variety Protection (Sistema Nacional de Proteção de Cultivares) and the Ministry of Agriculture, Livestock, and Supply (Ministério da Agricultura, Pecuária e Abastecimento). Private companies own 81.08% of the registered cultivars, as follows: Yoki (19 cultivars), Agristar (12 cultivars), Pipolino (7 cultivars), General Mills (7 cultivars), Sementes Boa Esperança (4 cultivars), Leandro Lodéa (4 cultivars), Seedco (3 cultivars), DSMM/CATI (2 cultivars), Feltrin and Vidasul (1 cultivars), and ATS (1 cultivars). Only two research institutions hold approximately 18.92% of the records: IAC (12) and State University of Northern Rio de Janeiro (UENF) (2 cultivars). These data reveal an extensive consumer market, exploited mainly by multinational companies. In contrast, the research institutions in Brazil, especially the public ones, have contributed little to the availability of quality seeds to small- and medium-scale producers.

Given the present situation of Brazilian cultivars and the great technological and scientific potential of breeding programs, precise and well-executed medium-term breeding strategies can be easily executed. A better understanding of the existing genetic variability in the genotypes that make up the germplasm bank of a breeding program would facilitate more accurate classification of the germplasm into heterotic groups. Owing to heterosis, evaluation of hybrid combinations raises the chances of obtaining consistent increases in the mean values of the traits of interest (Hallauer et al., 2010). In general, the current state of the genetic breeding of popcorn in Brazil is closely associated with the history of the crop, and it has suffered because heterotic groups have not been established. This situation contrasts that in the USA, where Larish and Brewbaker (1999) identified three heterotic groups at the beginning of the North-American popcorn breeding program. Their report spurred the development of hybrids with a high popping expansion rate. Miranda et al. (2008) studied genetic variability to identify heterotic groups in Brazilian popcorn populations. For the expansion trait, they found that both heterosis and genetic variability were lower in the populations than in commercial cultivars. This situation makes it difficult to directly obtain commercial hybrids based on local

varieties, as these have poor performance in terms of popping expansion.

Various strategies have been implemented in popcorn breeding to significantly increase the values of traits of commercial interest, such as productivity and expansion. In tropical countries, like Brazil, the common breeding practices include crossing local cultivars (Andrade, 1996), use of advanced generations of American hybrids (Coimbra, 2000), and, recently, the top-cross methodology, which exploits the combining ability among different testers (Hallauer et al., 2010; Barreto et al., 2012; Rodovalho et al., 2012). In temperate countries, the practice of crossing popcorn and flint corn has been adopted, mostly aimed at popping expansion, in addition to backcrossing of elite strains of common corn.

Hallauer et al. (2010) reported that corn populations displaying potential for use as parents in intra-population breeding programs must necessarily present high mean values and a wide genetic variability for the traits of interest. Based on this conclusion, since 1998, the popcorn breeding program of UENF has been developing a population with favorable agronomic traits, such as productivity and expansion. This population, derived from an indigenous compound, was initially crossed with an American popcorn variety (SAM) with high expansion rate. The filial generation of this cross was then selected and crossed with a popcorn variety resistant to spot blotch. After three cycles of backcrosses with the American cultivar and five cycles of recurrent intra-population selection, totaling 14 years of research, a new variety was established. It was eventually named UENF 14 and was recommended to producers from the North and Northwest Fluminense regions of Brazil (Amaral Júnior et al., 2013).

One of the important theories in quantitative genetics postulates that at any degree of dominance greater than zero, the heterosis manifested in an individual stems from the interaction of the allelic frequency between its parents, with a positive correlation existing between genetic divergence and heterosis (Falconer and Mackay, 1996). Based on this correlation, estimation of genetic diversity between strains using molecular markers would help reduce the number of studies using pollination, as well as allowing identification of heterotic groups and direct crosses for the generation of superior hybrids in terms of productivity and vigor (Laborda et al., 2005). The use of molecular markers is of great relevance to studies on genetic diversity between strains developed in breeding programs. Molecular marker analysis also provides new data on existing genetic variability, at the DNA level, by grouping the genotypes according to their genetic distances or dissimilarity measures, thus allowing determination of the best strategy for establishing promising hybrids (Kumar et al., 2009; Zhang et al., 2014).

Among the molecular markers available today, simple sequence repeats (SSR) have great applicability owing to their technical simplicity (e.g., use of polymerase chain reaction, PCR) and great resolving power (allowing detection of polymorphisms between fragments with up to one pair of bases). SSR markers are also favored for their high level of polymorphism, high reproducibility, cost-efficiency (requiring the availability of primers at specific sites), codominance, and site-specificity (Mohammadi and Prasanna, 2003; Kumar et al., 2009). The SSR markers originating from expressed sequence tags (EST) are also known as SSR-EST or genic SSR. Genic SSRs have the advantage over genomic SSRs of being intrinsically bound to genome coding sequences (Eujayl et al., 2002). Therefore, these markers may be developed based on transcribed regions of the genome, which can increase the chances of "marking" traits of interest and thus increasing the efficiency of marker-assisted selection (Yu et al., 1999; Zhang et al., 2014).

SSR-EST markers have been widely employed as a powerful genetic-molecular tool owing to their high transferability and close association with genes of known function. In addition, they can be developed at a low cost because of the availability of public databases

(Zhang et al., 2014). Use of genic molecular markers (EST-SSR) can provide information on both the degree of genetic diversity and the phenotype of a population. In the case of base populations for plant breeding programs, e.g., popcorn, such information becomes important to follow up with other work programs that are of interest.

The present study aimed to explore heterosis among 43 S_3 progenies and four elite varieties of popcorn by estimating the genetic diversity using SSR-EST molecular markers. With the results obtained, we intend to identify different heterotic groups for practical utilization of heterosis in future breeding strategies for this crop.

MATERIAL AND METHODS

Genetic material

Forty-three S_3 progenies originating from the UENF 14 popcorn variety and four genotypes with high genetic value, used as testers in top-cross experiments from the UENF breeding program, were utilized. The additional four genotypes used were two open-pollination varieties (BRS Angela and UENF 14), one top-cross hybrid (IAC 125), and one strain (P2).

S_3 progeny were obtained by three cycles of self-pollination: S_0 (UENF 14), S_1, and S_2. Each S_3 progeny was derived from a single ear remaining. The BRS Angela variety is the result from six cycles of recurrent intra-population selection in the popcorn compound CMS 43 from Embrapa Maize and Sorghum (Pacheco et al., 2001). UENF 14 (Amaral Júnior et al., 2013) is an open-pollination variety already adapted to and stable in the North and Northwest Fluminense regions of Brazil. IAC 125 is a top-cross hybrid (simple hybrid *vs*. variety) resulting from the cross of the simple modified hybrid IAC 112 and a synthetic hybrid obtained from a North American hybrid. Strain P2, named so by the breeding program of the State University of Maringá (UEM), was obtained after seven self-pollination cycles (S_7) of the CMS-42 compound, belonging to the germplasm bank of Embrapa Maize and Sorghum.

Genomic DNA extraction

Young leaves from each plot, corresponding to each S_3 progeny, were bulk-harvested, properly identified, covered with aluminum foil, and stored in an ultra-freezer (-80°C) until extraction. The genomic DNA was extracted by the CTAB method (Doyle and Doyle, 1990), with modifications as suggested by Daher et al. (2002). After the extraction, a 1% agarose gel electrophoresis was run, to evaluate the integrity of the material, as well as to quantify (in μg/μL) the extracted DNA. This was done using as comparison the High DNA Mass Ladder marker (Invitrogen, Carlsbad, Califórnia, USA), which is appropriate to estimate the DNA mass (μg/μL).

Selection of SSR-EST primers

The microsatellite markers (SSR-EST) were identified in the literature and in the data set of the National Center of Biotechnology Information (NCBI), based on sequences developed and mapped by Sharopova et al. (2002) for *Zea mays* L. During the NCBI marker search we sought to select a significant amount of primers in each of the ten species linking groups, for appropriate sampling of the genome. Initially, 228 pairs of primers were selected and subjected to annealing tests, using to the Primer Select application of Dnastar (www.dnastar.com/).

Polymerase chain reaction

To test the PCR conditions, a temperature gradient of 60 to 71.5°C was performed in which the optimal annealing temperatures for each of the synthesized markers (228 primer pairs) were selected. Subsequently, a screening was performed to select the primers based on the degree of polymorphism and quality of amplifications.

The amplification reactions were performed in a final volume of 10 µL per reaction containing 4.55 µL ultrapure water, 1 µL 10X reaction buffer, 0.6 µL 25 mM MgCl$_2$, 1 µL 2 mM dNTP, 0.8 µL 5 µM R+F primer, 0.05 µL Taq DNA polymerase (5 U/µL), and 2 µL DNA (5 ng/µL). Amplifications were performed using the respective ideal annealing temperatures of the markers in an Applied Biosystems/Veriti96 well thermocycler. A 35-cycle program was utilized, including the following temperatures and times: 94°C for 4 min (initial denaturation); 94°C for 1 min (cyclic denaturation); specific temperature of each marker for 1 min (annealing); 72°C for 3 min (cyclic extension); 72°C for 7 min (final extension); and 4°C until terminated.

Electrophoresis

The amplified DNA was subjected to electrophoresis with a current of 350 to 400 mA on 3% MetaPhor agarose gel, immersed in a TAE buffer [90 mM Tris-Acetate (pH 8.0) + 10 mM EDTA], and stained with GEL RED™. The results were then visualized using the MiniBis Progel documentation system, for analysis of the pattern of amplified bands in the four tester genotypes. To determine the size of the amplified fragments, a 250-bp ladder DNA marker was used. The primers that showed adequate amplification patterns in the tester genotypes were subjected to an additional analysis in the Fragment Analyzer capillary electrophoresis system. At this stage, 25 pairs of polymorphic primers were selected among the evaluated genotypes to be further tested in the 43 S$_3$ progenies. Information pertaining to the markers selected for the molecular analysis is given in Table 1.

Molecular data analysis

The 25 selected markers were analyzed in the entire population (in the four testers and the 43 S$_3$ progenies) and subjected to capillary electrophoresis, utilizing the Fragment Analyzer. A difference in amplified fragment size of at least five bp was utilized as a criterion for considering the alleles as distinct. The data obtained from the amplification of the SSR-EST markers were converted to a numerical code for each allele per locus. This numerical matrix was developed, assigning values of one up to the maximum number of alleles in the locus, as described in the following example: a locus with three alleles is represented as 11, 22, and 33 for the homozygous forms (A1A1, A2A2, and A3A3), and as 12, 13, and 23 for the heterozygous forms (A1A2, A1A3, and A2A3). Based on this numerical matrix, using Power Marker v. 3.25 (Liu and Muse, 2005), the number of alleles (N_A) and the number of effective alleles (N_E) for the group of genotypes in question (S$_3$ progenies) were estimated. In addition, the polymorphism information content (PIC) and the coefficient of inbreeding (F) were estimated. The genetic diversity (based on the Shannon index, I), as well as the observed (H_O) and expected (H_E) heterozygosities were obtained using POPGENE v.1.31 (Yeh et al., 1997).

The genetic dissimilarity matrix was estimated based on the Weighted Index proposed by Ramos et al. (2011). A dendrogram was constructed according to the dissimilarity values

using the hierarchical unweighted pair-group method average (UPGMA), aiming to illustrate the genetic relationships among the different S_3 progenies. The dendrogram was obtained by MEGA v. 5.0 (Tamura et al., 2011) and all statistical analyses were performed using Genes (Cruz, 2013).

Table 1. SSR-EST primers utilized in the molecular analyses with their respective locus, chromosomal location (Cr), centromere distance (cM), primer sequences, annealing temperature (Tm), and sequence motif.

Locus	Cr	cM	Sequence (5'→ 3')	Tm (°C)	Motif
umc2174	3	566.2	ACATAAATAAAACGTGTGCCGCAG GTACGTACGCAGCCACTTGTCAG	64.0	$(CGA)_4$
umc1153	5	607.7	CAGCATCTATAGCTTGCTTGCATT TGGGTTTTGTTTGTTTGTTTGTTG	62.5	$(TCA)_4$
umc1680	5	434.1	GGGGCTTATATGTCCCTTGAACTC TTAATAAAGGAGAGGGTGGGAACC	63.0	$(TC)_7$
umc1071	1	78.7	AGGAAGACACGAGAGACACCGTAG GTGGTTGTCGAGTTCGTCGTATT	63.0	$(TACGA)_5$
umc1780	3	11.4	CATGATGTACCCGCAACAAATG CTGTCCCCAGGTTGCTGTAGTAGT	63.0	$(ACC)_4$
umc1760	7	553.9	CATGTGGGTGTTAATAAGCAAGGG GCCTTGTTTGCTCTCTGAAACAAT	63.5	$(GA)_6$
umc1327	8	55.2	AGGGTTTTGCTCTTGGAATCTCTC GAGGAAGGAGGAGGTCGTATCGT	63.5	$(GCC)_4$
umc2165	6	448.9	AGAACACCAAATGGTGACGTTATGT CTAGCTCGTCTTCCCTGTGGTCT	63.0	$(TTC)_{12}$
umc1393	7	212	CCTTCTTCTTATTGTCACCGAACG GCCGATGAGATCTTTAACAACCTG	63.0	$(GTC)_4$
umc1450	7	270.3	ACTTTATTTAGCCCACGTCACTCG AGTATGACACGGGATTTTGCTGTT	63.0	$(AC)_{10}$
umc1221	5	329.5	AAACAGGCACAAAGCATGGATAG GCAACAGCAACTGGCAACAG	63.0	$(CT)_7$
umc1336	10	95.37	GTACAAATGATAAGCAAGGGGCAG CTCTGTTTTGGAAGAAGCTTTTGG	63.0	$(ACCAG)_5$
umc1415	8	119.56	GTGAGATATATCCCCGCCTTCC AGACTTCCTGAAGCTCGGTCCTA	63.0	$(GAC)_{10}$
umc1506	10	287.5	AAAAGAAACATGTTCAGTCGAGCG ATAAAGGTTGGCAAAACGTAGCCT	63.5	$(AACA)_4$
umc1594	3	685.1	CACTGCAGGCCACACATACATA GCCAGGGGAGAAATAAAATAAAGC	62.0	$(TA)_{10}$
umc1656	6	37.58	AGTTTTGACCGCGCAAAAGTTA GTACGAGCAGGCCATTAACCC	63.0	$(CGGT)_7$
umc2059	6	259.49	GGAAAAGGAGGAACAGTGTAAGCA AGCGTGATCAGACGTACAATGCTA	63.0	$(CAG)_8$
umc2152	3	738.7	TAGCTTCACCTGATGATCTTGCAC CCTTTGTCTTCCGCTATCTTCCTT	63.0	$(TG)_8$
umc2246	2	3.1	AGGCTCCAGCTCTAGGGGAGT GTGAACTGTGTAGCGTGGAGTTGT	63.0	$(CCT)_8$
umc2292	5	19.39	AGCAGAAGAGGACAAACCAGATTC ACTTCCGGCATGTCTTGTGTTT	62.5	$(CTGCCT)_4$
umc1448	2	232.5	ATCCTCTCATCTTTAGGTCCACCG CATATACAGTCTCTTCTGGCTGCTCA	63.5	$(GCT)_5$
gnl4	5	469.2	AGCAGAACGGCAAGGGCTACT TTTGGCACACCACGACGA	64.0	-
y1	6	36	CAAGAAGAGGAGAGGCCGGA TTGAGCAGGGTGGAGCACTG	64.5	$(CCA)_3$
umc1867	9	24.3	ATAAGCTCGTTGATCTCCTCCTCC TCTCGCTGTCCTTCGATTAGTACGG	63.5	$(TC)_8$
umc1380	10	16.7	CTGCTGATGTCTGGAAGAACCCT AGCATCATGCCAGCAGGTTTT	64.0	$(CTG)_5$

RESULTS AND DISCUSSION

Seventy-two alleles were identified for the 25 evaluated loci. The number of alleles per locus varied from two to five, with an average of 2.88 (Table 2). Of the 25 loci, seven (28%) were highly polymorphic, 15 (60%) were moderately polymorphic, and three (12%)

lacked informative content (Table 2). The maximum observed PIC value was 0.68 for locus *umc1680* and the lowest was attributed to locus *umc2246* (0.07). The average PIC for all loci was 0.40, indicating that the overall observed polymorphism level was moderate. Zhang et al. (2014) stated that several studies on diversity involving SSR-EST markers present alow-to-medium level of polymorphism, which is supported by the results observed in the present study. The cause of this low polymorphic rate has been attributed to the selection practiced against the possible variations in preserved regions of the SSR-EST markers, as stated by Scott et al. (2000).

Table 2. SSR-EST primers and their respective number of alleles (N_A), number of effective alleles (N_E), polymorphism information content (PIC), and putative function.

Locus	N_A	N_E	PIC	Putative function
umc2174	3	2.17	0.43	*NADPH-dependent reductase
umc1153	5	2.92	0.58	*Uncharacterized
umc1680	4	3.75	0.68	*Uncharacterized
umc1071	3	1.40	0.25	*Glutathione S-transferase12
umc1780	3	2.08	0.40	*Uncharacterized
umc1760	3	2.55	0.53	*Uncharacterized
umc1327	3	2.60	0.54	*U6 snRNA-associated Sm-like protein LSm4
umc2165	3	1.97	0.41	*Putative leucine-rich repeat receptor Kinase
umc1393	3	1.86	0.41	*Uncharacterized
umc1450	2	1.66	0.32	*Induced Stolen tip protein TUB8
umc1221	2	1.54	0.29	*Uncharacterized
umc1336	2	1.93	0.36	*Methionine adenosyltransferase isoform 1
umc1415	3	2.10	0.42	*Uncharacterized
umc1506	2	1.72	0.33	*Uncharacterized
umc1594	2	1.40	0.24	*Non-specific lipid-transfer protein 3
umc1656	2	1.28	0.19	*Uncharacterized
umc2059	3	2.08	0.45	*Uncharacterized
umc2152	3	2.53	0.53	*Uncharacterized
umc2246	2	1.08	0.07	*Uncharacterized
umc2292	5	2.57	0.54	*Uncharacterized
umc1448	3	2.26	0.46	*Uncharacterized
gnl4	2	1.94	0.36	*Uncharacterized
y1	5	2.68	0.55	*Chloroplast phytoene synthase
umc1867	2	1.86	0.35	*Uncharacterized
umc1380	2	1.99	0.37	*Uncharacterized
Average	2.88	2.08	0.40	
Total	72			

*Zea mays.

Absolute PIC values can indicate the discriminatory power of microsatellite loci, as they are directly related to the variability of the region for the analyzed genotypes. However, Warburton et al. (2002) reported that high discriminatory power is not always related to microsatellites with higher PIC values and that the composition of the population is one of the factors influencing the informative power of a marker. Of the seven highly polymorphic loci, two loci (*umc1327* and *y1*, with PIC values of 0.54 and 0.55, respectively) indicated a putative function in *Zea mays*: U6 snRNA-associated Sm-like protein LSm4 and chloroplast phytoene synthase.

For all analyzed loci, the mean H_O was lower than the H_E (Table 3), the former ranging from 0.20 (progeny 25) to 0.56 (IAC 125). The obtained H_E values were high, varying from 0.58 (progeny 47) to 0.70 (progenies 23, 37, and 45), with an overall mean of 0.65 (Table 3). Silva et al. (2009) used 30 SSR markers to evaluate the diversity and genetic structure of 31 popcorn populations in the UEM germplasm bank. They found that the mean H_O varied from

0.071 to 0.30, with genotype BOZM 260 (Cimmyt population) showing the highest proportion of heterozygous plants (0.30).

Table 3. Genetic diversity parameters of the four testers and S$_3$ progenies, based on the 25 SSR-EST loci tested.

Genotype	I	H_O	H_E	F
BRS Angela	0.99	0.40	0.61	0.35
IAC 125	1.03	0.56	0.59	0.06
P2	1.05	0.46	0.65	0.30
UENF14	1.18	0.44	0.67	0.35
1	1.13	0.29	0.66	0.57
4	1.12	0.38	0.65	0.43
6	1.11	0.29	0.65	0.56
9	1.05	0.28	0.60	0.54
10	1.10	0.33	0.64	0.49
11	1.15	0.28	0.65	0.58
12	1.15	0.28	0.65	0.58
13	1.18	0.24	0.68	0.65
14	1.15	0.32	0.68	0.53
15	1.02	0.24	0.60	0.61
17	1.20	0.28	0.69	0.60
18	1.06	0.28	0.62	0.55
19	1.23	0.28	0.69	0.60
20	1.12	0.24	0.62	0.62
21	1.03	0.28	0.61	0.55
22	1.14	0.32	0.67	0.53
23	1.30	0.33	0.70	0.53
24	1.13	0.26	0.66	0.61
25	1.04	0.20	0.63	0.69
26	1.15	0.24	0.68	0.65
27	1.13	0.29	0.66	0.56
28	1.10	0.42	0.63	0.35
29	1.17	0.32	0.69	0.55
30	1.05	0.32	0.65	0.52
32	1.14	0.28	0.67	0.59
33	1.06	0.36	0.62	0.42
34	1.08	0.24	0.63	0.63
35	1.16	0.28	0.66	0.58
36	1.14	0.33	0.67	0.51
37	1.24	0.24	0.70	0.66
38	1.15	0.36	0.68	0.48
39	1.17	0.28	0.66	0.58
40	1.05	0.29	0.64	0.55
41	1.10	0.52	0.65	0.20
42	1.14	0.28	0.67	0.59
43	1.05	0.44	0.61	0.28
44	1.13	0.40	0.66	0.41
45	1.24	0.36	0.70	0.49
46	1.07	0.48	0.63	0.25
47	1.03	0.29	0.58	0.51
48	1.20	0.28	0.67	0.59
49	1.06	0.28	0.62	0.55
50	1.14	0.36	0.66	0.46
Overall average	1.12	0.32	0.65	0.51

I = Shannon index; H_O and H_E = observed and expected heterozygosity, respectively; F = coefficient of inbreeding.

According to Amaral et al. (2013), the Shannon index (I), also known as the Shannon-Wiener diversity index, qualifies the degree of genetic diversity of a given genotype, which

increases as the index approaches unity. Based on the estimates of I, the analyzed genotypes showed a high degree of genetic diversity, ranging from 0.99 (BRS Angela) to 1.30 (progeny 23) and with an overall mean of 1.12 (Table 3). These results confirm a great genetic variability, especially among the studied S$_3$ progenies, and, consequently, in the UENF 14 variety. This broadens the possibilities in terms of exploratory potential of this population, and indicates the potential for significant genetic gains with the composition of future hybrids among S$_3$ progenies with strains obtained from the four testers. Galvão et al. (2015) evaluated the genetic diversity within and between populations of field corn (populations Cimmyt and Piranão) and obtained Shannon index estimates > 1, with an average of 1.04, for the Piranão population. These high values illustrate the efficiency of this index in studies aimed at increasing the genetic diversity of the target population.

Although popcorn belongs to the same species as field corn, comparative studies of the degree of genetic diversity in SSR loci between these two have revealed that popcorn has a smaller number of alleles per locus and lower mean heterozygosity values (Liu and Muse, 2005). A suggested explanation for the lower genetic diversity of popcorn than that of common corn is based on the hypothesis of the origin of popcorn. According to Erwin (1949), popcorn is a mutant of flint corn. Variations of flint corn probably appeared between the pre- and post-Columbian eras, influenced by the native American culture. However, these variations were recognized and propagated by the white man, specifically in North America.

The analysis of F found between testers and S$_3$ progenies revealed values ranging from 0.06 (IAC 125) to 0.69 (progeny 25), with an overall mean of 0.51 (Table 3). In population genetic studies, the estimate of F is considered one of the most important parameters. F estimates the balance between homozygotes and heterozygotes, and the mean values obtained are related to the allelic richness observed in the populations (Galvão et al., 2015). In breeding programs aimed at strain development, the analysis of F allows us to infer important information about genotype homozygosis level.

For comparison, the average H_O and F were 0.47 and 0.27, respectively for the four testers, and 0.31 and 0.53, respectively for the S$_3$ progenies (Table 3). The lower H_O value found for the S$_3$ progenies (0.31) is a result of a higher fixation of alleles, as indicated by the higher F (0.53) (Table 3).

In the cluster analysis (illustrated in the dendrogram in Figure 1), it could be observed that a distance of 0.20 (overall mean of the distance among all genotype pairs) was sufficient to separate the genotypes that were divided into six distinct groups (Figure 1). One of these groups contained 22 genotypes (21 S$_3$ progenies and tester BRS Angela), another group contained five progenies, three groups each contained four progenies (testers UENF 14 and IAC 125 were clustered separately within two of these groups), whereas the last and smallest group contained only two progenies. Apart from these groups, five progenies and tester P2 did not group with any other genotype. Variety BRS Angela was included in the first and largest group. This indicates that this variety shares a specific set of alleles with more than 21 S$_3$ progenies. Therefore, obtaining strains of genotypes belonging to this group aiming at the formation of heterotic groups would not be a very promising strategy.

The number of groups obtained, in relation to the number of genotypes evaluated, indicated that the sample size was relatively adequate for the study (Figure 1).

The groups consisting of four (45, 33, 04, and 28), five (01, 50, 49, 35, and 48), and two (29 and 32) S$_3$ progenies, respectively, are genotypes with a high potential for generating strains, since they did not form groups with any of the four testers analyzed. This characterizes

them as good candidates for the formation of complementary groups when combined with strains originating from the testers utilized in the present study.
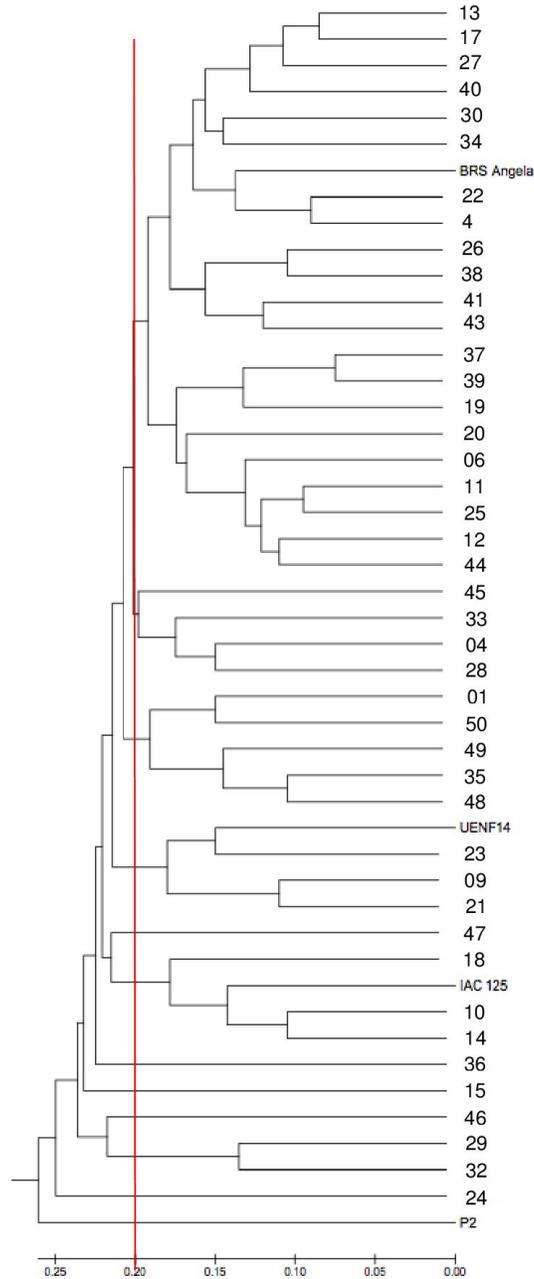
Figure 1. Dendrogram obtained using the UPGMA method including the testers and S_3 progenies. The cutoff point (indicated by a red line) was determined based on the overall mean of the distances (0.20).

Based on the formed groups, tester UENF 14 grouped with progenies 23, 09, and 21, whereas tester IAC 125 grouped with progenies 18, 10, and 14 (Figure 1). From this formation, it would be interesting to obtain hybrids of strains derived from genotypes present in the same group, because heterosis is little exploited in crosses between genetically similar materials. Among the remaining genotypes that did not group at all (progenies 47, 36, 15, 46, 24, and tester P2), tester P2 and progeny 24 stood out for showing the greatest genetic dissimilarity in relation to the other genotypes. Thus, obtaining strains from these materials would be the best strategy when aiming to exploit the heterosis effect to its maximum.

Resh et al. (2015) analyzed dominant molecular markers to estimate the genetic diversity of 26 genotypes of popcorn and calculated the genetic distance between the genotypes, using the Jaccard, Dice, and Rogers and Tanimoto coefficients. The dissimilarity matrix obtained from the genetic distances was used to construct dendrograms based on the three different coefficients. Four groups were formed according to the Jaccard and Dice coefficients, whereas only three groups were formed based of the Rogers and Tanimoto coefficient. In all three dendrograms, the same big group, containing 23 genotypes, was obtained. These results suggest that it is not trivial to clearly define distinct heterotic groups for the popcorn genotypes evaluated.

The limitations of the present study were limited by including a large number of measured samples and markers. Both the type of genetic marker (EST-SSR) and the analysis technique (Fragment Analyzer) used are considered robust tools for genetic diversity studies.

CONCLUSIONS

The analyses performed led us to conclude that there is high genetic diversity among the four testers and the 43 S$_3$ progenies included in this study. It can also be inferred that progeny 24 and tester P2 display lower genetic similarity in relation to the others and that obtaining strains from these genotypes for the development of hybrids with a high heterotic effect can be a promising alternative. However, the identification of heterotic groups in popcorn will depend on additional research utilizing a larger number of genotypes from the crop germplasm bank. This will provide a more comprehensive view of the genetic relationships and the identification of divergent groups with potential for exploitation of heterosis.

The results found in the present study can be directly used in the development of popcorn breeding programs in the production of hybrids as well as in genetic mapping studies. Furthermore, the functional genetic diversity indicates that there are informative genetic markers for popcorn. A more in-depth study is an interesting avenue of further applicability for this crop.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

Alexander DE and Creech RG (1977). Breeding special industrial and nutritional types. In: Corn and corn improvement (Sprague GF and Fuccilo DA, eds.). American Society of Agronomy, Madison, 363-386.

Amaral Júnior AT, Gonçalves LSA, Freitas Júnior SP, Candido LS, et al. (2013). UENF 14: a new popcorn cultivar. *Crop Breed. Appl. Biotechnol.* 13: 218-220.

Amaral LP, Ferreira RA, Lisboa GS, Longhi SJ, et al. (2013). Variabilidade espacial do índice de diversidade de Shannon-Wiener em floresta ombrófila mista. *Sci. For.* 41: 083-093.

Andrade RA (1996). Cruzamentos dialélicos entre seis variedades de milho pipoca. Master's thesis, Universidade Federal de Viçosa, Viçosa, MG.

Barreto RR, Scapim CA, Amaral Júnior AT, Rodovalho MA, et al. (2012). Avaliação da capacidade de combinação de famílias S_2 de milho-pipoca por meio de diferentes testadores. *Semina Cienc. Agrar.* 33: 873-890. http://dx.doi.org/10.5433/1679-0359.2012v33n3p873

Coimbra RR (2000). Seleção entre famílias de meios-irmãos da população DFT-1 Ribeirão de milho pipoca. Master's thesis, Universidade Federal de Viçosa, Viçosa, MG.

Cruz CD (2013). GENES - a software package for analysis in experimental statistics and quantitative genetics. *Acta Sci. Agron.* 35: 271-276. http://dx.doi.org/10.4025/actasciagron.v35i3.21251

Daher RF, Pereira MG, Tupinambá EA, Amaral Júnior AT, at el. (2002). Assessment of coconut tree genetic divergence by compound sample RAPD marker analysis. *Crop Breed. Appl. Biotechnol.* 2: 431-438. http://dx.doi.org/10.12702/1984-7033.v02n03a14

Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Erwin AT (1949). The origin and history of popcorn, *Zea mays* L. var. indurate (Sturt.) Bailey mut. Everta (Sturt.) Erwin. *Agron. J.* 41: 53-56. http://dx.doi.org/10.2134/agronj1949.00021962004100020001x

Eujayl I, Sorrells ME, Baum M, Wolters P, et al. (2002). Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theor. Appl. Genet.* 104: 399-407. http://dx.doi.org/10.1007/s001220100738

Falconer DS and Mackay TF (1996). Introduction to quantitative genetics. Longman Harrow, Essex.

Galvão KS, Ramos HC, Santos PH, Entringer GC, et al. (2015). Functional molecular markers (EST-SSR) in the full-sib reciprocal recurrent selection program of maize (*Zea mays* L.). *Genet. Mol. Res.* 14: 7344-7355. http://dx.doi.org/10.4238/2015.July.3.10

Hallauer AR, Carena MJ and Miranda Filho JB (2010). Quantitative genetics in maize breeding. Springer, New York.

Kumar P, Gupta VK, Misra AK, Modi DR, et al. (2009). Potential of molecular markers in plant biotechnology. *Plant Omics J.* 2: 141-162.

Laborda PR, Oliveira KM, Garcia AA, Paterniani ME, et al. (2005). Tropical maize germplasm: what can we say about its genetic diversity in the light of molecular markers? *Theor. Appl. Genet.* 111: 1288-1299. http://dx.doi.org/10.1007/s00122-005-0055-7

Larish LB and Brewbaker JL (1999). Diallel analyses of temperate and tropical popcorn. *Maydica* 44: 279-284.

Liu K and Muse SV (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128-2129. http://dx.doi.org/10.1093/bioinformatics/bti282

Miranda GV, Souza LV, Galvão JCC, Guimarães LJM, et al. (2008). Genetic variability and heterotic groups of Brazilian popcorn populations. *Euphytica* 162: 431-440. http://dx.doi.org/10.1007/s10681-007-9598-9

Mohammadi SA and Prasanna BM (2003). Analysis of genetic diversity in crop plants-salient statistical tools and considerations. *Crop Sci.* 43: 1235-1248. http://dx.doi.org/10.2135/cropsci2003.1235

Moterle LM, Braccini AL, Scapim CA, Pinto RJB, et al. (2012). Combining ability of popcorn lines for seed quality and agronomic traits. *Euphytica* 185: 337-347. http://dx.doi.org/10.1007/s10681-011-0458-2

Pacheco CAP, Gama EEG, Parentoni SN, Santos MX, et al. (2001). BRS Angela: variedade de milho-pipoca. EMBRAPA - Milho e Sorgo, Sete Lagoas.

Ramos HC, Pereira MG, Silva FF, Gonçalves LS, et al. (2011). Genetic characterization of papaya plants (*Carica papaya* L.) derived from the first backcross generation. *Genet. Mol. Res.* 10: 393-403. http://dx.doi.org/10.4238/vol10-1gmr1004

Resh FS, Scapim CA, Mangolin CA, Machado MF, et al. (2015). Genetic diversity of popcorn genotypes using molecular analysis. *Genet. Mol. Res.* 14: 9829-9840. http://dx.doi.org/10.4238/2015.August.19.16

Ribeiro RM, do Amaral Júnior AT, Gonçalves LS, Candido LS, et al. (2012). Genetic progress in the UNB-2U population of popcorn under recurrent selection in Rio de Janeiro, Brazil. *Genet. Mol. Res.* 11: 1417-1423. http://dx.doi.org/10.4238/2012.May.15.12

Rodovalho MA, Scapim CA, Pinto RJB, Barreto RR, et al. (2012). Comparação de testadores em famílias S$_2$ obtidas do híbrido simples de milho-pipoca IAC-112. *Biosci. J.* 28: 145-154.

Scott KD, Eggler P, Seaton G, Rossetto M, et al. (2000). Analysis of SSRs derived from grape ESTs. *Theor. Appl. Genet.* 100: 723-726. http://dx.doi.org/10.1007/s001220051344

Sharopova N, McMullen MD, Schultz L, Schroeder S, et al. (2002). Development and mapping of SSR markers for maize. *Plant Mol. Biol.* 48: 463-481. http://dx.doi.org/10.1023/A:1014868625533

Silva TA, Pinto RJB, Scapim CA, Mangolin CA, et al. (2009). Genetic divergence in popcorn genotypes using microsatellites in bulk genomic DNA. *Crop Breed. Appl. Biotechnol.* 9: 31-36. http://dx.doi.org/10.12702/1984-7033.v09n01a05

Silva VQR, Amaral Júnior AT, Scapim CA, Freitas Júnior SP, et al. (2010). Inheritance for economically important traits in popcorn from distinct heterotic groups by Hayman's diallel. *Cereal Res. Commun.* 38: 272-284. http://dx.doi.org/10.1556/CRC.38.2010.2.14

Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739. http://dx.doi.org/10.1093/molbev/msr121

Warburton ML, Xianchun X, Crossa J, Franco J, et al. (2002). Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. *Crop Sci.* 42: 1832-1840. http://dx.doi.org/10.2135/cropsci2002.1832

Yeh FC, Yang RC, Boyle TBJ, Ye ZH, et al. (1997). POPGENE, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Center, University of Alberta, Edmonton.

Yu K, Park SJ and Poysa V (1999). Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*). *Genome* 42: 27-34. http://dx.doi.org/10.1139/gen-42-1-27

Zhang M, Mao W, Zhang G and Wu F (2014). Development and characterization of polymorphic EST-SSR and genomic SSR markers for Tibetan annual wild barley. *PLoS One* 9: e94881. http://dx.doi.org/10.1371/journal.pone.0094881