



Genetic diversity of Cajazeira (*Spondias mombin* L.) in three geographic regions

B.M. Silva¹, A.A.B. Rossi¹, A.V. Tiago¹, K.F.M. Schmitt², J.F.E. Dardengo¹ and S.A.M. Souza¹

¹Programa de Pós-Graduação em Biodiversidade e Agroecossistemas Amazônicos, Programa de Pós-Graduação em Genética e Melhoramento de Plantas, PPGBionorte, Laboratório de Genética Vegetal e Biologia Molecular, Faculdade de Ciências Biológicas e Agrárias, Universidade do Estado de Mato Grosso, Campus de Alta Floresta, Alta Floresta, MT, Brasil

²Centro de Ciências e Tecnologias Agropecuárias, Universidade Estadual Norte Fluminense Darcy Ribeiro, UENF, Campos dos Goytacazes, RJ, Brasil

Corresponding author: A.A.B. Rossi
E-mail: anabanrossi@unemat.br

Genet. Mol. Res. 16 (1): gmr16018946

Received July 5, 2016

Accepted November 16, 2016

Published January 23, 2017

DOI <http://dx.doi.org/10.4238/gmr16018946>

Copyright © 2017 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

ABSTRACT. Cajazeira (*Spondias mombin* L.), of the family Anacardiaceae, is a species of fruit tree found in the Amazon region with fruits that have excellent prospects for commercial use. We aimed to evaluate the genetic diversity within and among natural populations of *S. mombin*, with natural occurrence in northern Mato Grosso State, by using inter-simple sequence repeat (ISSR) markers. Overall, 126 individuals were evaluated from three populations located Alta Floresta (AFL) 42, Marcelândia (MAR) 41, and Nova Bandeirantes (NBA) 43. The individuals were genotyped with 14 ISSR primers, which amplified 99 fragments. All markers, with the exception of DiGA3'A, presented a polymorphic information content above 0.25, and thus, are

recommended for diversity analyses in *S. mombin*. Genetic diversity of the AFL [Nei's diversity (H) = 0.2430 and Shannon index (I) = 0.3547] and MAR (H = 0.2062 and I = 0.2993) populations was higher when compared to the NBA population, which presented the lowest genetic diversity (H = 0.2002 and I = 0.2957). Analysis of molecular variance showed that 77.38% of the total genetic variation is found within populations while 22.62% is found among populations. AFL and NBA are genetically the most similar populations and also the closest "Structure" revealed genetic diversity among the genotypes of each population. As there is genetic variability in both populations, and there are no genetically identical individuals, both populations can be a source of genotypes for germplasm banks and for future commercial fruitful plantations *S. mombin*.

Key words: *Spondias mombin*; ISSR markers; Genetic variability

INTRODUCTION

Brazil offers a wide diversity of soil and climates, which favor the cultivation of various tropical fruit species, raising its fruit culture potential (Simão, 1998). Cajazeira (*Spondias mombin* L.) is a fruitful species belonging to Anacardiaceae family, including species such as cirigueleira, cajaraneira, and umbuzeiro (Souza, 2005). These species are scattered in the tropical regions of America, Africa, and Asia, and in Brazil, they are found mainly in the north and northeast regions.

The fruits have been given different names, including cajá, cajá verdadeiro, cajá-mirim, and taperebá (Sacramento and Souza, 2000). They exhibit high potential for use in agro-industry due to the high nutritional quality of their pulp, which is rich in carotenoids, sugars, and vitamins A and C (Barroso et al., 1999). The cultivation of cajazeiras represents a way for Brazil to expand its exports, due to the use of the fruit *in natura*, in the form of juice, ice cream, jam, and frozen pulp (Silva, 2003).

Due to the increased consumption of cajazeira fruits and their derivatives, these species have attracted interest for cultivation; however, there is still little information available regarding the planting of commercial orchards (Bosco et al., 2000).

According to Costa et al. (2011), information on the genetic variation of native species, both within and between populations, is fundamental for the development of strategies that permit the domestication and incorporation of these species into regional productive systems, as well as to support plans for the conservation of genetic resources.

Molecular techniques accelerate the analysis of variability and selection mainly when working with perennial species. Among the various molecular techniques, inter simple sequence repeats (ISSRs) have been used successfully to reveal molecular polymorphisms in several plant species, and are considered to be widely variable due to their widespread occurrence and distribution in the genome (Ellegren, 2004). In addition, this marker has the advantage of not requiring prior information of the genome and is highly polymorphic (Ziekiewicz et al., 1994). It has proven to be useful in genetic studies, especially when investigating the genetic diversity and relationships between closely related individuals (Salimath et al., 1995).

Several studies conducted with ISSR markers have demonstrated their efficiency in the analysis of genetic variability in populations of plants such as *Theobroma speciosum* (Giustina

et al., 2014), *Theobroma subincanum* (Rivas et al., 2013), *Mauritia flexuosa* (Rossi et al., 2014), and umbu-cajazeira (Santana et al., 2011). The above studies may provide important information for the conservation and creation of breeding programs aimed at the commercial cultivation of the species.

There is growing commercial interest in *S. mombin*, and conservation strategies aim to gather genetic information on this species to enable the genetic resources to be fully characterized in order to prevent the loss of genetic variability and to allow promising genetic gains for improvement. In this context, the present study aimed to evaluate the genetic diversity within and among populations of *S. mombin* with natural occurrence in the north of Mato Grosso State using ISSR markers.

MATERIAL AND METHODS

Vegetable material

To characterize the genetic diversity of *S. mombin*, three cities located in the northern region of the state of Mato Grosso were selected: Alta Floresta (AFL; 09°52'32"S and 56°05'10"W), Marcelândia (MAR; 11°05'22"S and 54°27'02"W), and Nova Bandeirantes (NBA; 09°48'50"S and 57°51'43"W). In each municipality, an area of the species natural occurrence was selected; each area was made up of one study population, for a total of three populations. A total of 126 individuals were sampled, which were distributed between the municipalities of AFL (N = 42), MAR (N = 41), and NBA (N = 43).

Young leaves of 126 individuals were collected from the three populations, the leaf material was identified and packaged *in silica* gel while still in the field, and was subsequently stored in a freezer at -20°C.

Extraction of total DNA, selecting of primers, optimization of PCR, and ISSR amplification

Total genomic DNA was extracted from approximately 100 mg leaves by following the cetyltrimethylammonium bromide (CTAB) as described by Doyle and Doyle (1987). The leaves were macerated with liquid nitrogen. They were later transferred to 2-mL micro tubes, identified, and 800 µL 5% CTAB Buffer (100 mM Tris-HCl, pH 8.0; 1.4 M sodium chloride; 20 mM EDTA), containing 2% PVP (polyvinylpyrrolidone) and 3% β-mercaptoethanol was added. Then, the samples were homogenized in a vortex mixer for 1 min and incubated in a water bath at 65°C for 30 min. After extraction, the DNA was resuspended in TE buffer and kept at -20°C.

The extracted DNA was quantified by comparing the size with DNA of known molecular mass by electrophoresis on 1% agarose gel stained with ethidium bromide (Ferreira and Grattapaglia, 1998). The gels were analyzed using a transilluminator under ultraviolet light and then photographed. After quantification, the DNA samples were diluted in autoclaved distilled water and standardized to a concentration of 20 ng/µL.

For ISSR amplification, a total of 50 ISSR primers (UBC primer set #9, Biotechnology Laboratory, University of British Columbia, Canada) were evaluated using four randomly chosen individuals. Among the 50 primers, 14 that yielded strong, reproducible bands were selected for use in subsequent experiments (Table 1).

Table 1. Number of fragments amplified by primers at the population level: Alta Floresta (AFL), Marcelândia (MAR), and Nova Bandeirantes (NBA).

Primers	Populations											
	AFL				MAR				NBA			
	NT	NP	P%	PIC	NT	NP	P%	PIC	NT	NP	P%	PIC
Di(AG) ₈ 3'C	7	5	71.43	0.15	7	6	85.71	0.21	7	5	71.43	0.34
Di(GA) ₈ 3'T	6	2	33.33	0.25	6	3	50.00	0.22	6	3	50.00	0.35
Di(GA) ₈ 3'A	5	2	40.00	0.16	5	1	20.00	0.10	5	2	40.00	0.11
Di(AC) ₈ 3'T	10	7	70.00	0.24	10	6	60.00	0.23	10	8	80.00	0.42
Di(AC) ₈ 3'C	8	6	75.00	0.45	7	5	71.43	0.43	9	7	77.78	0.48
Di(AC) ₈ 3'G	7	6	85.71	0.27	7	3	42.86	0.10	7	3	42.86	0.12
Di(AG) ₈ 3'YC*	6	4	66.67	0.31	6	2	33.33	0.10	6	2	33.33	0.05
Di(AC) ₈ 3'YT*	7	4	57.14	0.29	7	3	42.86	0.24	7	3	42.86	0.31
Tri(CTC) ₆	8	4	50.00	0.38	8	4	50.00	0.12	8	2	25.00	0.04
Tri(GAA) ₆	7	5	71.43	0.43	7	5	71.43	0.38	7	5	71.43	0.38
Di(CA) ₇ 5'BDB	8	4	50.00	0.41	8	4	50.00	0.34	7	3	42.86	0.43
Di(TG) ₇ 5'HVH	4	2	50.00	0.14	4	2	50.00	0.49	4	2	50.00	0.29
Tri(ACA) ₈ 3'RC	9	5	55.56	0.34	9	3	33.33	0.13	9	4	44.44	0.23
Di(GA) ₈ 3'C	6	4	66.67	0.21	6	3	50.00	0.15	6	4	66.66	0.44
Mean	7	4.28	61.21	0.29	6.92	3.57	50.78	0.23	7	3.78	52.76	0.28

*Y = C or T; B = C, G or T; R = A or G; V = A, C or G; H = A, C or T. NT = total number of amplified bands; NP = total number of polymorphic bands; P% = percentage polymorphism; PIC = polymorphic information content.

PCR amplifications were carried out in a final volume of 20 μ L, containing 1 μ L DNA (20 ng), 2 μ L buffer (10X 1 M KCl; 1 M Tris pH 8.3; 1 M MgCl₂; 10% Tween 20), 2 μ L MgC₂ (25 mM), 3 μ L primer (0.2 mM), 4 μ L dNTP (0.1 mM each dNTP), 1 μ L DMSO, and 0.2 μ L Taq polymerase (5 U/ μ L). Amplifications were performed in a thermocycler biocycler (MJ 96) under the following conditions: 15 min at 95°C (initial denaturation); 35 cycles of 30 s at 94°C (denaturation); 45 s at 48°-50°C (depending on the primer used) (annealing); 2 min at 72°C (extension), and 7 min at 72°C (final extension).

Amplified products were separated by electrophoresis on 1.5% agarose gel in running buffer 1X TBE (89.15 mM Tris basis; 88.95 mM boric acid; 2.23 mM EDTA), at a constant voltage of 90 V for approximately 4 h. Gels were stained with ethidium bromide (0.6 ng/ μ L) for 15 min. The sizes of the amplified fragments were estimated through comparison with a 100-bp DNA Ladder. The gel was then photographed upon irradiation with ultraviolet light through a UVB Transilluminator LTB-21x26 (Loccus Biotechnology®) and a digital camera (Sony®).

Analysis data

The ISSR fragments were treated as dominant markers and judged as binary characters: (1) presence or (0) absence of bands. Only robust and unambiguous bands were evaluated. Bands of weak intensity or those coalescent with other bands were excluded. The binary matrix showing the presence/absence of bands was analyzed using the program POPGENE (Population Genetic Analysis) version 1.32 (Yeh et al., 1999) to estimate the genetic diversity based on Nei's genetic diversity (1978) (H), the percentage of polymorphic loci (P%), and the Shannon index (I) for each population and for the set of populations. Genetic diversity was analyzed among populations using Nei's method (1978), estimating the total heterozygosity (Ht), the average gene diversity within populations (HS), the average genetic divergence between the populations (G_{ST}), and the gene flow (Nm).

Analysis of molecular variance (AMOVA) was used to reveal the distribution of the

genetic diversity within and between populations as described by Excoffier et al. (1992), with the aid of the program ARLEQUIN 3.01 (Excoffier et al., 2006).

Genetic diversity at the locus or the polymorphism information content (PIC) was used to evaluate the discriminatory power of one locus. The informativeness of locus p_i is the frequency of allele p in locus p_i , calculated using the equation, $PIC = 1 - \sum_i p_i^2$. The informativeness of primer p_{ij} is the frequency of allele p of locus i in primer j (Rezende et al., 2009) and is calculated using the equation, $PIC = 1 - \sum_i p_i^2$. The informativeness of primer p_{ij} is the frequency of allele p of locus i in the primer j (Rezende et al., 2009) and is calculated using the equation, $PIC_{primer} = 1 - \sum_j \sum_i p_{ij}^2$.

To form the matrix of dissimilarity, the arithmetic complement of the Jaccard index was used. The genetic dissimilarity matrix was used to analyze the grouping of genotypes by the unweighted pair-group method average (UPGMA) method, which was computed with the aid of the program GENES (Cruz, 2006).

The program "Structure" (Pritchard et al., 2000), based on Bayesian statistics, was used to infer the number of genetic groups (k). Twenty run-ins were performed for each value of k , 200,000 initial interactions ("burn-ins"), and 500,000 Monte Carlo simulations through the Markov chain. To determine the most probable value of k , we used the criteria proposed by Pritchard and Wen (2004) and Evanno et al. (2005).

RESULTS AND DISCUSSION

The 14 ISSR primers amplified a total of 99 fragments in the 126 Cajá genotypes, including 64 (64.65%) polymorphic and 35 (35.35%) monomorphic bands. The average number of bands amplified per primer was 6.97. Figure 1 shows the standard of amplification obtained using primer 24 in individuals from the MAR population.

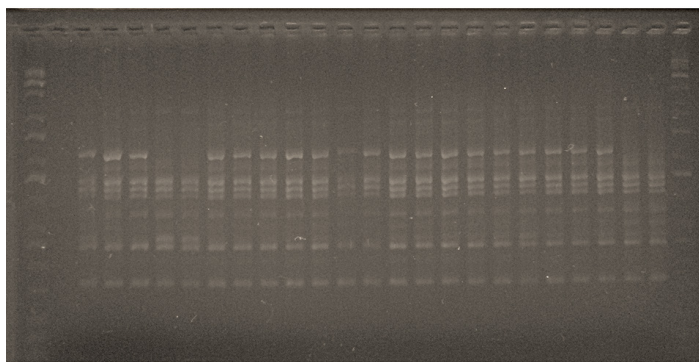


Figure 1. Standard amplification using the Di(AG)3'T primer in 24 individuals of the Marcelândia population.

According to Alves et al. (2007), the majority of tropical tree species present a large number of alleles per locus. This includes the average found in the present study (6.97), showing that *S. mombin* has variability and that its genotypes possess genetic material that may be used in breeding programs and for subsequent domestication. Santana et al. (2011) studied 17 accessions of umbu-cajazeiras belonging to BAG Tropical Fruit of EMBRAPA Mandioca e Fruticultura using ISSR markers, and found an average of 10 bands per primer.

This is higher than that found in the present study, in which an average 6.97 bands were found per primer, which is justified by the fact that this study was performed to evaluate natural populations without prior selection of genotypes.

The primer that presented the highest PIC value in each population was AFL (DiAC3'C: 0.45), MAR (DiTG5'HVH: 0.49), and NBA (DiAC3'C: 0.48). Those presenting the smallest values were AFL (DiTG5'HVH: 0.14), MAR (DiGA3'A/DiAC3'G/DiAG3'YC* both with 0.10), and NBA (TriCTC: 0.04) (Table 1).

According to the classification of Botstein et al. (1980), markers with PIC values above 0.5 are considered very informative, values between 0.25 and 0.50 are considered moderately informative, and values lower than 0.25 are considered little informative. According to Boza et al. (2013), highly polymorphic markers are useful in the identification of genetic diversity. In general, regardless of the population of the 14 markers used, except DiGA3'A, 13 presented PIC values above 0.25. These findings indicate that the primers used were moderately informative and are therefore, recommended for the analysis of diversity in cajazeira.

The highest index of diversity, estimated from H and I, was found in populations AFL (H = 0.2430 and I = 0.3547) and MAR (H = 0.2062 and I = 0.2993), while NBA had the lowest genetic diversity (H = 0.2002 and I = 0.2957); in level of species, without considering the populations separated, the diversity showed the following results: H = 0.2689 and I = 0.3971 (Table 2). The Shannon diversity index and Nei's genetic diversity vary from 0 to 1, with 0 considered as zero gene diversity and 1 considered the maximum gene diversity. The data obtained here show there is genetic diversity in natural populations of *S. mombin*.

Table 2. Genetic diversity in natural populations of *Spondias mombin* located in the municipalities of Alta Floresta (AFL), Marcelândia (MAR), and Nova Bandeirantes (NBA), state of Mato Grosso, Brazil.

Populations	N	P (%)	H	I
AFL	42	61.21	0.2430	0.3547
MAR	41	50.78	0.2062	0.2993
NBA	43	52.76	0.2002	0.2957
Species	126	71.72	0.2689	0.3971

Estimated by Nei's genetic diversity index (H) and the Shannon index (I). Percentage polymorphism (P%); sample size (N).

The diversity index in other species obtained under the conditions used in this study may aid in understanding H and I. Rossi et al. (2014) reported high diversity in *M. flexuosa* (H = 0.206 and I = 0.308), and Estopa et al. (2006) reported high diversity in natural populations of *Eremanthus erythropappus* (H = 0.490 and I = 0.330). These results are relatively similar to those obtained in the present study. However, the values of H and I described in those previous studies were higher than those obtained for *S. mombin*.

Among the analyzed populations, AFL and NBA showed a higher percentage of polymorphism, which was 61.21 and 52.76% respectively, when compared to that of the MAR population, which presented 50.78% polymorphism.

The results of AMOVA indicated that 22.62% of the total variance occurred among populations and 77.38% occurred within populations, demonstrating that the greatest genetic differentiation lies within the intrapopulation rather than the interpopulation component (Table 3). There was significant genetic differentiation ($P < 0.01$) among the populations.

Table 3. Analysis of molecular variance (AMOVA) of the three populations of *Spondias mombin* studied based on 14 ISSR markers.

Source of variation	d.f.*	SS*	CV*	TV (%)*	P value*
Between populations	2	174.19	1.91	22.62	<0.01
Between populations	123	808.99	6.56	77.38	
Total	125	981.19	8.47		

*d.f. = degrees of freedom, SS = sum of squares, CV = coefficient of variation, TV = total variation, and P = chances of a variance component greater than the observed values by chance. The probabilities were calculated using 1023 random permutations. $F_{ST} = 0.2262$.

These results are consistent with several studies on tropical species showing that greater genetic diversity occurs at the intrapopulation level (Rossi et al., 2009; Bertoni et al., 2010; Brandão et al., 2011). Studies on genetic variability in natural populations of plants in tropical regions have demonstrated that high levels of variability are preserved within populations (Paiva, 1998).

The results found for *S. mombin* in this study are consistent with the conclusion of Avise and Hamrick (1995), who stated that autogamous species have low genetic diversity within populations and high genetic differentiation among populations, when compared with allogamous species such as cajazeira.

For species with crossing reproduction (allogamous), estimates of genetic differentiation between natural populations based on data from AMOVA with random amplified polymorphic markers (a dominant marker similar to ISSR) are usually below 28%. For autogamous species, the estimated interpopulation genetic variation is normally above 70% (Nybon and Bartish, 2000). Genetic diversity among the populations in the present study was 22.62% (Table 3), which is within the estimates of Nybon and Bartish (2000) for species with crossing reproduction. According to Hartl (2008), allogamous species permit high rates of diversity to be found within populations, since allogamy favors recombination.

Another genetic parameter that enables genetic drift between populations to be estimated is the F_{ST} (interpopulational genetic divergence), which was estimated at 0.2262 in the present study (Table 3), indicating a moderate level of genetic divergence among the three populations studied. According to Wright (1965), F_{ST} values above 0.25 indicate high levels of genetic differentiation. Therefore, these results reinforce the existence of genetic differentiation in the *S. mombin* populations analyzed, making the collection of genotypes from each population necessary to compose germplasm banks and to maintain the diversity of the species. These banks will subsequently subsidize the selection of genotypes for commercial plantations.

The values for F_{ST} and estimated Nm, suggest that Nm has been sufficient to counteract the effects of drift, revealing a significant differentiation among populations, when measured by the fixation index, F_{ST} .

Ht estimated was 0.2688, indicating that the species in these populations represent a reserve of genetic variability. The percentage of G_{ST} was 0.1947.

In the populations of *S. mombin* studied, the estimated Nm or number of migrants per generation was 2.0684; therefore, it can be concluded that there was Nm between generations and that the populations are not genetically isolated.

Values of Nm lower than 1.0 indicate genetic isolation, and values greater than 1.0 are sufficient to prevent the random loss of alleles within populations. A value of Nm greater than four migrants per generation indicates homogenization of alleles within the population (Hartl and Clark, 1997).

According to Hamrick (2012), N_m among individuals within and between populations can have a big impact on the distribution of genetic variation. This is because the N_m , when introducing new variations, reduces the genetic differentiation among populations. In fact, one or more individual migrant between the populations is sufficient to prevent substantial differentiation between them.

The observed N_m may be associated with historic N_m when the populations were part of a single large population. Thus, data should be interpreted with caution, because this indicates that populations from different regions, although geographically isolated, feature alleles and similar frequencies as a result of historical N_m (Raposo et al., 2007).

Table 4 shows the estimated genetic identity and Nei (1978)'s genetic distance for all pair-to-pair comparisons among populations. The largest genetic identity was observed between the populations from AFL and NBA (0.9047), while the lowest was between the populations from MAR and NBA (0.9016). These results demonstrate that the populations from AFL and NBA are the most similar genetically (0.9047), whereas those from MAR and NBA are genetically more distant (0.1036), showing that the populations with greater genetic similarity are the nearest geographically. This suggests that the genetic structure of *S. mombin* can be determined by the geographical distance between populations.

Table 4. Geographical distance, genetic distance, and Nei (1978)'s genetic identity among three populations of *Spondias mombin*: Alta Floresta (AFL), Marcelândia (MAR) and Nova Bandeirantes (NBA), Mato Grosso State, Brazil.

Population	Genetic distance	Nei's identity	Geographical distance* (km)
AFL with MAR	0.1018	0.9032	220
AFL with NBA	0.1002	0.9047	190
MAR with NBA	0.1036	0.9016	386

*Distance provided by Google Earth, in a straight line.

The dendrogram generated based on UPGMA of the 126 analyzed genotypes formed nine groups. Group I was allocated 21 genotypes of the MAR population; Group II contained six genotypes from AFL; Group III was formed by three genotypes of the MAR population; Group IV contained six genotypes from AFL; Group V contained 15 MAR genotypes; Group VI was composed of 26 genotypes from AFL; Group VII was allocated 44 genotypes, including 43 individuals of the NBA population and one genotype of the AFL population; Group VIII merged two MAR genotypes; and Group IX contains three genotypes from AFL.

All of the individuals formed well-defined groupings with genotypes in the same population, except for genotype 40 of the AFL population, which was allocated to Group VII from the NBA population.

The cophenetic correlation showed an association of 76% ($CCC = 0.7600$) between the distances obtained in the dissimilarity matrix (complement of Jaccard) and the cophenetic matrix. According to Patto et al. (2004), cophenetic correlations higher than 0.56 reflect good concordance between the arrays.

The number of genetic groups (k) was also determined using the Structure program (Pritchard et al., 2000), which is based on Bayesian statistics. The best k value found relating to the three populations was represented by three groups ($k = 3$), which verified that the individuals are clearly allocated according to their place of origin (Figure 2).

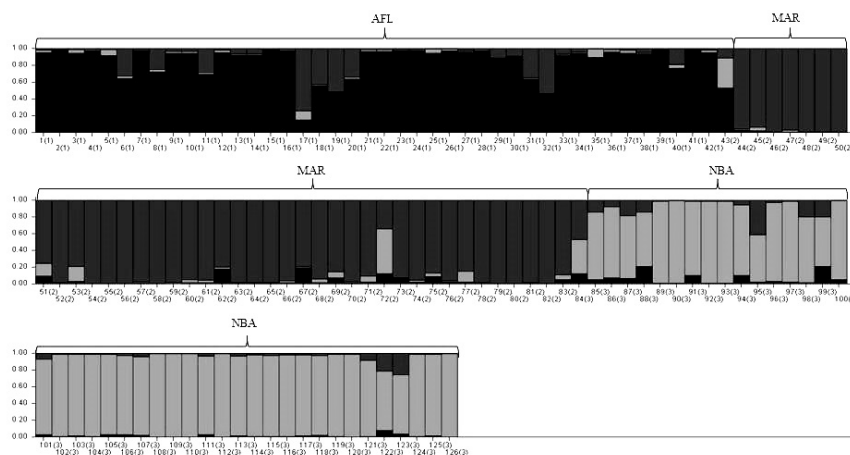


Figure 2. Distribution of 126 *Spondias mombin* individuals into three groups based on molecular data from 14 ISSR (inter-simple sequence repeats) primers using the Structure program ($k = 3$), in accordance with the original populations: Alta Floresta (AFL), Marcelândia (MAR) and Nova Bandeirantes (NBA). Individuals are represented by vertical columns and shaded according to their group.

In conclusion, molecular characterization revealed that the genetic diversity is greater at the intrapopulation than at the interpopulation level in the populations studied, and that the populations with greater genetic similarity are the closest geographically. It is proposed that individuals from populations AFL, MAR and NBA are retained in order to preserve the intrapopulation genetic diversity, and to maintain genetic variability and the effective conservation of these populations. As there is genetic variability in both populations, and no individuals are genetically very close or identical, both may represent a source of genotypes for germplasm banks and for future commercial fruitful plantations of *S. mombin*.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors wish to thank FAPEMAT (Projeto: Conservação e uso de espécies vegetais nativas da região Amazônica com potencial econômico para região Norte do estado de Mato Grosso; process #166159/2014). Bruna Mezzalira da Silva received a Master's scholarship from FAPEMAT/CAPES.

REFERENCES

- Alves RM, Sebbenn AM, Artero AS, Clement C, et al. (2007). High levels of genetic divergence and inbreeding in populations of cupassu (*Theobroma grandiflorum*). *Tree Genet. Genomes* 3: 289-298. <http://dx.doi.org/10.1007/s11295-006-0066-9>
- Avise JC and Hamrick JL (1995). Conservation genetics: case histories from nature. Norwel: Kluwer Academic Publishers.
- Barroso GM, Morim MP, Peixoto AL and Ichaso CLF (1999). Frutos e sementes: morfologia aplicada à sistemática de dicotiledôneas. Universidade Federal de Viçosa, Viçosa.

- Bertoni BW, de C Telles MP, Malosso MG, Torres SC, et al. (2010). Genetic diversity in natural populations of *Jacaranda decurrens* Cham. determined using RAPD and AFLP markers. *Genet. Mol. Biol.* 33: 532-538. <http://dx.doi.org/10.1590/S1415-47572010005000068>
- Bosco J, Soares KT, Aguiar Filho SD and Barros RV (2000). A cultura da cajazeira. João Pessoa: Editora Emepa.
- Botstein D, White RL, Skolnick M and Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.
- Boza EJ, Irish BM, Meerow AW, Tondo CL, et al. (2013). Genetic diversity, conservation, and utilization of *Theobroma cacao* L.: genetic resources in the Dominican Republic. *Genet. Resour. Crop Evol.* 60: 605-619. <http://dx.doi.org/10.1007/s10722-012-9860-4>
- Brandão MM, Vieira FA and Carvalho D (2011). Estrutura genética em microescala espacial de *Myrcia splendens* (Myrtaceae). *Rev. Arvore* 35: 957-964. <http://dx.doi.org/10.1590/S0100-67622011000600001>
- Costa TS, Silva AD, Léo ADS, Santos AD, et al. (2011). Diversidade genética de acessos do banco de germoplasma de mangaba em Sergipe. *Pesqui. Agropecu. Bras.* 46: 499-508. <http://dx.doi.org/10.1590/S0100-204X2011000500007>
- Cruz CD (2006). Programa Genes: Biometria. 1st edn. UFV, Viçosa.
- Doyle JJ and Doyle JL (1987). A rapid DNA isolation procedure for small amounts of fresh leaf tissue. *Phytochem. Bull.* 19: 11-15.
- Ellegren H (2004). Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* 5: 435-445. <http://dx.doi.org/10.1038/nrg1348>
- Estopa RA, Souza AMD, Moura MCDO, Botrel MCG, et al. (2006). Diversidade genética em populações naturais de candeia (*Eremanthus erythropappus* (DC.) MacLeish). *Sci. Forum* 70: 97-106.
- Evanno G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620. <http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Smouse PE and Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.
- Excoffier L, Laval G and Schneider S (2006). Arlequinver 3.01. An integrated software package for population genetics data analysis. Computational and Molecular Population Genetics Lab (CMPG). Institute of Zoology. University of Berne.
- Ferreira ME and Grattapaglia D (1998). Introdução ao Uso de Marcadores Moleculares em Análise Genética. 3rd edn. Embrapa, Brasília.
- Giustina LD, Luz LN, Vieira FS, Rossi FS, et al. (2014). Population structure and genetic diversity in natural populations of *Theobroma speciosum* Willd. Ex Spreng (Malvaceae). *Genet. Mol. Res.* 13: 3510-3519. <http://dx.doi.org/10.4238/2014.February.14.5>
- Hamrick JL (2012). Tropical breeding systems: one and done? *Heredity (Edinb)* 109: 330-331. <http://dx.doi.org/10.1038/hdy.2012.47>
- Hartl DL (2008). Princípios da genética de população. FUNPEC, Ribeirão Preto.
- Hartl DL and Clark AG (1997). Principles of population genetics. Sunderland: Sinauer Associates.
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.
- Nybon H and Bartish I (2000). Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Perspect. Plant Ecol. Evol. Syst.* 3: 93-114. <http://dx.doi.org/10.1078/1433-8319-00006>
- Paiva JR (1998). Melhoramento genético de espécies agroindustriais na Amazônia: estratégias e novas abordagens. Brasília, DF. Embrapa-SPI, Embrapa-CNPAT.
- Patto MV, Satovic Z, Pego S and Fevereiro P (2004). Assessing the genetic diversity of Portuguese maize germplasm using microsatellite markers. *Euphytica* 137: 63-72. <http://dx.doi.org/10.1023/B:EUPH.0000040503.48448.97>
- Pritchard JK and Wen W (2004). Documentation for structure software: Version 2.1. Chicago, USA. Available at [<http://pritch.bsd.uchicago.edu>]. Accessed October 21, 2015.
- Pritchard JK, Stephens M and Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Raposo A, Martins K, Ciampi AY, Wadt LD, et al. (2007). Diversidade genética de populações de andiroba no Baixo Acre. *Pesqui. Agropecu. Bras.* 42: 1291-1298. <http://dx.doi.org/10.1590/S0100-204X2007000900011>
- Rezende RKS, Paiva LV, Paiva R, Junior AC, et al. (2009). Divergência genética entre cultivares de gerbera utilizando marcadores RAPD. *Cienc. Rural* 39: 61-72. <http://dx.doi.org/10.1590/S0103-84782009005000176>
- Rivas LH, Giustina LD, Luz LN, Karsburg IV, et al. (2013). Genetic diversity in natural populations of *Theobroma subincanum* Mart. in the Brazilian Amazon. *Genet. Mol. Res.* 12: 4998-5006. <http://dx.doi.org/10.4238/2013.October.24.12>
- Rossi AAB, de Oliveira LO, Venturini BA and dos Santos Silva R (2009). Genetic diversity and geographic differentiation of disjunct Atlantic and Amazonian populations of *Psychotria ipecacuanha* (Rubiaceae). *Genetica* 136: 57-67. <http://dx.doi.org/10.1007/s10709-008-9309-z>

- Rossi FS, Rossi AAB, Dardengo JDFE, Brauwers LR, et al. (2014). Diversidade genética em populações naturais de *Mauritia flexuosa* L. f. (Arecaceae) com uso de marcadores ISSR. *Sci. Forum* 42: 631-639.
- Sacramento CK and Souza FX (2000). Cajá (*Spondias mombin* L.). FUNEP, Jaboticabal.
- Salimath SS, de Oliveira AC, Godwin ID and Bennetzen JL (1995). Assessment of genome origins and genetic diversity in the genus *Eleusine* with DNA markers. *Genome* 38: 757-763. <http://dx.doi.org/10.1139/g95-096>
- Santana IBB, Oliveira ED, Soares Filho WS, Ritzinger R, et al. (2011). Variabilidade genética entre acessos de umbu-cajazeira mediante análise de marcadores ISSR. *Rev. Bras. Frutic.* 33: 868-876. <http://dx.doi.org/10.1590/S0100-29452011005000090>
- Silva LM (2003). Superação de dormência de diásporos de cajazeira (*Spondias mombin*). Master's thesis, Universidade Federal de Lavras, Minas Gerais.
- Simão S (1998). Tratado de fruticultura. FEALQ, Piracicaba.
- Souza FX (2005). Crescimento e Desenvolvimento de Clones enxertados de Cajazeira na Chapada do Apodi, Ceará. *Rev. Bras. Frutic.* 28: 414-420. <http://dx.doi.org/10.1590/S0100-29452006000300017>
- Wright S (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19: 395-420. <http://dx.doi.org/10.2307/2406450>
- Yeh FC, Yang RC and Boyle T (1999). POPGENE: Microsoft Windows-based freeware for population genetic analysis, Release 1.31. University of Alberta, Canada.
- Zietkiewicz E, Rafalski A and Labuda D (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20: 176-183. <http://dx.doi.org/10.1006/geno.1994.1151>