Genetic diversity in native *Genipa americana* (Rubiaceae) populations in Sergipe, Brazil

M.N. Cardoso¹, A.L.S. Nascimento¹, L.A.R. de Oliveira¹, D.A. de Assunção², C.A. Machado¹, A.C.A. de Oliveira¹, A.S. de Jesus¹, A.S. Lédo², R.S. Archimínio³, A.R.C. Rabbani and A.V.C. Silva³

¹Programa de Pós-Graduação em Agricultura e Biodiversidade, Universidade Federal de Sergipe, São Cristóvão, SE, Brasil
²Universidade Federal de Sergipe, Aracaju, SE, Brasil
³Embrapa Tabuleiros Costeiros, Aracaju, Sergipe, Brasil

Corresponding author: M.N. Cardoso
E-mail: milenacardoso.bio@gmail.com

Genet. Mol. Res. 18 (1): gmr18119
Received August 17, 2018
Accepted February 05, 2019
Published February 26, 2019
DOI http://dx.doi.org/10.4238/gmr18119

**ABSTRACT.** Genipap (*Genipa americana* - Rubiaceae) is native to tropical America, occurring in Brazil in practically all biomes, except in the Pampas. It has socioeconomic importance, mainly due to the use of fruits in the manufacture of sweets and juices besides the medicinal properties of its leaves. We evaluated the genetic diversity of 73 individuals from 15 natural populations in the state of Sergipe using ISSR markers. The choice of areas was made randomly and genetic material was extracted from young leaves. PCR analysis using nine markers generated 113 fragments, which were used to estimate genetic diversity. The Shannon index was 0.36 and the similarity of individuals varied between 0.11 and 0.81. The individuals Salgado 3 (SAL3) and Salgado 4 (SAL4) were the most similar (0.81) and seven pairs had low similarity (0.11). The genetic distance the UPGMA analysis, which divided individuals into two distinct groups. The level of genetic variability found allows differentiation between genotypes that can be used to enrich the Genipap Active GermPlasm Bank in Sergipe, and the information generated will be useful for the conservation of these genetic resources and in future breeding programs for this species.
INTRODUCTION

Genipap, \( \textit{Genipa americana} \), also known in Brazil as jenipapo, janapabeiro and janipaba, is a native species cultivated in tropical America, extending from Mexico to Argentina (Delprete et al., 2005). It occurs in almost all Brazilian biomes, except the Pampas, being found in various types of forest formations located in humid or wet floodplains (Lorenzi, 2008). It belongs to the Rubiaceae family, with cosmopolitan distribution concentrated in the tropics, including approximately 550 genera and 9,000 species. In Brazil, there are about 120 genera and 1,400 species, corresponding to the most populous family of the Cerrado biome and it is an important element in almost all natural formations (Souza and Lorenzi, 2012; Ruzza et al., 2018).

Trees of this species are functionally monoecious and dioecious perennials, with reproduction preferentially by crosses (Sebbenn et al, 1998). It is an allogamous species, with great flexibility of genetic structuring, which allows better adaptation of plants to large-scale changes in the environment (Allard, 1971), generating variability within populations.

In folk medicine, leaves, bark, fruits and roots are employed, usually in the form of teas; in addition to medicinal use, it has food, dyeing and ornamental purposes (Souza et al., 2013). The usefulness of this species is a function of its chemical composition, mainly composed of iridoids such as genipin and geniposide (Alves et al., 2017), steroids (Conceição et al., 2011), fatty acids (Costa et al., 2011) and tannins (Moura et al, 2016).

The leaves have anticonvulsive (Nonato et al., 2018), antimalarial (Deharo et al., 2001) and trypanocidal properties (Souza et al., 2018). Phenolic extracts of fruits have the ability to inhibit the proliferation of cells responsible for human hepatocarcinoma (Finco and Groeve, 2013). The blue pigment extracted from green fruits is a promising alternative for industrial applications in food products due to the similarity with commercial dyes (Sigurdson et al., 2017), providing fabric staining, ninhydrin replacement in the chromatographic identification of amino acids and as a dye for colorimetric and fluorimetric analysis (Tokareva et al., 2017). The use of this species has also been recommended for the recovery of degraded areas due to its adaptive capacity under adverse conditions, such as soil flooding (Barbosa et al., 2007). Its importance justifies efforts towards conservation of this species.

Genetic erosion occurs as a consequence of forest fragmentation, resulting from factors such as urbanization, causing the reduction and isolation of natural vegetation, reducing intra- and inter-population variability (Young and Boyle, 2000). Allele fixation or loss occurs in small populations that can survive with few copies of certain alleles, increasing the likelihood of loss due to random events. Consequently, alleles disappear over time, making the population less genetically diversified (Bartlewicz et al., 2015). Consequently, studies involving diversity and genetic structure of remaining populations are important so that conservation strategies can be planned. In order to evaluate genetic variability, ISSR (Inter Simple Sequence Repeat) markers have been used to estimate the genetic diversity, based on the PCR technique, with a high polymorphism rate, reproducibility and widely used in population genetics (Wakte et al., 2012). Molecular
Genetic diversity of genipap markers in the estimation of the genetic diversity in fruit trees have also been used in mangaba (Soares et al., 2016); mulungú (Gonçalves et al., 2014) and cambui (Santana et al., 2016); they are fundamental in the generation of data that can be used in breeding programs and planning of conservation strategies, such as germplasm banks, which is a way to combine conservation of agronomic species with sustainable development (Nass, 2007).

In order to preserve the genetic resources of genipap, the Germplasm Active Bank of Embrapa Tabuleiros Costeiros, was implemented in 2009 in the city of Nossa Senhora das Dores, state of Sergipe. Currently, it is constituted of 24 accessions of different states of the federation and performs prospection, collection, enrichment and ex situ conservation activities. The first study on genetic diversity in this germplasm was carried out by Silva et al. (2014) using the first 18 accessions of this collection. We examined the genetic diversity of natural *G. americana* populations for inclusion in and eventual enrichment of the Active Germplasm Bank of Embrapa Tabuleiros Costeiros.

**MATERIAL AND METHODS**

**DNA Sampling, Collection and Extraction**

Young leaves were collected from 15 natural populations in the state of Sergipe, totaling 73 individuals (Figure 1). The collection expedition occurred in October 2016 and the choice of areas was random.

![Figure 1. Collection of *Genipa americana* accessions for evaluation of the genetic diversity.](image)

The leaves of each tree were collected separately, packaged in ice to avoid oxidation and then transported to the laboratory. Samples were kept in a freezer at -80°C until extraction of the genomic DNA, which was performed based on a method described by Doyle and Doyle (1990), modified by Alzate-Marín et al. (2009). DNA quantification was performed by spectrophotometry using Nanodrop 2000c (Thermo Scientific, USA). Samples were diluted in TE (Tris EDTA) at a concentration of 10 ng μL⁻¹, with a final volume of 50 μL per sample, and then stored at -20°C for subsequent use in PCR reactions.
PCR Amplification

For PCR assays, the total reaction volume was 20 μL, containing: 1 μL of genomic DNA solution (10 ng.μL⁻¹), 1 μL of each primer (10 μM), together with a composite mix of 2 μL 10X PCR buffer; 0.4 μl Dntp (10 mM); 0.6 μl MgCl₂ (50 mM); 0.2μL of Taq DNA polymerase (5U / μL) Ludwig and 14.8 μL of ultrapure water. For reaction amplification, an Axygen Maxygene thermocycler was programmed so that samples were denatured at 95°C for five min, followed by 45 cycles of amplification. At each cycle, samples were denatured at 94°C for 1 min, annealing at different temperatures for 1 min, first extension at 72°C for 45 s, and a final extension at 72°C for 10 min.

The amplification result was submitted to horizontal electrophoresis in a 2% agarose gel at a constant voltage of 182V, 91 mA and 17 W for 115 min. Then, gels were immersed in a solution containing ethidium bromide (0.5 μL.mL⁻¹ of water) for 1 hour and visualized under ultraviolet light. For measurement of the fragment patterns, 5 μl of the 100 bp molecular weight marker was used (Promega, Madison, South Dakota, USA). The visualization of results was performed in Gel doc L pix photodocumentation equipment (Loccus Biotecnologia, Cotia, SP).

Data analysis

ISSR markers were converted into a binary matrix based on the presence (1) or absence (0) of fragments. The Jaccard’s coefficients used to estimate the genetic distance and the dendogram using the unweighted pair group method with arithmetic mean (UPGMA) were estimated by the TreeView software (Page, 1996). To determine the dendrogram robustness, data were run through 10,000 replicates using the FreeTree software (Pavlicek et al., 1999).

To determine the ideal number of amplified fragments necessary to study genetic diversity in natural populations, the correlation value (r) of the similarity and stress matrix values (E) was estimated. According to Oliveira et al. (2017), these values indicate the fit between the original matrix and the simulated matrix. The optimal number of fragments was calculated by the GENES software (Cruz, 2007), being considered satisfactory when the stress value was less than 0.05 (Kruskal, 1964) and the correlation was close to 1. The genetic variability distribution in each population and in each region was estimated based on the following variables: Shannon index (I), expected heterozygosity (He), number of expected alleles (Na), number of effective alleles (Ne) and principal coordinate analysis method (PCoA), based on the Jaccard’s similarity coefficients, using the Genalex software version 6.5.

RESULTS AND DISCUSSION

A total of 113 polymorphic fragments were generated from nine ISSR markers used to estimate the genetic diversity of natural populations (Table 1). The number of fragments ranged from five (HB9) to 20 (17898 B), with a mean of 12.5 fragments per primer.

Various other studies using ISSR markers obtained high polymorphism rates, allowing the evaluation of genetic divergence. In a natural cambui population, Santana et al. (2016) found 71 bands using 10 markers. Evaluating the genetic diversity in Capparis...
Genetic diversity of genipap (Capparaceae) in different geographic distributions, Tamboli et al. (2018) obtained 97 loci using four markers. Jimenez et al. (2015) characterized the genetic diversity of individuals from natural mangaba populations (*Harconia speciosa*) from six markers, resulting in 83 polymorphic bands.

**Table 1.** ISSR markers used in natural *G. americana* populations of the state of Sergipe, Brazil, with their respective sequences, annealing temperatures (AT), number of fragmented bands (NFB), polymorphic fragments (PF) and polymorphism percentage (%P).

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequences 5’ → 3’</th>
<th>AT</th>
<th>NFB</th>
<th>PF</th>
<th>%P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1789B</td>
<td>CACACACACACAGT</td>
<td>49.7</td>
<td>20</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>UBC 807</td>
<td>AGAGAGAGAGAGAGAGT</td>
<td>47.0</td>
<td>14</td>
<td>14</td>
<td>100</td>
</tr>
<tr>
<td>HB9</td>
<td>GTGTTGTTGTTGTTG</td>
<td>52.6</td>
<td>5</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>HB10</td>
<td>GAGAGAGAGAGACC</td>
<td>52.6</td>
<td>10</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>HB11</td>
<td>GTGTTGTTGTTGTC</td>
<td>52.6</td>
<td>8</td>
<td>8</td>
<td>100</td>
</tr>
<tr>
<td>HB13</td>
<td>GAGGAGGAGGC</td>
<td>49.3</td>
<td>10</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>UBC 811</td>
<td>GAGAGAGAGAGAGAC</td>
<td>46.8</td>
<td>11</td>
<td>11</td>
<td>100</td>
</tr>
<tr>
<td>ISSR 13</td>
<td>ACACACACACACACAGA</td>
<td>48.0</td>
<td>16</td>
<td>16</td>
<td>100</td>
</tr>
<tr>
<td>UBC 841</td>
<td>GAGAGAGAGAGAGAYC</td>
<td>58.8</td>
<td>19</td>
<td>19</td>
<td>100</td>
</tr>
<tr>
<td>Total average</td>
<td></td>
<td>12.55</td>
<td>12.55</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

The demonstration of efficiency in the number of ISSR markers used in the diversity of *G. americana* was obtained from 111 polymorphic fragments, where the correlation estimate (r) was 0.996 and the stress value was 0.0346 (Figure 2). Stress values less than or equal to 0.05 indicate that estimates are accurate (Kruskal, 1964).

![Figure 2](image_url)

**Figure 2.** Correlation coefficient and number of fragments obtained from the use of nine ISSR markers in the genetic diversity of natural *Genipa americana* populations obtained by the bootstrap method.

There was an inversely proportional relationship between the number of fragments analyzed and the variance (Figure 3). With analysis using the DBOOT software, it was observed that when 110 polymorphic fragments were generated, there was stabilization of the variation coefficient (<0.01%). Quantification of the minimum number of markers provides data that optimizes the use of resources and time, resulting in smaller number of representative markers of genome sampling for genetic diversity characterization (Gonçalves et al., 2014).
Studies on *G. americana* are all recent. In order to evaluate accessions from the Germplasm Bank of Embrapa Tabuleiros Costeiros, Silva et al. (2014) used 12 ISSR markers and found 100% polymorphism. Rabbani et al. (2012) used 12 RAPD markers to estimate the genetic variability among individuals in the Baixo São Francisco region of Sergipe and found 74% polymorphism. Such results indicate that molecular markers are efficient in examining heterogeneity among populations. The Nei diversity index (He), assuming Hardy-Weinberg equilibrium, was 0.22. The observed allele index (Na) was 1.98, effective alleles (Ne) was 1.34 and the Shannon index (I) was 0.36. These values reveal intermediate genetic diversity levels among populations, a fact related to the geographic proximity among matrices in which the plant material was collected. Under natural conditions, the H value is never zero, as there is incorporation of new alleles through crosses, as well as losses due to genetic drift (Silva et al., 2014).

ISSR markers have shown efficacy in studies with several families, among them genetic diversity in the Rubiaceae family (Gaafar et al., 2014; Neuba et al., 2014). In Apocynaceae, such as mangaba (*Hancornia speciosa* Gomes), Soares et al. (2016) observed a diversity index of 0.12 in natural populations of Sergipe. In the medicinal *Myrcia lundiana* species of the Myrtaceae family, a moderate diversity index of 0.15 was observed in a native population (Alves et al., 2017).

Based on ISSR markers, the similarity matrix presented values ranging from 0.11 to 0.81 (Table 2), with an overall mean of 0.33. The pairs formed by individuals Salgado 3 (SAL3) and Salgado 4 (SAL4) were the most genetically similar, with a value of 0.81. On the other hand, Estância 3 (EST3) and Pirambú 4 (PIR4), Estância 3 (EST3) and Pirambú 5 (PIR5), Estância 4 (EST4) and Japaratuba 1 (JAT1), Estância 3 (EST3) and Japaratuba 1 (JAT1), Estância 4 (EST4) and Pirambú 2 (PIR2), Salgado 2 (SAL2) and Barra dos Coqueiros 2 (BC2), Lagarto 5 (LAG5) and Pirambú 5 (PIR5) pairs had the lowest values (0.11), indicating greater differentiation among individuals.

![Figure 3. Variance of polymorphic fragments using ISSR markers among 73 *Genipa americana* genotypes.](image)

**Table 2.** Pairs of genotypes with higher and lower similarity values (S) based on the Jaccard’s Coefficient, calculated from nine markers in 73 *Genipa americana* individuals.

<table>
<thead>
<tr>
<th>Order</th>
<th>Highest similarity</th>
<th>S</th>
<th>Smallest similarity</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Salgado 3 x Salgado 4</td>
<td>0.81</td>
<td>Estância 3 x Pirambú 4</td>
<td>0.11</td>
</tr>
<tr>
<td>2</td>
<td>Salgado 3 x Salgado 7</td>
<td>0.80</td>
<td>Estância 3 x Pirambú 5</td>
<td>0.11</td>
</tr>
<tr>
<td>3</td>
<td>Pirambú 1 x Barra dos Coqueiros 3</td>
<td>0.75</td>
<td>Estância 4 x Japaratuba 1</td>
<td>0.11</td>
</tr>
<tr>
<td>4</td>
<td>Capela 1 x Pacatuba 4</td>
<td>0.74</td>
<td>Estância 3 x Japaratuba 1</td>
<td>0.11</td>
</tr>
<tr>
<td>5</td>
<td>Barra dos Coqueiros 4 x Barra dos Coqueiros 2</td>
<td>0.73</td>
<td>Estância 4 x Pirambú 2</td>
<td>0.11</td>
</tr>
<tr>
<td>6</td>
<td>Salgado 4 x Salgado 6</td>
<td>0.72</td>
<td>Salgado 2 x Barra dos Coqueiros 2</td>
<td>0.11</td>
</tr>
<tr>
<td>7</td>
<td>Lagarto 4 x Lagarto 3</td>
<td>0.72</td>
<td>Lagarto 5 x Pirambú 5</td>
<td>0.11</td>
</tr>
</tbody>
</table>
Crosses influence the genetic variability levels because they reduce the distance and increase the divergence among some genotypes (Silva et al., 2017). The finding of heterogeneity among the study individuals using ISSR markers demonstrates the efficacy of this technique for genetic diversity studies (Santana et al., 2016).

The genetic distance among 73 individuals from different locations was used to construct a dendrogram using the UPGMA method (Figure 4). Based on the Jaccard’s coefficient, two groups were identified. The largest group formed was C1, with 37 individuals. SAL3 and SAL4 (0.81) pairs were the closest individuals, found in group C2, composed of 36 individuals. The highest diversity pairs (0.11) were composed of individuals belonging to both groups. The high variability observed in groups can be explained by the preference for allogamy of species that are not yet domesticated (Oliveira et al., 2009). The subgroups formed suggest differentiation among divergent accessions with a common ancestor, a fact that can occur due to the geographic distance of matrixes (Silva et al., 2014).

Consistency in genetic similarity data is important for the establishment of an ex situ collection or germplasm exchange, since in such cases, only one individual would be selected (Santana et al., 2016). Therefore, such information can be used to enrich the Genipap germplasm bank.

Figure 4. Phylogenetic representation of the Unweighted Pair Group Method using the arithmetic mean (UPGMA) cluster algorithm estimated by the genetic similarity of Jaccard’s coefficient and bootstrap analysis (10,000x) and groups (C1 and C2) of 73 Genipa americana individuals.
Genetic distances were also used in principal component analysis. Two clusters were identified, and the sum of the first two coordinates of the generated graph explained 58.47% of the variability (Figure 5). These results demonstrate the efficiency of the genetic diversity study using ISSR markers.

Figure 5. Principal Component Analysis (PCA) of natural *Genipa americana* populations of the states of Sergipe and Bahia, Brazil.

The use of more than one clustering method, according to hierarchization, optimization and ordering of groups allows a more concise classification, since it uses criteria for each technique used, preventing erroneous inferences about the classification of the material within a certain subgroup of genotypes (Silva et al., 2012).

**CONCLUSIONS**

We found intermediate genetic diversity in natural *G. americana* populations in the Brazilian state of Sergipe. The genetic reliability level obtained through analyses via ISSR markers allows differentiation among genotypes that can be used to enrich the Genipap Active Bank, and among the specimens evaluated, EST3, PIR4, PIR5, EST4, JAT1, PIR2, SAL2, BC2 and LAG5 individuals are recommended because they are genetically more distant.

**ACKNOWLEDGMENTS**

Research supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária).

**REFERENCES**

Genetic diversity of genipap


Genetics and Molecular Research 18 (1): gmr18119 ©FUNPEC-RP www.funpecrp.com.br


