Diversity and genetic structure of the native Brazil nut tree (*Bertholletia excelsa* Bonpl.) population

J.C. Cabral¹, A.B. Baldoni², H. Tonini², V.C.R. Azevedo³, L.D. Giustina¹, A.V. Tiago¹ and A.A.B. Rossi¹

¹Laboratório de Genética Vegetal e Biologia Molecular, Faculdade de Ciências Biológicas e Agrárias, Centro de Tecnologia da Amazônia Meridional, Universidade do Estado de Mato Grosso, Campus de Alta Floresta, Alta Floresta, MT, Brasil
²Embrapa Agrossilvipastoril, Sinop, MT, Brasil
³Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, Brasil
⁴Programa de Pós-Graduação em Biotecnologia e Biodiversidade da Rede em Associação Ampla, Pró-Centro Oeste Universidade Federal de Mato Grosso, Cuiabá, MT, Brasil

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

Genet. Mol. Res. 16 (3): gmr16039702
Received April 17, 2017
Accepted June 13, 2017
Published July 6, 2017
DOI http://dx.doi.org/10.4238/gmr16039702

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.** The present study was carried out in a native Brazil nut tree population (*Bertholletia excelsa* Bonpl., Lecythidaceae) to assess its genetic diversity and structure. Ten microsatellite markers were used to genotype 198 adult trees (*B. excelsa*). The population presented high genetic diversity and inbreeding absence rates. The empirical Bayesian method showed three distinct groups in the structure of this population. Molecular analysis of variance showed 98% variability within groups, and 2% between groups. The genetic divergence (*F*<sub>ST</sub>) indicated little
difference between groups; thus, suggesting efficient gene flow between the analyzed *B. excelsa* adult trees.

**Key words:** *Bertholletia excelsa*; Genetic diversity; Amazonia

**INTRODUCTION**

The Brazil nut tree (*Bertholletia excelsa* Bonpl.) is native to the Amazon and presents large and irregular distribution in “terra firma” (not flooded) forests throughout the Amazon Basin (Mori and Prance, 1990). In the Brazilian Amazon, Brazil nut trees are found in Pará, Rondônia, Amazonas, and Acre States - Brazilian Amazon, as well as in Amapá, Roraima, Maranhão, and Mato Grosso States (Müller et al., 1995).

Brazil nut tree seeds have high nutritional value and play a relevant role in the Amazonian international trade because it is one of the main extractive products in regions’ exportation agenda (Salomão, 2009). The harvest often takes place in natural populations and may represent up to 93% of the fruit picking from exploited trees (Zuidema and Boot René, 2002). The intense nut harvesting over decades is a threat to the species’ natural regeneration. However, the ecological sustainability of Brazil nut harvesting was analyzed in several studies (Wadt et al., 2008; Scoles and Gribel, 2011; Ribeiro et al., 2014; Scoles and Gribel, 2015) according to which the current regeneration of the species is sufficient for its persistence in the study sites; nonetheless, genetic diversity conservation, besides maintaining tree population, is fundamental. Effective species conservation strategies must take the genetic, ecological, and population data into account since the genetic variability maintenance is a key to this process (Yeh et al., 1996). Knowing genetic diversity within and between populations helps to understand their history because the current genetic diversity levels can influence the success of future populations (Erickson et al., 2004). The awareness of population genetic structures and diversity levels is an essential instrument for breeding programs (Erickson et al., 2004).

There are several techniques available to detect genetic diversity at DNA sequence level. Molecular markers can be used in studies focused on the genetic diversity between individuals, and within and between populations (Souza et al., 2008). Microsatellite markers, also known as SSRs, are one of the most available polymorphic nowadays. They are codominant expressions of high replicability and present frequent and random distribution, which allows wide genome coverage (Caixeta et al., 2016).

Diversity and genetic structure studies using microsatellite markers to assess *B. excelsa* (Sujii et al., 2015) were successful but remain limited if one considers the species’ great socio-economic importance. Accordingly, the aim of the present study was to assess the genetic diversity and population structure of a native *B. excelsa* population by using microsatellite markers.

**MATERIAL AND METHODS**

**Study site and plant material**

The study was carried out in Itaúba County, Mato Grosso State (MT) that is an area presenting a large extension of original Brazil nut tree formations. The experimental area is located approximately 30 km from Itaúba County, at Santo Ângelo Farm legal reserve.
Genetic diversity of Brazil nut trees (*Bertholletia excelsa*) belonging to the Dal Pai Group. An 18.8 hectare (400 x 470 m) sub-area was installed in the site; 198 adult Brazil nut trees were sampled within it. A vascular cambium disk was extracted from each tree, approximately 1.30 m up in the stem and transferred to 2.0 mL specimen transport medium (STM; Digene Corporation) (300 μL 2% CTAB, 700 μL absolute ethanol). The material was stored at -20°C.

**DNA extraction**

The total genomic DNA extraction followed the CTAB method proposed by Doyle and Doyle (1987), with modifications by Machado et al. (2002). The DNA verification was performed in 1% agarose gel electrophoresis and stained with ethidium bromide at a concentration of 0.6 ng/μL. The DNA quantification was performed in small volume spectrophotometer (NanoDrop 2000 system, Thermo Scientific) and the DNA samples were diluted to a concentration of 3 ng/μL. The laboratory analyzes were performed in the Molecular Biology Laboratory of Embrapa Agrossilvipastoril, Sinop, MT, and at Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF.

**Genotyping**

Ten microsatellite markers (primers) selected for *B. excelsa* were used in genotyping; four of them were developed by Suji et al. (2013) and identified as “Bet”; six were developed by Reis et al. (2009) and identified as “Bex” (Table 1).

<table>
<thead>
<tr>
<th>Loci</th>
<th>Reason</th>
<th>Classification</th>
<th>Sequence primer (5'→3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bet12</td>
<td>(TC)7</td>
<td>Simple perfect</td>
<td>F: ATAAGGACCGCCCACTCATC R: ATACGGGAGCAAGACCTGGTGAAC</td>
</tr>
<tr>
<td>Bet14</td>
<td>(AG)14</td>
<td>Simple perfect</td>
<td>F: GGTGACTTCTGCTGGGGAAC R: CCGAGTCTATTACCAAACT</td>
</tr>
<tr>
<td>Bet15</td>
<td>(AG)18 (AGA)13</td>
<td>Perfect composite</td>
<td>F: ACTGGCATCACCAGCATGAG R: GTCTTCCTGTTGTCACAAAT</td>
</tr>
<tr>
<td>Bet16</td>
<td>(CT)7 (CCCT)3</td>
<td>Perfect composite</td>
<td>F: TTGATCTTCCAGGTCGGTGTAG R: ACTCTCTAATCCATGAGT</td>
</tr>
<tr>
<td>Bex02</td>
<td>(CT)7 (CA)3 (CT)7 (CA)3</td>
<td>Perfect composite</td>
<td>F: GCCATGTTCCTACAGTCCTG R: AGTCGGCATCCTGGTCCT</td>
</tr>
<tr>
<td>Bex09</td>
<td>(CT)12</td>
<td>Simple perfect</td>
<td>F: TAYTCCAGTTCTCTGGGTAG R: AYCTAATCATCTTCAAGAGT</td>
</tr>
<tr>
<td>Bex22</td>
<td>(CT)10</td>
<td>Simple perfect</td>
<td>F: GCATCTCTCATTTTCCGGCTG R: CCCTAGCAATGGTCTCTCTC</td>
</tr>
<tr>
<td>Bex27</td>
<td>(GA)6</td>
<td>Simple perfect</td>
<td>F: ACTGTGATGACGCAGCATG R: TTTGGCATCAGAATACT</td>
</tr>
<tr>
<td>Bex33</td>
<td>(CT)7</td>
<td>Simple perfect</td>
<td>F: CAGGTTCATCGACAGATGTC R: ACCAGTGTCAGACAGATGTC</td>
</tr>
<tr>
<td>Bex37</td>
<td>(CT)9</td>
<td>Simple perfect</td>
<td>F: TGGATCTGCTTGATCTGCTG R: CACGCAACCTCAGTCTGGT</td>
</tr>
</tbody>
</table>

The amplification reactions were performed in a total volume of 12 μL containing 3 μL DNA (3 ng/μL), 1.2 μL 10X PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 1.2 μL dNTP (2.5 mM), 1.92 μL BSA (2.5 mg/mL), 0.24 μL Taq DNA Polymerase (5 U/μL), 1.5 μL of each primer (1 μM), and 1.44 μL ultrapure water. They were conducted in a thermal cycler (Applied Biosystems) under the following conditions: an initial denaturation cycle, at 94°C, for 5 min,
which was followed by 30 cycles of a 1-min denaturation, at 94°C, 1-min annealing, at 64°C, and an extension at 72°C, for 30 s, and a 7-min final extension at 72°C (Sujii et al., 2013). Primers were labeled using different fluorochromes to enable analyzing the PCR products and the duplex and multiplex systems in the ABI 3100 DNA analyzer (Applied Biosystems). The amplified fragments were detected in the GeneMapper software (v.4., Applied Biosystems).

Data analysis

The genetic diversity of the population was characterized by the mean number of alleles per locus \( (N_A) \), allelic frequency, observed heterozygosity \( (H_O) \) and expected heterozygosity \( (H_E) \) through the Hardy-Weinberg equilibrium. The Cervus 3.0.3 software was used to estimate and find the fixation index (Kalinowski et al., 2007).

The Bayesian analysis was performed in the Structure software (Pritchard et al., 2000). The number of K-groups was adjusted for 1 to 5 variations, with 20 independent interactions, using 250,000 burn-ins and 750,000 Markov chain Monte Carlo simulations, in each interaction. The genetic structure was analyzed in the Structure Harvester software (Earl and VonHoldt, 2012), according to the criterion by Evanno et al. (2005). Analysis of molecular variance (AMOVA) applied to the groups set the Structure was calculated in the GenAlEx 6.5 software (Peakall and Smouse, 2012).

RESULTS AND DISCUSSION

The sample of 10 microsatellite loci presented 38 alleles, in total, which ranged from 2 (Bet16 and Bex02) to 7 (Bex33), and 3.8 alleles per locus, on average (Table 2). Reis et al. (2009) found 10.4 alleles per \( B. \) excelsa locus, on average, a number that is larger than the one identified in the present study. On the other hand, Sujii et al. (2015) studied the genetic diversity structure of five \( B. \) excelsa populations and found a range from 3.36 to 6.18 alleles per locus, on average, thus corroborating the herein found results.

<table>
<thead>
<tr>
<th>Loci</th>
<th>Amplitude amplification (bp)</th>
<th>N</th>
<th>( N_A )</th>
<th>( H_O )</th>
<th>( H_E )</th>
<th>PIC</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bet12</td>
<td>112-118</td>
<td>3</td>
<td>189</td>
<td>0.646</td>
<td>0.582</td>
<td>0.498</td>
<td>-0.073</td>
</tr>
<tr>
<td>Bet14</td>
<td>105-121</td>
<td>5</td>
<td>190</td>
<td>0.432</td>
<td>0.617</td>
<td>0.540</td>
<td>0.165</td>
</tr>
<tr>
<td>Bet15</td>
<td>188-200</td>
<td>3</td>
<td>185</td>
<td>0.481</td>
<td>0.475</td>
<td>0.369</td>
<td>-0.008</td>
</tr>
<tr>
<td>Bet16</td>
<td>128-130</td>
<td>2</td>
<td>188</td>
<td>0.947</td>
<td>0.501</td>
<td>0.375</td>
<td>-0.308</td>
</tr>
<tr>
<td>Bex02</td>
<td>108-112</td>
<td>2</td>
<td>187</td>
<td>0.171</td>
<td>0.157</td>
<td>0.144</td>
<td>-0.037</td>
</tr>
<tr>
<td>Bex09</td>
<td>124-138</td>
<td>4</td>
<td>193</td>
<td>0.622</td>
<td>0.654</td>
<td>0.599</td>
<td>0.028</td>
</tr>
<tr>
<td>Bex22</td>
<td>128-150</td>
<td>3</td>
<td>193</td>
<td>0.591</td>
<td>0.616</td>
<td>0.535</td>
<td>0.014</td>
</tr>
<tr>
<td>Bex27</td>
<td>117-137</td>
<td>4</td>
<td>181</td>
<td>0.365</td>
<td>0.368</td>
<td>0.340</td>
<td>-0.015</td>
</tr>
<tr>
<td>Bex33</td>
<td>213-247</td>
<td>7</td>
<td>192</td>
<td>0.432</td>
<td>0.467</td>
<td>0.423</td>
<td>0.036</td>
</tr>
<tr>
<td>Bex37</td>
<td>185-207</td>
<td>5</td>
<td>173</td>
<td>0.434</td>
<td>0.482</td>
<td>0.424</td>
<td>0.041</td>
</tr>
<tr>
<td>Average</td>
<td>112-118</td>
<td>3.8</td>
<td>187.7</td>
<td>0.512</td>
<td>0.491</td>
<td>0.424</td>
<td>-0.016</td>
</tr>
</tbody>
</table>

N: number of individuals; \( N_A \): number of alleles; \( H_O \): observed heterozygosity; \( H_E \): expected heterozygosity; PIC: polymorphism information content, F: fixation index.

Polymorphism information content (PIC) ranged from 0.144 (Bex02) to 0.599 (Bex09); mean PIC = 0.424 (Table 2). According to Botstein et al. (1980), PIC values below 0.25 are less informative, those between 0.25 and 0.50 are moderately informative, and the values above 0.50 are highly informative. Thus, the Bex02 locus is not informative or
Genetic diversity of Brazil nut trees (*Bertholletia excelsa*)

recommended for further studies about the species. The loci Bet14, Bex09, and Bex22 are highly informative because they presented values higher than 0.50, and the remaining loci were moderately informative (Table 2).

The mean $H_o$ and $H_e$ (0.512 and 0.491, respectively) indicate population genetic diversity. Sujii et al. (2015) studied the genetic diversity of *B. excelsa* and observed higher $H_o$ than $H_e$. The high heterozygosity rate presented by Brazil nut populations may be related to the breeding system. Cavalcante et al. (2012) noticed *B. excelsa* self-incompatibility and the absence of fruit generation when the plant receives pollen from its flowers. Self-incompatibility helps to maintain genetic diversity within populations since it favors cross-breeding. Other studies involving arboreal species have indicated the possible selection for the heterozygote of self-incompatible species (Sebbenn et al., 2000; Ribas and Kageyama, 2004).

Table 2 shows mean fixation index = -0.016, thus indicating the high heterozygosity and the low inbreeding of the herein studied *B. excelsa* population. According to Kageyama et al. (2003), the fixation index is one of the most important population genetic parameters, since it measures the balance between homozygotes and heterozygotes.

Based on the Bayesian analysis performed in the STRUCTURE software and on the ΔK method described by Evanno et al. (2005), we identified three groups of structures in the *B. excelsa* population, as well as overlaps between them. The ΔK value was higher than that of the remaining K in K = 3 (Figure 1A). Group I comprised 73 individuals; group II, 68; and group III, 57 (Figure 1B).

![Figure 1. A. ΔK graphic representation of each K value (group) based on the 10 SSR locus information found in the Structure Harvester software. B. Representation of the 198 individuals from the native *Bertholletia excelsa* population. They are represented by the vertical color bars (red, green, and blue), depending on the genetic group that they belong.](image-url)
AMOVA of the groups formed in the Structure software showed 98% variability within groups and 2% variability between groups (Table 3).

Such genetic structure pattern meets the studies in the literature: tropical tree species reproducing through allogamy maintain greater genetic variability proportion within populations than between populations (Zucchi et al., 2005).

Costa et al. (2015) studied the genetic structure of a Nectandra megapotamica population and found greater diversity within groups (75.11%) than among groups (24.89%). Soares (2016) observed the same effect when he studied the genetic diversity of a native Hancornia speciosa population; the greatest genetic variation occurred within groups (83.39%), rather than between groups (16.61%).

Table 3. Analysis of molecular variance (AMOVA) of native Bertholletia excelsa populations divided into groups, according to the molecular data from 10 SSR loci found in the “Structure” software.

<table>
<thead>
<tr>
<th>Variation</th>
<th>d.f.</th>
<th>SS</th>
<th>CV</th>
<th>TV (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>2</td>
<td>20.194</td>
<td>0.056</td>
<td>2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within groups</td>
<td>393</td>
<td>1055.826</td>
<td>2.687</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>395</td>
<td>1076.020</td>
<td>2.743</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

d.f. = degrees of freedom; SS = sum of squares; CV = coefficient of variation; TV = total variation; P = variation probabilities greater than the randomly observed values. Odds were calculated for 1000 random permutations.

The $F_{ST}$ value, a differentiation parameter responsible for identifying the allelic frequencies and the proportion of total variation of allelic frequencies occurring between subpopulations, was 0.021. It is also used to measure the effects of population subdivisions and the subpopulation heterozygosity reduction caused by genetic drift. The $F_{ST}$ values lower than 0.05 indicate little genetic differentiation and those higher than 0.5 confirm high genetic differentiation. The $F_{ST}$ value of the herein studied population revealed little differentiation between groups, thus corroborating AMOVA results and suggesting efficient gene flow between individuals belonging to the studied B. excelsa population. Zucchi et al. (2005) assumed that population variations have direct implications for conservation, and it indicates that a larger number of variations need to be sampled when the $F_{ST}$ value is high; on the other hand if it is low, a larger number of individuals per population should be sampled. Therefore, a large number of individuals had to be kept in situ or sampled ex situ for conservation in the current study.

The studied B. excelsa population presented genetic diversity and mean $H_o$ higher than $H_e$. The population structure was divided into three groups, and the greatest diversity was observed within groups. Therefore, the population conservation demands that a large number of individuals be maintained in situ or sampled ex situ for conservation. The absence of population inbreeding suggests that the species is not reproductively isolated and presents effective gene flow.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors wish to thank Conselho Nacional de Desenvolvimento Científico e
Genetic diversity of Brazil nut trees (Bertholletia excelsa)

Tecnológico (CNPq Process #472988/2013-1). J.C. Cabral received a Master’s scholarship from FAPEMAT/CAPES.

REFERENCES


Genetics and Molecular Research 16 (3): gmr16039702


