Genetic diversity in forest populations from conservation units in the Atlantic Rainforest in northeast Brazil

Ageu da Silva Monteiro Freire, Cristiane Gouveia Fajardo, Kyvia Pontes Teixeira das Chagas, Luciana Gomes Pinheiro, Fernanda Moura Fonseca Lucas, Fábio de Almeida Vieira

1 Universidade Federal do Rio Grande do Norte, Unidade Acadêmica Especializada em Ciências Agrárias, Macaíba, RN, Brazil. E-mail: ageu_monteiro_123@hotmail.com (ORCID: 0000-0002-5364-0859); genegoista00@gmail.com (ORCID: 0000-0001-6202-7143); kyviapontes@gmail.com (ORCID: 0000-0003-1361-3204); luciana.gpinheiro@yahoo.com.br (ORCID: 0000-0002-9115-4210); fernanda-sonseca@hotmail.com (ORCID: 0000-0002-3181-2568); vieirafa@gmail.com (ORCID: 0000-0003-3347-255X)

ABSTRACT: Fragmentation leads to changes in the size and dynamics of forest populations. Studies of genetic diversity support strategies for preserving the genetic resources, and the data to date for the Atlantic Forest in northeast Brazil is scarce. Thus, the objective of this study was to evaluate the genetic diversity of Protium heptaphyllum (Burseraceae) tree species in three populations. We sampled 64 individuals in this study, for which seven primers used generated 90 loci, with 93% being polymorphic. The average number of alleles was 1.59, the effective number of alleles was 1.34, Nei’s genetic diversity was 0.20, and the Shannon index was 0.30. AMOVA indicated higher genetic variation among individuals within populations than among populations. Bayesian analysis revealed the existence of two distinct genetic groups, in which the MAC (Macaíba) population was the most differentiated. Genetic bottleneck tests indicated that all populations presented significant reductions in the population size. In situ conservation strategies must be adequately maintained considering the current isolation of studied populations, the genetic bottleneck, and the low average genetic diversity detected. Also, further recovery of degraded areas in the vicinity and ex situ conservation are recommended.

Key words: bottleneck; forest fragmentation; ISSR; Protium heptaphyllum

Diversidade genética em populações florestais em unidades de conservação da Mata Atlântica, Nordeste do Brasil

RESUMO: A fragmentação ocasiona alterações no tamanho e na dinâmica das populações florestais. Estudos de diversidade genética subsidiam estratégias de conservação dos recursos genéticos, sendo escassos os dados para a Mata Atlântica no Nordeste do Brasil. O objetivo foi avaliar a diversidade genética da espécie arbórea Protium heptaphyllum (Burseraceae) em três populações. Foram amostrados 64 indivíduos, para os quais sete iniciadores utilizados forneceram 90 loci, sendo 93% polimórficos. O número médio de alelos observados foi de 1,59, o número de alelos efetivos foi 1,34, a diversidade genética média de Nei foi 0,20 e o índice de Shannon foi 0,30. A AMOVA indicou maior variação genética entre indivíduos dentro das populações do que entre as populações. A análise Bayesiana mostrou a existência de dois grupos genéticos distintos, onde a população MAC (Macaíba) foi a mais diferenciada. Os testes de gargalo genético indicaram que todas as populações apresentaram reduções significativas no tamanho populacional. As estratégias de conservação in situ devem ser adequadamente mantidas, considerando-se o atual isolamento das populações estudadas, o gargalo genético e a baixa diversidade genética média detectada. Além disso, recomenda-se a recuperação das áreas degradadas nas proximidades e a conservação ex situ.

Palavras-chave: gargalo genético; fragmentação; ISSR; Protium heptaphyllum
Introduction

*Protium heptaphyllum* (AUBL.) March is a native tree which belongs to the Buergeraceae botanical family, with a distribution in the phytoecological domains of Amazonia, Caatinga, Cerrado, and Atlantic Forest (Daly, 2015). It is popularly known as almecega and breu, and like other *Protium* species it has aromatic elemis resin triterpenes (Gigante, 2005). Although *P. heptaphyllum* is classified as unthreatened, it is listed among species with data deficiency, and therefore should be included in research and conservation studies (Martinelli & Moraes, 2013).

Its resin is used in popular medicine such as an analgesic, for healing and as an expectorant. It is also used in the industry as a varnish, in the caulking of vessels, and as incense in religious rituals. The essential oil of the resin has constituents such as limonene, trans-β-ocimene, α, β-amyrin, eucalyptol, and p-cymene, which have inflammatory (Amaral et al., 2009) and antimutagenic (De Lima et al., 2016) properties, and it has been known to reduce obesity (Carvalho et al., 2017). Due to this pharmacological potential, there is an immediate need to evaluate the remaining genetic resources of *P. heptaphyllum* for *in situ* and *ex situ* conservation strategies.

*P. heptaphyllum* has wide distribution in the Atlantic Forest, a biome that incorporates significant biological diversity in South America. However, the majority (80%) of the remaining forest fragments are less than 50 hectares (Ribeiro et al., 2009) in extent. The reduction and fragmentation of the Atlantic Forest began in the colonial period, where the loss of the biome was a consequence of the urbanisation processes, industrialisation and agricultural expansion (Pinto et al., 2014). The intense anthropogenic pressure continues today; the remnants suffer from the exploitation of the biodiversity resources (Fajardo et al., 2017).

Fragmentation causes changes in population size and dynamics (Laurance & Vasconcelos, 2009), modifying the mechanisms of pollen and seed flow, with impacts on genetic diversity (Vranckx et al., 2012). Studies related to population genetics are essential for defining species’ conservation strategies and management (Fajardo et al., 2017). Due to the immense richness of forest species in the tropics, studies regarding the genetic diversity with species of ecological, economic, and medicinal interest are priorities. In fact, these kinds of species are threatened by the loss of natural habitats and intense collection, making it necessary to elaborate strategies for the conservation of the remaining genetic resources (Vieira et al., 2015).

Molecular markers enable detecting genetic polymorphism in DNA, which studies point out as an essential tool in studies of genetic diversity, aiding in parameters such as variability, structure and both intrapopulational and interpopulational distribution (Adhikari et al., 2017). Among commonly used DNA markers to analyse genetic diversity, ISSR markers (inter-simple sequence repeat) are lower cost markers with acceptable and reproducible polymorphism (Chagas et al., 2015).

Therefore, we asked the following questions regarding the genetic diversity of *P. heptaphyllum* in natural populations: What are the genetic diversity levels across the studied fragmented Atlantic forest? Has there been a reduction in effective population size (i.e. genetic bottlenecks)? Do populations located in conservation units have similar genetic diversity? Consequently, the objective of this study was to analyse the genetic diversity of natural *P. heptaphyllum* populations, evaluating the genetic structure and the occurrence of population/genetic bottlenecks.

Material and Methods

Sampling

We sampled 64 individuals with a minimum distance of five meters between them, distributed in three natural populations in fragments of Atlantic Forest located in Rio Grande do Norte state, Brazil (Table 1 and Figure 1). Young leaf samples from all plants were conditioned in 2 mL plastic tubes containing 2X CTAB (cetyltrimethylammonium bromide). We registered the samples in the "National System for Management of Genetic Heritage and Associated Traditional Knowledge” (Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado - SISBIO)” under the registration code A0FBC35.

The MAC population is in a small forest area with sandy soil and near a weir, having trails, fields for crops and houses

![Figure 1. Remnants of Atlantic Forest in the Rio Grande do Norte, and location of the three fragments where Protium heptaphyllum populations were sampled. Source: SOS Mata Atlântica (2018).](image)

Table 1. Populations and sample size (n) for Protium heptaphyllum, area and altitude of the forest fragments.

<table>
<thead>
<tr>
<th>Population</th>
<th>Codes</th>
<th>Geographical coordinates</th>
<th>n</th>
<th>Area (ha)</th>
<th>Altitude (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macaíba</td>
<td>MAC</td>
<td>5°53’25”S, 35°21’06”W</td>
<td>19</td>
<td>10</td>
<td>55</td>
</tr>
<tr>
<td>Nisia Floresta</td>
<td>NIF*</td>
<td>6°05’09”S, 35°11’04”W</td>
<td>20</td>
<td>160</td>
<td>65</td>
</tr>
<tr>
<td>Baía Formosa</td>
<td>BAF*</td>
<td>6°22’31”S, 35°01’09”W</td>
<td>25</td>
<td>1800</td>
<td>53</td>
</tr>
</tbody>
</table>

* Brazilian Conservation Units
in its neighbourhood. Among three populations, MAC is the furthest from an urban centre (4 km) and the Atlantic Ocean (20 km). This population is the only one that is not located in any conservation unit.

The NIF population is in a Brazilian National Forest Conservation Unit (FLONA). It has an average distance from an urban centre of 3 km and is 9.5 km from the Atlantic Ocean. It is in a location with sandy soil and some declivities with water accumulation. The surrounding environment is impacted by roads, some houses and areas used for crops.

Finally, the BAF population is in a Brazilian conservation unit classified as Private Reserve of Natural Heritage (RPPN). It has the largest area among the three populations, with greater proximity to an urban centre (0.4 km) and the Atlantic Ocean (0.3 km). The environment is humid due to slopes which favour an accumulation of water, although it only has dunes. The neighbourhood area is the most fragmented among the three populations, with several roads and excessive tourist movement. It also has extensive areas used for sugarcane cultivation in the neighbourhood.

DNA Extraction and PCR
We utilized the method described by Doyle & Doyle (1987) to extract DNA. DNA quantification was assessed by spectrophotometer (Epoch™, Bio Tek Instruments Inc, Winooski, USA) and diluted in TE (10 mM Tris-HCl; mM EDTA pH 8.0), with final concentration for the use of 50 ng μL⁻¹.

The PCR was performed using seven ISSR primers (Table 2) after previous selection of 30 primers (UBC primer set #9, Vancouver, Canada). The annealing temperature was 47 °C. The PCR mix was composed of buffer (10 X), BSA (1 mg mL⁻¹), MgCl₂ (50 mM), dNTP (2.5 mM), primer (2 μM), Taq polimerase (5.0 U μL⁻¹), DNA and ultrapure water, with a final volume of 12 μL. The PCRs were performed in a thermocycler (Veriti®, Applied Biosystems, Woodlands, Singapore) at 94 °C for 5 minutes, followed by 37 amplification cycles at 94 °C (15 seconds), 47 °C (30 seconds) and 72 °C (1 minute). We included a final extension at 72 °C (7 minutes) with cooling at 4 °C.

PCR products were subjected to electrophoresis on 1.5% agarose gel (v/v) in TAE buffer (1.0 X, EDTA Tris-acetate) and stained with GelRed™. We used a molecular weight DNA ladder of 1 kb. The gels were photographed with ultraviolet light (E-Box™ VX2, Vilber Lourmat, Marne la Valle, France). We selected primers with better resolution and a more significant number of fragments, excluding those which presented bands with low intensity or low definition.

Data analysis
The polymorphism information content (PIC) was calculated using the Equation 1:

\[
PIC_i = 1 - \sum_{j=1}^{n} \left( \frac{P_{ij}}{n} \right) ^2
\]

where \( P_{ij} \) is the frequency of the allele “\( j \)” in the marker “\( i \)” (Anderson et al., 1993). We used the PopGene 1.32 program (Yeh et al., 1997) to obtain the percentage of polymorphic loci (%P), number of observed alleles (Na), number of effective alleles (Ne), Nei’s genetic diversity (h) and Shannon index of genetic diversity (I) for each population. The diversity indexes h and I were submitted to analysis of variance (ANOVA) using BioEstat 5.3 program (Ayres et al., 2007) to verify the statistical difference between the populations. We used the Kruskal-Wallis ANOVA for non-parametric data, which were confirmed by the Lilliefors normality test (Ayres et al., 2007).

The analysis of molecular variance (AMOVA) was estimated using the Alerquim 3.5 program (Excoffier & Lischer, 2010). We constructed a dendrogram of the genetic identity of Nei (1978) among the populations using an Unweighted Pair Group Method with Arithmetic mean (UPGMA). The consistency of the cluster nodes was analysed through 1,000 permutations in the TFPGA 1.3 program (Miller, 1997). We analysed the correlation between the Nei’s genetic distances and geographic distance using the Mantel test and GenAlex 6.503 program (Peakall & Smouse, 2012). The Monte Carlo test was performed with 999 permutations.

We used a Bayesian analysis approach to determine the number of genetic groups (K) representing the populations. The value of K varied from 1 to 3 and was estimated through the admixture ancestry model based on the frequency of the correlated alleles using Structure 2.3.4 program (Pritchard et al., 2000). Ten independent runs were performed for each K using 250,000 Markov chain Monte Carlo simulations (MCMC), and a burn-in of 500,000. The value of K was measured according to the ΔK method (Evanno et al., 2005) as implemented in the Structure Harvester program (Earl & Vonholdt, 2012).

The significant reductions in the effective size population were evaluated in the Bottleneck 1.2.02 program (Cornuet & Luikart, 1996). We utilised the IAM (infinite allele model) and SMM (stepwise mutation model) models. The significance was evaluated through the sign test (\( \alpha = 0.05 \)) from the frequency of alleles (Cornuet & Luikart, 1996).

Results
We analysed 90 loci using seven primers (Table 2), ranging from 10 to 16 loci per primer and with a mean of 13. Among the 90 observed loci, 84 (93%) were polymorphic (Table 3). The NIF population presented the highest percentage (66%) of
polymorphic loci. The average Nei’s genetic diversity (h) and Shannon (I) indexes were respectively 0.20 (± 0.04) and 0.30 (± 0.06). The three populations showed no statistical differences between the genetic diversity indexes of h (H = 3.51, p = 0.17) and I (H = 3.51, p = 0.17) according to the ANOVA test (Kruskal-Wallis).

The lowest genetic distance was found between NIF and BAF populations (0.08), and the highest genetic distance was between the MAC and BAF populations (0.21) (Table 4). The Mantel test indicated a positive, however not significant, correlation between the genetic and geographical distances of the populations (r = 0.27; p = 0.50).

AMOVA indicated that there is greater genetic variation among individuals within populations (55.64%) than among populations (44.36%) (Table 5). However, the genetic differentiation between populations (ΦST = 0.44) was still considered to be high.

The UPGMA method of clustering illustrates the greater similarity between the NIF and BAF populations, which are divergent from the MAC population (Figure 2).

Two distinct genetic groups (K = 2) were additionally observed according to the δK values of the Bayesian analysis (Figure 3) and proportion of genotypes (Figure 4).

Bottleneck tests indicated that all populations had a genetic bottleneck in the SMM model (p < 0.05). In the IAM model, a bottleneck was only detected in the BAF population (Table 6).

**Discussion**

The observed loci average per primer (13) was similar to some other studies with forest species. Rossi et al. (2014) studied *Mauritia flexuosa* and found 97 loci from nine primers, with an average of 11 loci per primer. Chagas et al. (2015)
Table 6. Bottleneck tests for the Protium heptaphyllum populations using IAM and SMM models.

<table>
<thead>
<tr>
<th>Populations</th>
<th>IAM</th>
<th>SMM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Hd/He</td>
</tr>
<tr>
<td>MAC</td>
<td>42.87</td>
<td>36</td>
</tr>
<tr>
<td>NIF</td>
<td>42.74</td>
<td>36</td>
</tr>
<tr>
<td>BAF</td>
<td>41.20</td>
<td>30</td>
</tr>
</tbody>
</table>

n = expected number of loci with excess heterozygosity under the respective model; Hd/He = number of loci with heterozygosity deficit and excess heterozygosity, respectively; p = probability; * = significant at 5% probability.

identified an average of 11 when observing 63 loci in Elaeis guineensis through six primers.

The PIC value ranged from 0.34 to 0.50 (Table 2), classifying all primers as moderately informative. The UBC 842 primer presented a higher PIC value, similarly to Costa et al. (2015) who also observed greater PIC value in this primer, indicating higher polymorphic content in comparison to others.

The observed average genetic diversity of 0.20 was lower than expected for species which show a similar life story. According to Nybom (2004), the expected genetic diversity for long-lived perennial species is 0.25. Genetic diversity was also lower than expected for cross-fertilisation species (0.27) and with seed dispersal through ingestion (0.24) (Nybom, 2004). The Protium genus is composed of dioecious species (Vieira et al., 2010), with this also being confirmed for Protium heptaphyllum (Silva et al., 2014). Although no deviations in the 1:1 sexual ratio were observed in the study area (MAC population) (Silva et al., 2014), the fragmented landscape is a threat to maintain the remaining genetic diversity. Vranckx et al. (2012) showed that forest fragmentation could cause the loss of genetic variability due to population decrease.

Furthermore, P. heptaphyllum has fruits which are dispersed by birds (Pizo, 2004), which can be affected in a fragmented landscape resulting in negative consequences (Khimoun et al., 2017). Over the generations, the loss of alleles may increase the genetic divergence between populations and reduce the evolutionary potential. Thus, the geographic isolation, crossings rising between related individuals, and gene flow among populations are amongst the main factors to consider in the maintenance of genetic diversity.

Although NIF and BAF populations have a greater geographic distance, both are more genetically similar and occur in conservation units near the coastline, and in humid environments. In fact, the Bayesian analysis showed that these populations have well-defined genotypes, with MAC being the most differentiated from the others, and so it is necessary to seek conservationist actions for the area since it does not have protection measures due to it being the only one which is not a conservation unit. P. heptaphyllum presented high genetic differentiation among the populations (44.36%), which should be considered in conservation strategies. The current scenario indicates a higher likelihood of cross-breeding occurrence between related plants, considering the high degree of deforestation and geographical isolation of the populations. Thus, studies that evaluate the contemporary gene flow within and among the populations are suggested, aiming at understanding the dynamics of the allele frequencies in the fragmented landscape.

The coastal region in northeast Brazil is marked by intense human occupation and land use since the colonial period (Pinto et al., 2014). Although the BAF population is located in the largest fragment and conservation unit, it presented a genetic bottleneck in both tested models, in addition to having the lowest genetic polymorphism. The neighbouring BAF population area is predominantly urban with extensive sugar cane cultivation areas. In addition to the Atlantic Forest deforestation over the years, some studies have already demonstrated the damaging effects of monoculture in the region (Oliveira & Mattos, 2014; Govindin & Miller, 2015). Therefore, urgent actions are needed to conserve the remaining genetic resources of P. heptaphyllum, as well as to restore the degraded environments.

Conclusions

We observed low genetic diversity in P. heptaphyllum populations located in a fragmented landscape. There is an immediate need for the effective preservation of forest remnants in the Atlantic Forest region to maintain the genetic diversity of P. heptaphyllum and other species. Also, the creation of P. heptaphyllum germplasm banks to constitute the genetic basis for degraded area recovery programs are necessary for the long-term. The MAC population occurs in the smallest fragment and outside the conservation unit, and has the most significant genetic differentiation with other populations, making it a priority for conservation.

Acknowledgements

We thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for providing a scientific initiation scholarship for ASMF and for FMFL. This work was supported by the Fundação de Apoio à Pesquisa do Rio Grande do Norte (FAPERN) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), process no. 23038.004469/2012 and Finance Code 001. We also thank Professor Michael B. Usher for assistance in editing the English of the manuscript.

Literature Cited


