Research Article

Using Random Amplified Polymorphic DNA to Assess Genetic Diversity and Structure of Natural Calophyllum brasiliense (Clusiaceae) Populations in Riparian Forests

Evânia Galvão Mendonça,1 Anderson Marcos de Souza,2 Fábio de Almeida Vieira,3 Regiane Abjaud Estopa,4 Cristiane Aparecida Fioravante Reis,5 and Dulcinéia de Carvalho6

1 Departamento de Engenharia Florestal, Universidade Federal Rural do Rio de Janeiro, 23890-000 Rio de Janeiro, RN, Brazil
2 Departamento de Engenharia Florestal, Universidade de Brasília, 70919-900 Brasília, DF, Brazil
3 Unidade Acadêmica Especializada em Ciências Agrárias, Universidade Federal do Rio Grande do Norte, 59280-000 Macaíba, RN, Brazil
4 Klabin, Avenida Brasil 26, Bairro Harmonia, 84275-000 Têlmaco Borba, PR, Brazil
5 Embrapa Florestas, 74001-970 Goiânia, GO, Brazil
6 Departamento de Ciências Florestais, Universidade Federal de Lavras, 37200-000 Lavras, MG, Brazil

Correspondence should be addressed to Fábio de Almeida Vieira; vieirafa@yahoo.com.br

Received 28 April 2014; Revised 26 July 2014; Accepted 28 July 2014; Published 21 August 2014

Academic Editor: Robin Reich

Copyright © 2014 Evânia Galvão Mendonça et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The objective of this study was to assess the genetic variability in two natural populations of Calophyllum brasiliense located along two different rivers in the state of Minas Gerais, Brazil, using RAPD molecular markers. Eighty-two polymorphic fragments were amplified using 27 primers. The values obtained for Shannon index (I) were 0.513 and 0.530 for the populations located on the margins of the Rio Grande and Rio das Mortes, respectively, demonstrating the high genetic diversity in the studied populations. Nei’s genetic diversity (Hₑ) was 0.341 for the Rio Grande population and 0.357 for the Rio das Mortes population. These results were not significantly different between populations and suggest a large proportion of heterozygote individuals within both populations.

AMOVA showed that 70.42% of the genetic variability is found within populations and 29.58% is found among populations (Bₛ₆₇₅ = 0.2958). The analysis of kinship coefficients detected the existence of family structures in both populations. Average kinship coefficients between neighboring individuals were 0.053 (P < 0.001) in Rio das Mortes and 0.040 (P < 0.001) in Rio Grande. This could be due to restricted pollen and seed dispersal and the history of anthropogenic disturbance in the area. These factors are likely to contribute to the relatedness observed among these genotypes.

1. Introduction

Of all the ecosystems that constitute the Brazilian Semideciduous Forests, riparian forests contribute significantly to the conservation of biodiversity, mainly due to the relationship between ecological corridors and ecosystem functioning [1]. Because of their interconnectedness, riparian forests play a central role in biogeographical and evolutionary shifts as they facilitate seed dispersal [2]. In Brazil, even though riparian forests are protected by law in the Forest Code (Law 4.771 of 1965), these environments suffer the consequences of human activity, mainly resulting from fragmentation and degradation. The reduction of population sizes and the isolation of populations are direct consequences of anthropogenic activities, resulting in inbreeding and genetic drift [3, 4]. Understanding the distribution of genetic variability of tree populations in these areas is fundamental in management programs aimed at conservation and the survival of riparian forest species.

There are a large number of tree species that occur in riparian forests and as such the choice of the target species for genetic and ecological studies is crucial. Among them,
Calophyllum brasiliense Camb. (Clusiaceae), a Neotropical tree commonly known as “guanandi,” stands out as a tree species that occurs preferentially in soils with high humidity saturation and that is sensitive to environmental changes and susceptible to local extinction. The species is widely distributed throughout Central and South America. In Brazil, the wide distribution of C. brasiliense ensures its presence in watersheds of several phytoecological regions.

C. brasiliense presents both male and hermaphroditic flowers. The small “pollen flowers” (<1 cm diameter) are mainly pollinated by halictid bees (Halictidae). Its medium sized, spherical seeds (18–23 mm diameter) are dispersed by gravity and by phylllostomid bats [5]. Bats eat the pulp and discard seeds beneath their feeding roosts [6]. Seeds can also be dispersed secondarily by water, being carried and deposited downstream [7]. Besides the fact that C. brasiliense is a visually attractive species, it is also economically important as it is used for animal feed. It has been widely used in traditional medicine and recent studies show that xanthones from C. brasiliense can potentially decrease the oxidative damage produced by some prooxidants [8]. The species is commonly used in reforestation programs for environmental restoration because its fruit are attractive to fauna and it is also used in tree plantations in areas where the soil is permanently waterlogged [9]. Due to its various uses and importance, C. brasiliense is currently being exploited unsustainably which may eventually alter important genetic parameters in the species’ gene pool.

Information about the structure and genetic diversity of populations as well as understanding other related factors is crucial for the development and adoption of more adequate management strategies to ensure its genetic conservation. Genetic markers, like random amplified polymorphic DNA (RAPD), are valuable in studies assessing genetic variation in plant species when using carefully scored RAPD markers [10,11]. This technique is based on the amplification of a DNA fragment, which occurs when a pair of nucleotides hybridizes in opposite directions from the target sequence; this is achieved through denaturation cycles, primer annealing, and enzyme extension of Taq polymerase [12]. Thus, the objective of this study is to evaluate inter- and intrapopulation genetic variability to provide information that can be used in the maintenance and conservation of C. brasiliense.

2. Materials and Methods

2.1. Sampled Populations. The populations sampled in this study are located along two important rivers in Minas Gerais State: the Rio Grande, in the municipality of Itumirim (21°13′08″S 44°48′14″W), and Rio das Mortes, in Bom Sucesso (21°07′20″S 44°48′22″W), both at an altitude of 850 m. The riparian fragments included in the study are made up of semideciduous forest types. A total of 120 reproductive trees were sampled throughout the study areas. Young leaves were collected and stored at the Tree Species Genetic Conservation Laboratory at the Federal University of Lavras (UFLA), at ~80°C, until genomic DNA extraction. All trees had their geographical coordinates recorded with GPS (Garmin 720) to conduct a spatial analysis of the genetic structure of the studied populations.

2.2. DNA Extraction and PCR-RAPD Reactions. Genomic DNA was extracted from 250 mg of young leaves. In order to equalize the concentrations of DNA to 10 ng·mL⁻¹, genotypes were diluted in a buffer according to the formula \( V = (V_sC_s) - (10V_s)/10 \), where \( V \) is the volume of the TE solution (μL), \( V_s \) is the sample volume, and \( C_s \) is the sample concentration. Subsequently, samples were submitted to PCR-RAPD reactions. The DNA was amplified by the RAPD method and reactions were conducted in a GeneAmp PCR system 9700, using 20 “Operon technologies” primers (California, USA). Each RAPD reaction occurred in a volume of 13 μL containing 32 ng of genomic DNA, 0.2 mM of each dNTP, 4 μM of individual primer, 1 unit of Taq polymerase (Phoneutria), PCR reaction buffer (50 mM of Tris pH 8.3, 20 mM of KCl, 2 mM of MgCl₂, 5 μg/μL of BSA, 0.25% of Ficol 400, 10 mM of tartrazine), and ultrapure water. A PCR 9700 thermocycler was used with the following program: 45 cycles with a denaturation step of 94°C for 15 seconds, annealing of the primer at 42°C for 30 seconds, extension at 72°C for 30 seconds, and a final extension step at 72°C for 5 min.

Amplification products were separated in 1.0% (w/v) agarose gel in TBE buffer, running at 200 V for 2 h, and stained with ethidium bromide; bands were estimated by comparison with a 250-bp ladder (Gibco BRL). Negative controls were consistently used to detect amplification of bands from contaminated DNA. Bands with the same molecular weight and mobility were considered identical. The gels were photographed under ultraviolet light using a BioRad transilluminator.

In order to determine the band frequency in each population, images of the gels were carefully analyzed to construct a phenotype matrix composed of 0 and 1, where 0 corresponds to a band’s absence and 1 its presence. The bands that presented weak coloration and minimal definition were discarded. A locus was considered polymorphic when at least two different bands were detected.

2.3. Data Analysis

2.3.1. Identification of the Optimal Number of Markers. In order to verify that the numbers of polymorphic bands generated by the 20 RAPD primers were sufficient to determine genetic similarity, a bootstrap analysis of 10,000 permutations was conducted using GENES software [13]. The squared sum of deviations in relation to the new samples and the stress value (\( E \)), indicating the fit between the original matrix and the sampled matrix, was calculated. The number of polymorphic bands is considered precise when the stress value is less than 0.05 [14].

2.3.2. Genetic Structure. To assess genetic variability, the software POPGENE (version 1.32) was used to calculate the statistics of genetic variation considering the use of dominant markers and diploid data [15]. Keeping in mind
that the software takes into account the dominance of the alleles, this program assumes that for each estimate of allelic frequencies, the loci are in Hardy-Weinberg equilibrium. The allelic frequencies were estimated from the square root of the number of occurrences of the null genotype (recessive).

Analysis of molecular variance, AMOVA [15], was carried out using the software ARLEQUIN version 3.5.1.2 [16]. The total variance is partitioned into covariance components as they relate to intra- and interindividual and/or population differences [17].

2.3.3. UPGMA. The estimate of genetic similarity between each pair of genotypes was calculated using PC-ORD 4.14 software [17]. The values of genetic distance were calculated based on Jaccard’s coefficient [18]. We used dendrograms of the unweighted pair-group method with arithmetic average (UPGMA) to represent genetic similarities [19]. Group consistency was analyzed by calculating the cophenetic correlation coefficient among matrices based on the arithmetic complement \(d_{ij}^{\text{cp}}\) and cophenetic values, known as Mantel test \((Z)\) [20], using NTSYS-pc software, version 2.1 [21]. This test compares two matrices, element by element, resulting in a correlation value \((r_c)\) that represents the level of relation among the matrices, with the significance of \(r_c\) tested by permutations [22].

2.3.4. Fine-Scale Spatial Genetic Structure (SGS). Kinship coefficients \((F_{ij})\) were calculated from RAPD data according to the dominant marker estimator described by Hardy [23]. This coefficient is the probability that two genes drawn randomly from two individuals are identical by descent; that is, the coefficient estimates the genetic coancestry between individuals. The estimate of \(F_{ij}\) with dominant markers requires the calculation of the inbreeding coefficient, \(F_{IS}\) [23]. We estimated \(F_{IS}\) based on allozymic data [24]. The inbreeding coefficient \((F_{IS})\) was −0.046 for the Rio das Mortes population and 0.311 for the Rio Grande. Spatial genetic patterns were further quantified by the \(SP\) statistic [25] for each population based on kinship analyses, using the program SPAGEDi v. 1.2g [26]. The \(SP\) value was estimated as \(SP = -b_{\log} / (1 - F_{IS})\), where \(b_{\log}\) is the regression slope and \(F_{IS}\) is the mean kinship coefficient between individuals belonging to the first distance class. A jackknife procedure (across loci) was used to estimate standard errors for each distance class and 1,000 randomizations of spatial location were conducted to test for overall spatial structure under the null hypothesis \((b_{\log} = 0)\).

3. Results

3.1. Profile of RAPD Markers. Of the total 27 primers selected, interpretation was possible for only 20 primers, resulting in a total 91 and 92 markers for the Rio Grande and Rio das Mortes populations, respectively (Table 1). Each primer revealed from 1.0 to 7.0 polymorphic loci, producing an average of 5.2 polymorphic loci per primer.

3.2. Evaluating the Optimal Number of Markers. In order to evaluate if the number of amplified polymorphic fragments was able to represent the species’ existent genetic diversity, we conducted a bootstrap resampling analysis. In this analysis, we used the 92 generated fragments from the Rio Grande population (Figure 1) and 91 fragments generated for the Rio das Mortes population (Figure 2). The optimal number of bands was 81 in both populations. With this number of fragments, the correlation’s magnitude value was very close to the maximum for the trees sampled in Rio Grande \((r = 0.97)\) and in Rio das Mortes \((r = 0.98)\).

3.3. Interpopulation Genetic Diversity. Nei’s genetic diversity \((H_e)\) was 0.341 for the Rio Grande population and 0.357 for the Rio das Mortes population, and they were not significantly different (Table 2). The estimated values for the Shannon index \((I)\) suggest high genetic diversity within and among C. brasiliense populations, with a value of 0.513 for Rio Grande, 0.530 for Rio das Mortes, and 0.588 for the entire population.

3.4. Population Genetic Structure. AMOVA showed that 70.42% of the genetic variability of C. brasiliense is found within populations and 29.58% among populations, with an \(\Phi_{ST}\) value of 0.2958 (Table 3).

3.5. Genetic Similarity. We observed that only individuals 22, 23, 37, and 53 sampled from the Rio Grande population were grouped with individuals from the Rio das Mortes population.
analysis detected the existence of family structures in both studied populations (Figure 4). The average kinship coefficient between neighboring individuals (represented by $F_{ij}$) was 0.053 ($P < 0.001$) in the Rio das Mortes population and 0.040 ($P < 0.001$) in the Rio Grande population, indicating that neighboring individuals have a higher level of genetic relatedness than random pairs of individuals. Negative values of $F_{ij}$ occurred with increased distances between individuals for both populations. Overall slopes ($b_{log}$) of two correlograms were significantly different from the null hypothesis of no spatial genetic structure ($b_{log} = 0$): $b_{log} = -0.0223 \pm 0.0036$ (s.e.) for Rio das Mortes and $b_{log} = -0.0228 \pm 0.0040$ (s.e.) for Rio Grande. $Sp$ values revealed strong patterns of SGS for the total sample and they were similar for both populations ($Sp = 0.0236$ for Rio das Mortes and $Sp = 0.0238$ for Rio Grande).

4. Discussion

4.1. Genetic Diversity. The number of RAPD markers used to evaluate genetic variability in plants is highly variable and not dependent on the species' level of domestication. Pigato and Lopes [27] when evaluating the genetic variability of four generations of Eucalyptus urophylla identified 86 amplified fragments. Zimback et al. [28] detected 72 polymorphic fragments when analyzing the genetic diversity of Trichilia pallida, a Meliaceae climax species. Wang et al. [29], investigating the genetic structure of Neolitsea sericea, found 50 polymorphic fragments. Torezan et al. [30], studying Aspidosperma polyneuron using RAPD markers, encountered 98 polymorphic fragments, while Phong et al. [31], in studying the genetic variability of Dalbergia oliveri, a leguminous tree of the Fabaceae family, found 24 polymorphic fragments when using 29 RAPD primers. Considering the number of polymorphic fragments described in the many published studies on natural tree species populations, the number of fragments obtained in this study of 92 and 91 (Table 2) can be considered satisfactory, enabling an accurate evaluation of genetic diversity for this species.

The sufficiency of the number of polymorphic bands for genetic similarity, calculated by resampling, showed low stress values ($E$), 0.045 and 0.040 at Rio Grande and Rio das Mortes, respectively (Figures 1 and 2). According to Kruskal [14], stress values lower than 0.05 are an excellent indicator of precision. Therefore, the values obtained in our study of 0.045 for Rio Grande and 0.040 for Rio das Mortes, respectively, validate the results obtained for number of fragments.

The interpopulation genetic divergence indexes obtained in our analysis demonstrate that the values of $He$ (0.341 and 0.357) were similar to the values obtained in other studies of natural tree species populations. Torezan et al. [30], studying six natural populations of Aspidosperma polyneuron, obtained an $He$ of 0.299 for adults and seedlings. Estopa et al. [32] found $He$ values ranging from 0.299 to 0.333 in Eremanthus elytropappus, while Zimback et al. [28] found a variation of 0.27 to 0.33 for $He$ when studying Trichilia pallida. The high genetic diversity within natural populations of C. brasiliense is consistent with other studies of tree species when comparing the values of Shannon's index ($I$), where similar methodologies were used. Lacerda et al. [33] observed

![Figure 1](image1.jpg)  
**Figure 1:** Resampling analysis (bootstrap) and the correlation ($r$) with number of fragments from the Rio Grande population. SQd is the squared sum of the deviations in relation to the new samples; $E$ is the stress value.

![Figure 2](image2.jpg)  
**Figure 2:** Resampling analysis (bootstrap) and the correlation ($r$) with number of fragments from the Rio das Mortes population. SQd is the squared sum of the deviations in relation to the new samples; $E$ is the stress value.

<table>
<thead>
<tr>
<th>Population</th>
<th>$n$</th>
<th>Genetic diversity ($He$)</th>
<th>Shannon index ($I$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rio Grande</td>
<td>60</td>
<td>0.341</td>
<td>0.513</td>
</tr>
<tr>
<td>Rio das Mortes</td>
<td>60</td>
<td>0.357</td>
<td>0.530</td>
</tr>
<tr>
<td>Overall</td>
<td>120</td>
<td>0.403</td>
<td>0.588</td>
</tr>
</tbody>
</table>

$n$, sample size.

(Figure 3). This result underscores the level of divergence among trees sampled in both studied populations. The results confirmed that the polymorphic bands obtained in these two populations, using 20 RAPD primers, were able to assess the genetic similarity of *C. brasiliense* trees. We can therefore assume that the grouping of the individuals based on the cophenetic correlation is highly reliable (Figure 3).

3.6. Fine-Scale Spatial Genetic Structure. Kinship coefficient analysis detected the existence of family structures in both
Table 3: Analysis of molecular variance (AMOVA) for 120 individuals sampled from two populations of *Calophyllum brasiliense* using RAPD molecular markers.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>SSD</th>
<th>Variance component</th>
<th>% Total</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among populations</td>
<td>1</td>
<td>384.866</td>
<td>6.16959</td>
<td>29.58</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within populations</td>
<td>118</td>
<td>1733.533</td>
<td>14.69096</td>
<td>70.42</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

The data show the degrees of freedom (d.f.), sum of squared deviation (SSD), variance component estimates, the percentage of total variance contributed to each component (% total), and the probability (P value).

![Genetic identity](image)

Figure 3: Hierarchical cluster analysis of *Calophyllum brasiliense* genotypes in populations from Rio Grande (RG) and Rio das Mortes (RM) using UPGMA, based on data obtained with RAPD. RC: cophenetic correlation.

$I$ values ranging from 0.301 to 0.367 for *Plathymenia reticulata*. Torezan et al. [30] obtained an $I$ value equal to 0.410 for adult individuals of *Aspidosperma polyneuron*, and Zang et al. [34] found diversity values of 0.4382 using RADP markers in natural populations of *Larix gmelinii*. Comparatively, these studies suggest that there is strong evidence that *C. brasiliense* presents high levels of genetic variability with $H_e$ values of 0.341 for the Rio Grande population and 0.357 for the Rio das Mortes population. This can be related to factors such as insect-pollinated flowers and a predominantly outcrossed
Confidence intervals around each jackknife procedure across all loci. In this study, the observation of significantly positive kinship coefficients over a short distance indicates a clear family structure, that is, genetic cluster, within the two populations. A similar pattern between populations was found based on the correlograms. The significant SGS found in both populations suggests restricted seed dispersal of *C. brasiliense*, a primary factor responsible for the observed SGS of this species [41]. The significant genetic clustering at short distances in both populations likely reflects seedling recruitment around the maternal plant [42, 43]. According to Wright [44], if seed dispersal is localized at the scale of investigation, it will result in spatial clustering of genetically related individuals and the development of significant SGS.

The *Sp* statistic is a robust method to compare fine-scale genetic patterns among plant species, independent of sampling strategies, life histories, and population densities [25]. In comparison to mean values for other species, the estimated *Sp* values in our study (average *Sp* = 0.024) are within the range of those reported in predominantly mixed mating (*Sp* = 0.0372 ± 0.0367), pollen animal-dispersed (*Sp* = 0.0171 ± 0.0142), and seed gravity-dispersed (*Sp* = 0.0281 ± 0.0166) species [25].

### 4.3. Implications for Genetic Conservation

Appropriate conservation and management strategies are necessary in order to ensure the long-term genetic variability of *C. brasiliense*. The distribution of genetic variation across the landscape is a prime factor to consider in the conservation and management of natural plant populations [45]. Our results show the existence of family structures and a detailed depiction of fine-scale genetic variation patterns. This spatial information can be used to determine future sampling strategies that ensure appropriate distances between sampled genotypes.

For example, sampling methods for seed banks (ex situ conservation) and definitions of the size of an area for in situ conservation can benefit greatly from this information. Based on the correlogram of average *F*$_{ij}$ in local *C. brasiliense* populations, individuals located within an area of 535 m (at Rio das Mortes) and 206 m (at Rio Grande) are likely to be genetically similar. Therefore, to optimize sampling design and limit redundant genotypes, only individuals located at greater distances than 535 m or 206 m should be collected. Furthermore, samples should be collected from both

---

**Figure 4**: Correlograms of average coancestry coefficients (*F*$_{ij}$) per distance class in two *Calophylhum brasiliense* populations. Confidence intervals around each *F*$_{ij}$-value were obtained using a jackknife procedure across all loci.
populations since our results show that they are significantly different.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for financial support. The authors also thank Dr. Evelyn Nimmo for editing the English of the paper.

References


Submit your manuscripts at http://www.hindawi.com