Molecular characterization as strategy for ex situ conservation of *Anadenanthera peregrina* (L.) Speg.

Caracterização molecular como estratégia para conservação ex situ de *Anadenanthera peregrina* (L.) Speg.

Maressa Albuquerque Cortelete1, Adelson Lemes da Silva Júnior1, Maria Lucia Silva Pereira1, Fábio Demolinari de Miranda1, Marcos Vinicius Winckler Caldeira2

1Universidade Federal do Espírito Santo – UFES, Alegre, ES, Brasil
2Universidade Federal do Espírito Santo – UFES, Jerônimo Monteiro, ES, Brasil


Abstract

The diversity and genetic structure were used as strategies for the ex situ conservation of *Anadenanthera peregrina*, a tree with great ecological and economic potential. From a planted population, 166 individuals were evaluated using molecular markers Simple Sequence Repeats (SSR). All six SSR loci were polymorphic, of which 42 alleles were detected. The PIC values ranged from 0.61 to 0.84, classifying loci as very informative. Values of *H*<sub>e</sub>, *H*<sub>o</sub> and *R*<sub>a</sub> indicated a high degree of heterozygosity and genetic diversity, while the relationship between *H*<sub>e</sub> and *H*<sub>o</sub> generated negative values for the fixation index (*F*<sub>IS</sub>), indicating escape of inbreeding in the population. The analyzes performed by the UPGMA method and the Bayesian approach structured the population into six and two genetic groups, respectively. The high genetic diversity and population structure revealed the importance of ex situ conservation of the evaluated genotypes, allowing future use of the population as a seed orchard.

Keywords: Genetic diversity; Population structure; SSR; Seed orchard.

1. INTRODUCTION

Ecosystems and populations of plant species are known for their ability to survive and reproduce in the face of disturbances. As a preponderant factor for this maintenance and
adaptation, the term genetic diversity has been highlighted today in effective decision making, as it reflects the presence of different alleles in the gene pool and, therefore, different genotypes within the populations (Ellegren & Galtier, 2016; Mukhopadhyay & Bhattacharjee, 2016; Torres-Florez et al., 2018).

The implementation of this mechanism proceeds from the serious anthropic damages generated to the natural habitats at the global level, causing the reduction and fragmentation of the vegetal cover, allied consequently the genetic erosion by the increase of the inbreeding, limitations to the gene flow and stochastic events (Aavik & Helm, 2018).

Knowing the causes and the resulting damages, global policies have demanded actions that help in the conservation, restoration and recovery of ecosystems and their species, combining issues such as expansion of the vegetation cover and demand for wood (Freitas et al., 2015; Potter et al., 2017). As a viable strategy to meet these purposes, a seed orchard is an ex situ conservation method, characterized as a planting planned, delineated, managed, established with superior matrices and isolated from external propagules (Brasil, 2004), capable of providing seeds in greater quantity and quality, maximizing the adaptive and evolutionary capacity of future populations, by the use of seeds with a broad genetic base (Broadhurst et al., 2017).

For this and other purposes, in recent decades, there has been a significant increase worldwide in the use of molecular marker methods to evaluate variability and genetic diversity in plant species (Muchugi et al., 2008). Among the different types of molecular markers, we highlight the Simple Sequence Repeats (SSR), which have wide acceptance in research with forest species (Resende et al., 2011; Mangaravite et al., 2016; Mora et al., 2017), due to the fact that it is an easy technique in terms of reproducibility, has a multi-allelic character, codominant and mainly because they are abundant in the genome (Turchetto-Zolet et al., 2017). Studies involving the use of SSR markers in forest species, among the various purposes, are important in the search for knowledge of the genetic composition of the species, and how it is organized (structured) in their populations, being fundamental for management and conservation actions, since intra and interspecific genetic diversity provides the adaptive / evolutionary potential of a species (Galetti Junior et al., 2008).

Anadenanthera peregrina (L.) Speg., popularly known as Angico-curtidor or Red-Angico, belongs to the Fabaceae family (Carvalho, 2003). It is a species capable of developing in the most diverse environments, found in Brazil in biomes such as Caatinga and Atlantic Forest, in different forest formations such as montane, sub-montane, ombrophilous and seasonal forests, also encompassing other countries such as Peru, Bolivia, Uruguay and Argentina (Lorenzi, 2002). Characterized as an allogamous species, it presents cross pollination by bees, favored by the high degree of genetic self-incompatibility. The ripening pattern of fruits and seeds occurs during or at the end of the dry season, with anemochoric seed dispersal (Costa et al., 2003).

Regarding the economic and ecological potential, the species presents moderate to rapid growth, reaching up to 20 m in height and 60 cm in diameter, with robust and very durable core and, average productivity of about 25.55 m³ha⁻¹ year (Araújo et al., 2006; Souza et al., 2012). Due to the durability of the wood, it can be used in construction, furniture manufacturing and energy supply through firewood and charcoal (Lorenzi, 2008). It also has medicinal use because it has tannin, an antioxidant substance, also used in leather tanning. Ecologically, the use of the species in the restoration and recovery of degraded environments is highly efficient because it has abundance of seeds and the ability to colonize the most diverse environments, aiding the nutritional condition since it is a nitrogen fixer (Araújo et al., 2006).

Despite the potential of the species, there are no data from individuals genetically selected for the composition of a germplasm bank. Knowledge about the variability and existing genetic structure will bring important information in decision making for ex situ conservation and production of seedlings with confirmed genetic variability. In this sense, we aim to characterize the diversity and genetic structure of a planted population of A. peregrina.
in the southern region of the state of Espírito Santo, Brazil, using molecular markers Simple Sequence Repeats (SSR).

2. MATERIAL AND METHODS

2.1. Sampling and collection of plant material

The study area is characterized as a planted population of *A. peregrina*, located at the Instituto Federal de Educação Ciência e Tecnologia do Espírito Santo, Campus of Alegre, ES, on highway ES 482, km 47, district of Rive, Alegre, ES, Brazil (20°45'46.8 South latitude and 041°27'38.4 West longitude). The planting was carried out in June 2011, through seedlings donated by the Reserva Vale in Linhares-ES, which is characterized as a well-preserved Atlantic Forest fragment, with intermediate vegetation between the semi-deciduous seasonal and dense ombrophilous typologies (Kierulff et al., 2014), however, there is no information available on the number of matrices used in seed collections or their genetic basis.

Young leaves with good phytosanitary appearance were collected from 166 individuals georeferenced and randomly selected, covering the whole area of the plantation (Figure 1).

![Figure 1: Location map of the study area and sampling of the selection and collection of *Anadenanthera peregrina* individuals.](image)

2.2. Extraction and quantification of genomic DNA

The genomic DNA was obtained by the protocol of Doyle & Doyle (1990), adjusting the concentrations of 1% polyvinylpyrrolidone (PVP) and 2% cetyl trimethylammonium bromide (CTAB), proposed by the Instituto Agronômico de Campinas (IAC). The determination of the concentration and purity of the samples were performed by spectrophotometry using the NanoDrop 2000 apparatus (Thermo Fisher Scientific, Waltham, MA, USA), and the quality definition was based on the A260 / A280 ratio, where ratios lower than 1.8 indicate contamination by proteins and higher than 2.0 indicates contamination with chloroform or phenol (Barbosa, 1998).

2.3. Microsatellite markers (SSR) and PCR amplification conditions

A set of six microsatellite markers developed for *Anadenanthera colubrina* with heterologous amplification for *Anadenanthera peregrina* (Feres et al., 2012), were used for the genetic characterization of the individuals. The amplification reactions were carried out at the Laboratório de Bioquímica e Biologia Molecular, Universidade Federal do Espírito Santo, UFES (Brazil). The final volumes of the reactions were adjusted to 15 μl: 1x amplification buffer (10 mM Tris-HCl pH 8.5 and 50 mM KCl); 1.5 mM MgCl2; 1.0 mM dNTPs (dGTP, dATP, dCTP, dTTP); 0.3μM primer F; 0.3μM primer R; 1 unit of Taq DNA polymerase and 10 ng of DNA.

The amplification conditions followed the one proposed by Feres et al. (2012), with initial denaturation at 96 ºC for 4 min, followed by 30 denaturation cycles (94 ºC for 40 sec), annealing
specific for each primer pair (ranging from 52-58 °C for 1 min) (Table 1) and alloying (72 °C for 1 min), final extension of 72 °C for 7 min.

Table 1. Sequences of the primers and their respective temperature conditions used for *Anadenanthera peregrina*. Source: Feres et al. (2012).

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5'-3') F</th>
<th>Sequence (5'-3') R</th>
<th>Ta °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acol 09</td>
<td>F: CCAGGGTCCTCTCAGATTG</td>
<td>R: ATGCTCCTCTCAACACACC</td>
<td>58</td>
</tr>
<tr>
<td>Acol 10</td>
<td>F: CGATGAGCTGAGCTAATATGC</td>
<td>R: GCATGAAAGAAAGCTAATCCTC</td>
<td>54</td>
</tr>
<tr>
<td>Acol 13</td>
<td>F: CATATGCAATCAGGTTAGAC</td>
<td>R: CTCATTTGGAAGCTAAGC</td>
<td>58</td>
</tr>
<tr>
<td>Acol 15</td>
<td>F: GTCAGGTATTCAATCTTCTACGT</td>
<td>R: CACCATACCTCCACTCAGT</td>
<td>52</td>
</tr>
<tr>
<td>Acol 18</td>
<td>F: GACTGAAACCATCTTCTTTATC</td>
<td>R: TGTTGCGTATGTGGGTG</td>
<td>52</td>
</tr>
</tbody>
</table>

Ta °C = Annealing temperature.

The amplification products were separated by 10% polyacrylamide gel electrophoresis, at a voltage of 100 V, for approximately 4 h. The gels were then stained by immersion in ethidium bromide solution (0.25 μg / ml) for 20 minutes and photo-documented under UV light, allowing visualization of the bands and molecular weight determination by Ladder 50pb marker.

### 2.4. Data analysis

The molecular data recording was performed from polymorphisms of the PCR products; then an encoding matrix of the individual genotypes was generated, considering the size of the alleles in each locus.

The number of alleles (A), allelic richness (R_a), expected heterozygosity (H_e), observed heterozygosity (H_o), Hardy-Weinberg equilibria deviations, and the fixation index (F_{IS}) estimated according to Weir (1996) of the FSTAT 2.9.3.2 software (Goudet, 2002) were calculated. By means of the GENES program (Cruz, 2013), genetic dissimilarity estimates (dii') were performed according to the arithmetic complement of the weighted index and organized into matrices for the UPGMA cluster analysis (Unweighted pair-groups method using arithmetic averages), where the number of groups was obtained from the cut-off point proposed by Mojena (1977), which describes the formula as \( P_c = m + kp \); where \( m \) = the mean distance values of the melting levels corresponding to the stages; \( k = 1.25; \) \( dp \) = standard deviation. Subsequently, the output matrix with dissimilarity values were exported to program R (R Development Core Team, 2018) to obtain the dendrogram.

The genetic structuring was evaluated by the Bayesian approach Monte Carlo via Markov Chains (MCMC) using STRUCTURE 2.3 software (Pritchard et al., 2000), with 20 runs for each K value, establishing the number of groups (K) ranging from K = 1 to K = 4, with 7,500 interactions and an initial burn-in of 2,500 interactions. The number of genetic groups was determined by the ad hoc \( \Delta K \) method proposed by Evanno et al. (2005) using STRUCTURE HARVESTER software (Earl & Vonholdt, 2012).
3. RESULTS AND DISCUSSION

The results obtained from the microsatellite markers were used to generate base matrices for the analyses of the variability and genetic diversity characterization. All loci presented polymorphism and the number of alleles ranged from four to nine, with an average of seven alleles. Of the 42 alleles detected, the loci with the highest number of alleles were Acol 10 and Acol 19, with nine alleles each, while the Acol 18 locus had the lowest polymorphism, with only four alleles observed (Table 2).

Table 2. Results of the genetic parameters of the microsatellite markers in the population of *Anadenanthera peregrina*. A = number of alleles, Ra = Allelic Richness, He = Expected Heterozygosity, Ho = Observed heterozygosity, FIS = Fixing Index and PIC = Polymorphic Information Content.

<table>
<thead>
<tr>
<th>Loci</th>
<th>A</th>
<th>Ra</th>
<th>He</th>
<th>Ho</th>
<th>FIS</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acol 09</td>
<td>6</td>
<td>6</td>
<td>0.64</td>
<td>0.68</td>
<td>-0.06</td>
<td>0.61</td>
</tr>
<tr>
<td>Acol 10</td>
<td>9</td>
<td>11</td>
<td>0.85</td>
<td>0.93</td>
<td>-0.08</td>
<td>0.84</td>
</tr>
<tr>
<td>Acol 13</td>
<td>6</td>
<td>6</td>
<td>0.77</td>
<td>0.79</td>
<td>-0.01</td>
<td>0.74</td>
</tr>
<tr>
<td>Acol 15</td>
<td>8</td>
<td>8.85</td>
<td>0.74</td>
<td>0.51</td>
<td>0.24</td>
<td>0.70</td>
</tr>
<tr>
<td>Acol 18</td>
<td>4</td>
<td>5</td>
<td>0.71</td>
<td>0.65</td>
<td>0.07</td>
<td>0.65</td>
</tr>
<tr>
<td>Acol 19</td>
<td>9</td>
<td>9</td>
<td>0.83</td>
<td>0.90</td>
<td>-0.08</td>
<td>0.81</td>
</tr>
<tr>
<td>Average</td>
<td>7</td>
<td>7.64</td>
<td>0.76</td>
<td>0.74</td>
<td>0.001</td>
<td>0.72</td>
</tr>
</tbody>
</table>

* All Fixation Indices were non-significant (5% probability).

Compared with the study by Feres et al. (2012), who evaluated the genetic characterization in two natural populations of *Anadenanthera colubrina*, the number of alleles present in Acol 15 (14 alleles), Acol 18 (7 alleles) and Acol 19 (10 alleles) were higher than those found in the present study, corroborating the values for the Acol 09, Acol 10 and Acol 13 loci.

The polymorphic information content (PIC) indicating marker quality for genetic studies is dependent only on the number of alleles and their frequency for each marker (Guo & Elston, 1999). PIC values above 0.50 are considered to be very informative, values between 0.25 and 0.50 are considered informative and values below 0.25 are considered to be less informative (Botstein et al., 1980). According to this classification we can say that all the loci used in this study were very informative.

The high values of He and Ho indicate a high degree of heterozygosity and diversity in the *A. peregrina* population when compared to other studies involving tree species (Guidugli et al., 2010; Feres et al., 2012; Schwarcz et al., 2014). The relationship between the observed and expected heterozygosity values generated negative values for the fixation index (FIS) in most loci, which indicates escape of inbreeding in the population, except for the Acol 15 and Acol 18 loci, which presented positive indices, in which, according to Wright (1965), positive values represent excess homozygotes. In addition, high Ra (0.76) as well as He is also indicative of high genetic diversity. The parameters of genetic diversity calculated in this study point to the ex situ conservation of the species and the possible use of the population as a seed orchard.

All loci of the present study did not differ significantly from the Hardy-Weinberg Equilibrium, thus indicating an absence of inbreeding, suggesting that Costa et al. (2003) found in their study that *A. peregrina* features genetic self-incompatibility, which is the inability of the fertile plant to form seeds when fertilized by its own pollen, favoring cross-pollination and maintaining genetic variability. This result is confirmed by means of the non-significant index of fixation (FIS) for the studied population, from which we can infer that the individuals evaluated were obtained from seeds originated in an unstructured population where the crosses occurred in a random manner.

Estimates of genetic dissimilarity (dii’) among the sampled individuals generated 89 access pairs with 100% dissimilarity, while individuals 82 and 83 had the lowest value with 7.14%. This similarity between individuals 82 and 83 may be associated with seed collection for seedling production in the same matrix tree. However, it was observed that the majority of the individuals presented a high degree of dissimilarity, indicating that the seeds for the production of the seedlings came from different mother trees.
The genetic dissimilarity dendrogram performed in the R program, based on the genetic distances generated in the Genes program for 166 of *Anadenanthera peregrina* individuals, obtained by the UPGMA method, revealed six distinct groups of the cutoff point defined by m + kdp = 0.7446 (Figure 2). The analysis of the dendrogram allows to define the genetic proximity between the individuals, allowing the selection of more and less discrepant in the near future, guiding to obtain of seeds lots with greater genetic variability.

*Figure 2.* Dendrogram of genetic dissimilarity performed in the R program, based on the genetic distances generated in the Genes program for 166 individuals of *Anadenanthera peregrina*.

In relation to the Bayesian approach, associated to the ΔK method described by Evanno et al. (2005), it was found that the most probable number of K groupings is two (K = 2) (Figure 3).

*Figure 3.* Genetic structuring of the *Anadenanthera peregrina* population using the Bayesian approach performed in the Struture program, indicating a better value of ΔK, with K = 2.
The number of k clusters reveals how many genetic groups partition the analyzed samples, so the genetic structuring performed by the Bayesian approach allows us to obtain detailed information such as the possibility of verifying heterogeneity; which helps in the maintenance and adaptation of trees against adverse characteristics. In a study with *Cedrela fissilis*, samples from 18 populations collected from almost all of Brazil were evaluated for genetic structuring by the Bayesian approach which also indicated the most probable number of K groupings equal to two (K = 2) (Mangaravite et al., 2016). In comparison, it can be interpreted that the species under study has a good genetic structure and that the trees evaluated to compose the seed orchard are capable of generating genetically structured seeds.

Despite the disparate values between the UPGMA method and the Bayesian approach, there is normality in the data for evaluating different statistics, and we can add that strategies involving population structuring are performed by more than one tool which can be used together.

4. CONCLUSION

We highlight the efficiency of molecular markers in evaluating existing polymorphisms in species in general, and the use of genetic diversity as a basis for maintenance and adaptation of future populations.

The high diversity and structure in at least two genetic groups revealed the importance of conserving the evaluated genotypes ex situ, allowing the future use of the population as a seed orchard. However, although future matrices will be able to produce genetically dissimilar seeds, further studies and evaluations are required for seed orchard establishment, such as progeny testing of selected matrices, thinning of unselected or similar individuals and the inclusion of new provenances which may increase this variability.

5. ACKNOWLEDGMENTS

The authors are grateful to the Fundação de Amparo à Pesquisa do Espírito Santo (FAPES) and the Universidade Federal do Espírito Santo (UFES) for the award of scientific grants and research support. The authors also grateful the Núcleo de Pesquisa Científica e Tecnológica em Meio Ambiente, Silvicultura e Ecologia (Nupemase) for the research support.

6. REFERENCES


Molecular characterization as strategy for ex situ conservation of Anadenanthera peregrina (L.) Speg.
Molecular characterization as strategy for ex situ conservation of *Anadenanthera peregrina* (L.) Speg.


Authors' contributions: MAC: Conceptualization, data curation, formal analysis, investigation methodology; ALSJ: Data curation, formal analysis, methodology, investigation, writing – original draft, writing – review & editing; MLSP: Formal analysis, investigation, methodology; FDM: Conceptualization, methodology, supervision, writing – review & editing; MVWC: Funding acquisition, project administration, resources, supervision, writing – review & editing.