



Genetic structure of natural populations of *Theobroma* in the Juruena National Park, Mato Grosso State, Brazil

B.M. Silva¹, A.A.B. Rossi¹, J.F.E. Dardengo¹, C.R. Silva², I.V. Silva³, M.L. Silva⁴ and C.J. Silva⁵

¹Laboratório de Genética Vegetal e Biologia Molecular, Rede Bionorte, Faculdade de Ciências Biológicas e Agrárias, Universidade do Estado de Mato Grosso, Campus de Alta Floresta, MT, Brasil

²Departamento de Agronomia, Universidade Federal Rural da Amazônia, Campus de Parauapebas, Parauapebas, PA, Brasil

³Laboratório de Biologia Vegetal, Faculdade de Ciências Biológicas e Agrárias, Universidade do Estado de Mato Grosso, Campus de Alta Floresta, MT, Brasil

⁴Laboratório de Genética/Cultura de Tecido Vegetal, Centro de Pesquisa, Estudos e Desenvolvimento Agro-Ambientais, Faculdade de Ciências Biológicas e Agrárias, Universidade do Estado de Mato Grosso, Campus de Tangará da Serra, MT, Brasil

⁵Centro de Pesquisa em Liminologia, Biodiversidade e Etnobiologia do Pantanal, Rede Bionorte, Universidade do Estado de Mato Grosso, Campus de Cáceres, MT, Brasil

Corresponding author: A.A.B. Rossi
E-mail: anabanrossi@gmail.com

Genet. Mol. Res. 14 (3): 10365-10375 (2015)

Received January 31, 2014

Accepted May 8, 2015

Published September 1, 2015

DOI <http://dx.doi.org/10.4238/2015.September.1.3>

ABSTRACT. Several species within the genus *Theobroma* have particularly high economic value, including *T. cacao* and *T. grandiflorum*. Other species in this genus, such as *T. speciosum* and *T. subincanum*, have potential value for use in the conservation of

genetic diversity in breeding programs. These latter species could also be domesticated or improved to produce commercial products. Using 13 simple sequence repeat loci, the population structure and genetic diversity of *T. speciosum* and *T. subincanum* natural populations in the Juruena National Park, Mato Grosso State, Brazil, was studied. We sampled all individuals of each species (N = 25) present inside a designated research area established by the Program for Research on Biodiversity. The average number of alleles per locus was 5 for *T. speciosum* and 6.69 for *T. subincanum*, with average PIC values above 0.5 in both species. All evaluated individuals varied genetically. Seeds from the individuals analyzed will be useful for the development of germplasm banks and for establishment of breeding programs.

Key words: Genetic variability; *Theobroma speciosum*; Cacaú; *Theobroma subincanum*; Cupuí

INTRODUCTION

The Amazon forest contains the largest natural reserve of plant diversity on the planet, and each forest environment contains a rich and varied floristic composition and structure that is often unique (Oliveira and Amaral, 2004). The Amazon is an extraordinary resource for the Brazilian population and offers incalculable benefits to the world (Barreto et al., 2005).

The sustainable use of resources in the Amazon is a significant challenge presently and also for future generations. The lack of biological analysis and understanding of the majority of floral species in the Amazon is a particular problem because many species are being lost as the result of land conversion for agricultural production and uncontrolled natural resource exploitation in the region (Bekessy et al., 2002). The majority of tropical forest in the Amazon has been exploited unsustainably and this is evident from the loss of vegetation cover and, consequently, the loss of species diversity, without even a basic understanding of the importance and richness of many of these species (Souza et al., 2006).

One of the primary functions of conservation areas is the protection of biodiversity. In the early 1990s, the determination to preserve biodiversity resulted in the creation of the Convention on Biological Diversity (CBD), under which the role of protected areas for *in situ* conservation of biodiversity was reinforced and expanded. The CBD also established a set of activities to be undertaken to improve and better represent these protected areas (CDB, 1992).

The Juruena National Park (PnJu) is a fully protected conservation unit that is managed by the Chico Mendes Biodiversity Institute (ICMBio). The Park is located in the north of Mato Grosso State and the southeast of Amazonas State, between the municipalities of Apuí and Maués (AM) and Apicás, Cotriguaçu and Nova Bandeirantes (MT). Its establishment in 2006 represented one of the largest planning and management initiatives for environmental conservation in Brazil (PNJU, 2013).

The PnJU includes members of the Malvaceae sensu lato family, which contains economically important fruit tree species, such as genus *Theobroma* L. Members of this family typically grow in Neotropical areas and are distributed throughout the humid forests of the Western Hemisphere, extending from Mexico to the limits of the Amazon Forest. The center of origin

of the genus is possibly in the State of Amazonas and surrounding regions (Cuatrecasas, 1964).

The *Theobroma* genus includes 22 species, all found in the tropical regions of the Americas (Cuatrecasas, 1964). Of these, eight species are found in the Brazilian Amazon, including trees locally called cacau (*T. cacao* L.), cupuaçu [*T. grandiflorum* (Willd.) Schum.], cacauí (*T. speciosum* Willd.), and cupuí (*T. subincanum* Mart.). Cacau and cupuaçu trees are cultivated on a large scale; although the fruits of both cacauí and cupuí are harvested by native populations, both species suffer from the consequences of habitat fragmentation.

Due to the intense forest fragmentation in the Amazon region, preservation of genetic diversity has become the principal objective of most conservation programs (Bekessy et al., 2002). The main implications of fragmentation for a species are reduction in total number of individuals, reduction in average population size, restriction of populations into small fragmented groups, and spatial isolation of remnant populations (Young and Boyle, 2000).

In the Amazon, the need for strategies for biodiversity conservation, along with global pressure to implement less aggressive environmental resource exploitation, has encouraged the establishment of research priorities on genetic resources. In this context, information obtained through collection, evaluation and characterization of germplasm has received much attention. One outcome of this interest in germplasm research is the increased potential for using this information on genetic variability in technological development programs for native species (Sousa et al., 2009).

According to Estopa et al. (2006), genetic studies of natural populations should seek to evaluate and quantify how genetic variability is distributed over time and space. Information on variability within and between natural populations can provide a better understanding of how selection functions in terms of adaptability; the general presumption is that the greater the genetic variability within a population, the greater the likelihood of adaptation and perpetuation of the species.

Genetic diversity is important for species within an ecosystem as it enables the species to adapt and also provides a foundation for genetic improvement. This diversity is responsible for the variation in productivity and reproduction among individuals of the same species (Lowe et al., 2005).

Molecular markers are widely used in genetic analyses because of the relative ease of amplifying specific regions of the genome in order to characterize genetic variability and genetic structure in populations (Sena et al., 2007).

Microsatellite markers, such as simple sequence repeats (SSRs), are highly informative and have been used to study natural populations because they are multi-allelic, highly polymorphic, codominant, abundant, evenly distributed throughout the genome, and easily obtained using PCR (Gaiotto et al., 2003).

The objective of this study was to characterize population structure and genetic diversity of *T. speciosum* and *T. subincanum* in the PnJu, Mato Grosso, Brazil. This information is essential for developing strategies for conserving and improving these species.

MATERIAL AND METHODS

Study area

The study was conducted at the PnJu, Mato Grosso State, Brazil (Figure 1), in two plots established by the PPBio Project (Biodiversity Research Program). Samples were col-

lected from two species of *Theobroma* that occur most frequently in the plots: *T. speciosum* and *T. subincanum*. A total of 25 *T. speciosum* trees and 25 *T. subincanum* trees were identified, sampled, and georeferenced in the study area.

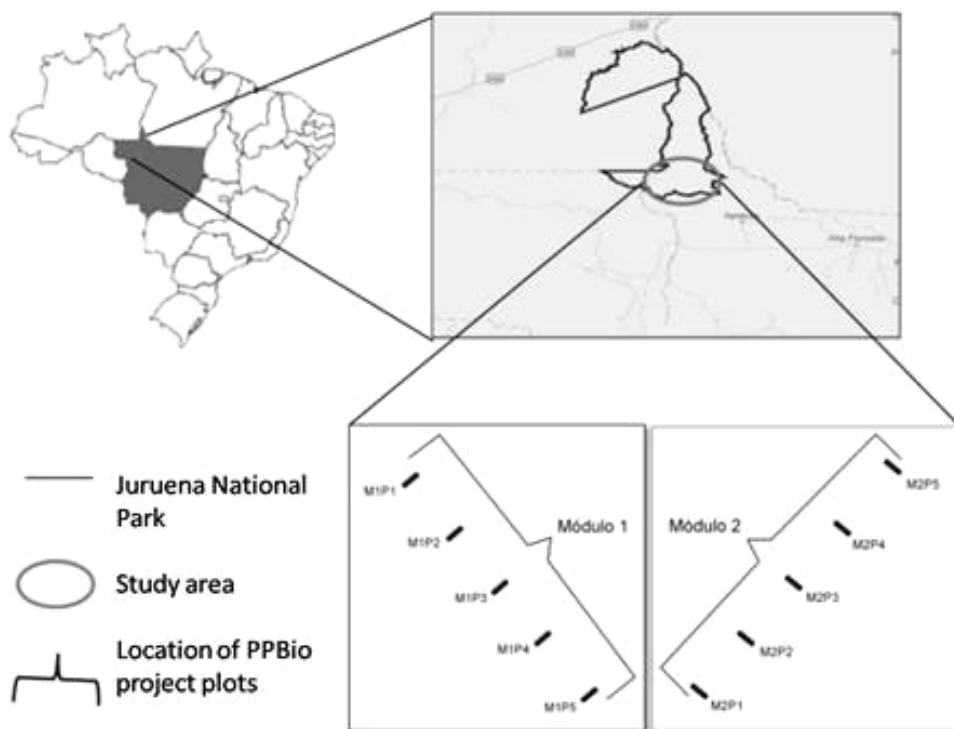


Figure 1. Location of the study area within the Juruena National Park and the plots established by the PPBio Project (Biodiversity Research Program).

Plant sampling

Leaf tissue samples were collected for DNA extraction from all adults and juveniles found within the study plots. The samples collected were stored in containers with silica gel in the field. The samples were later stored at -20°C in the Genetics Laboratory of Campus Universitário de Alta Floresta, Mato Grosso.

Total DNA extraction

Total genomic DNA was extracted using the cetyltrimethylammonium bromide method as described by Doyle and Doyle (1990), with modifications following Lemes et al. (2010). For quantification, DNA was applied to an agarose gel 1% (w/v), stained with ethidium bromide. Bands were compared to standard DNA (lambda phage) of known concentrations. Gels were analyzed with a UV transilluminator and photographed. After quantification, the

extracted DNA samples were diluted in autoclaved distilled water.

Selection of primers and amplification by PCR

Twenty-three microsatellite loci (SSR), previously isolated and characterized by Lanaud et al. (1999), were tested in an initial PCR amplification using one *T. speciosum* and one *T. subincanum* plant chosen after DNA quantification. Of the 23 loci tested, 13 were selected for analysis of genetic diversity of the species.

The amplification protocol followed that described by Lanaud et al. (1999), with some modification: one initial cycle of denaturation at 94°C for 4 min, followed by 32 cycles of 94°C for 30 s, 46° or 51°C (depending on the primer used) for 1 min, 72°C for 1 min, and one final extension cycle at 72°C for 5 min. The amplification products were separated by electrophoresis on a 2% (w/v) agarose gel in 1X TBE running buffer, at a constant 85 V for approximately 5 h. The gel was stained with ethidium bromide (0.2 mg/mL) and subsequently photographed under ultraviolet light using the Transilluminator UVB LTB-21x26 (Loccus Biotecnologia) and a Sony digital camera.

Data analysis

Fragments (amplification products) were analyzed using the Gel QuantExpress program to develop a matrix based on the fragment size of the bands.

The allelic frequency, genetic diversity, observed heterozygosity, expected heterozygosity, and the polymorphism information content (PIC) were assessed using the program Power Marker (Liu and Muse, 2005). Nei's (1973) matrix of genetic distance between individuals was estimated using the same program. This matrix was imported by MEGA 3.1 (Kumar et al., 2004) to construct a dendrogram of average distance using the unweighted pair group method with arithmetic mean (UPGMA).

The program "Structure" (Pritchard et al., 2000), based on Bayesian statistics, was used to infer the number of groups (k). We conducted 10 runs for each K value, with 200,000 burn-ins and 500,000 Markov chain Monte Carlo simulations. To determine the most probable value of K, we used the criteria proposed by Pritchard and Wen (2004) and Evano et al. (2005).

RESULTS

For *T. speciosum*, 22 of the 23 loci tested amplified alleles. Of these, we selected 13 loci to be used in the analysis: *mTcCIR2*, *mTcCIR3*, *mTcCIR4*, *mTcCIR7*, *mTcCIR9*, *mTcCIR10*, *mTcCIR11*, *mTcCIR13*, *mTcCIR18*, *mTcCIR19*, *mTcCIR22*, *mTcCIR26*, and *mTcCIR28*.

For *T. subincanum*, all 23 loci amplified an allele. Of these, 13 loci were selected for use in this study: *mTcCIR1*, *mTcCIR2*, *mTcCIR3*, *mTcCIR4*, *mTcCIR7*, *mTcCIR9*, *mTcCIR10*, *mTcCIR11*, *mTcCIR17*, *mTcCIR19*, *mTcCIR22*, *mTcCIR26*, and *mTcCIR28*.

With one exception, *mTcCIR11* for *T. speciosum*, the primers used in the study identified polymorphisms at their target loci. In *T. speciosum*, a total of 65 alleles were detected, with an average of 5.0 per locus (Table 1). For *T. subincanum*, 87 alleles were detected with an average of 6.69 alleles per locus (Table 2).

Table 1. Allelic frequency (Fa), number of alleles (N_A), genetic diversity (Dg), expected heterozygosity (H_E), observed heterozygosity (H_O), and polymorphism information content (PIC) of 13 SSR primers in 25 *Theobroma speciosum* individuals from the Juruena National Park, Mato Grosso State, Brazil.

Primer	Fa	N_A	Dg	H_E	H_O	PIC
<i>mTcCIR2</i>	0.91	4	0.15	0.96	0.04	0.15
<i>mTcCIR3</i>	0.52	3	0.60	0.92	0.00	0.53
<i>mTcCIR4</i>	0.23	8	0.84	0.92	0.08	0.82
<i>mTcCIR7</i>	0.64	4	0.51	0.96	0.04	0.45
<i>mTcCIR9</i>	0.25	8	0.83	0.96	0.33	0.81
<i>mTcCIR10</i>	0.32	5	0.77	0.92	0.73	0.73
<i>mTcCIR11</i>	1.00	1	0.00	1.00	0.00	0.00
<i>mTcCIR13</i>	0.48	4	0.66	1.00	0.00	0.61
<i>mTcCIR18</i>	0.41	6	0.69	0.96	0.83	0.64
<i>mTcCIR19</i>	0.82	2	0.29	1.00	0.36	0.25
<i>mTcCIR22</i>	0.28	7	0.81	1.00	0.12	0.79
<i>mTcCIR26</i>	0.24	6	0.82	1.00	0.00	0.79
<i>mTcCIR28</i>	0.28	7	0.82	1.00	0.00	0.80
Average	0.49	5.00	0.60	0.96	0.19	0.55

Table 2. Allelic frequency (Fa), number of alleles (N_A), genetic diversity (Dg), expected heterozygosity (H_E), observed heterozygosity (H_O), and polymorphism information content (PIC) for 13 SSR primers in 25 *Theobroma subincanum* individuals from the Juruena National Park, Mato Grosso State, Brazil.

Primer	Fa	N_A	Dg	H_E	H_O	PIC
<i>mTcCIR1</i>	0.56	4	0.62	1.00	0.00	0.57
<i>mTcCIR2</i>	0.18	9	0.86	1.00	0.52	0.85
<i>mTcCIR3</i>	0.31	5	0.75	0.88	0.00	0.71
<i>mTcCIR4</i>	0.34	6	0.76	0.92	0.00	0.73
<i>mTcCIR7</i>	0.36	5	0.73	1.00	0.00	0.69
<i>mTcCIR9</i>	0.26	6	0.80	0.92	0.00	0.77
<i>mTcCIR10</i>	0.19	8	0.86	0.84	0.00	0.85
<i>mTcCIR11</i>	0.31	6	0.75	0.96	0.16	0.71
<i>mTcCIR17</i>	0.25	10	0.85	0.96	0.50	0.83
<i>mTcCIR19</i>	0.31	6	0.76	0.96	0.08	0.72
<i>mTcCIR22</i>	0.33	8	0.80	0.96	0.12	0.77
<i>mTcCIR26</i>	0.27	7	0.81	0.96	0.37	0.79
<i>mTcCIR28</i>	0.32	7	0.80	1.00	0.40	0.77
Average	0.30	6.69	0.78	0.95	0.16	0.75

For *T. speciosum*, the largest number of alleles was found at the *mTcCIR4* and *mTcCIR9* loci with 8 alleles each; as mentioned above, *mTcCIR11* had only one allele. In *T. subincanum*, the largest number of alleles was found at *mTcCIR10* and *mTcCIR2* (10 and 9 alleles, respectively). Compared to that in *T. speciosum*, in *T. subincanum*, *mTcCIR4* identified only a single allele.

PIC values varied between 0.0 (*mTcCIR11*) and 0.82 (*mTcCIR4*) for *T. speciosum* and between 0.57 (*mTcCIR1*) and 0.85 (*mTcCIR2* and *mTcCIR10*) for *T. subincanum* (Tables 1 and 2).

The PIC analysis in *T. speciosum* showed that nine markers (*mTcCIR3*, *mTcCIR4*, *mTcCIR9*, *mTcCIR10*, *mTcCIR13*, *mTcCIR18*, *mTcCIR22*, *mTcCIR26*, and *mTcCIR28*) produced values above 0.5, two markers (*mTcCIR7* and *mTcCIR19*) produced values between 0.25 and 0.5, while two markers (*mTcCIR2* e *mTcCIR11*) produced values below 0.25. The average PIC value per locus was 0.55. For *T. subincanum*, all markers used produced values higher than 0.5, with an average of 0.75 per locus.

For all loci analyzed in both species, the observed heterozygosity was lower than the expected heterozygosity. The average heterozygosity observed for *T. speciosum* was 0.19, ranging from 0 (*mTcCIR3*, *mTcCIR11*, *mTcCIR13*, *mTcCIR26*, and *mTcCIR28*) to 0.83 (*mTcCIR18*), while the average expected heterozygosity was 0.96. For *T. subincanum*, the average observed heterozygosity was 0.15, ranging from 0 (*mTcCIR1*, *mTcCIR3*, *mTcCIR4*, *mTcCIR7*, *mTcCIR9*, and *mTcCIR10*) to 0.83 (*mTcCIR2*), while the average expected heterozygosity was 0.95. Of the 13 markers used in this study 11 (85%) presented low levels of heterozygosity for both species, with values lower than 50%.

The dendrogram analysis of *T. speciosum* individuals revealed two principal groups (Figure 2). Group I included only two individuals (18 and 19), while Group II contained the remaining 23 *T. speciosum* plants (92%). Group II was subdivided into two subgroups: Group IIA, which contained six individuals, and Group IIB, which included 17 individuals or 68% of the total analyzed sample.

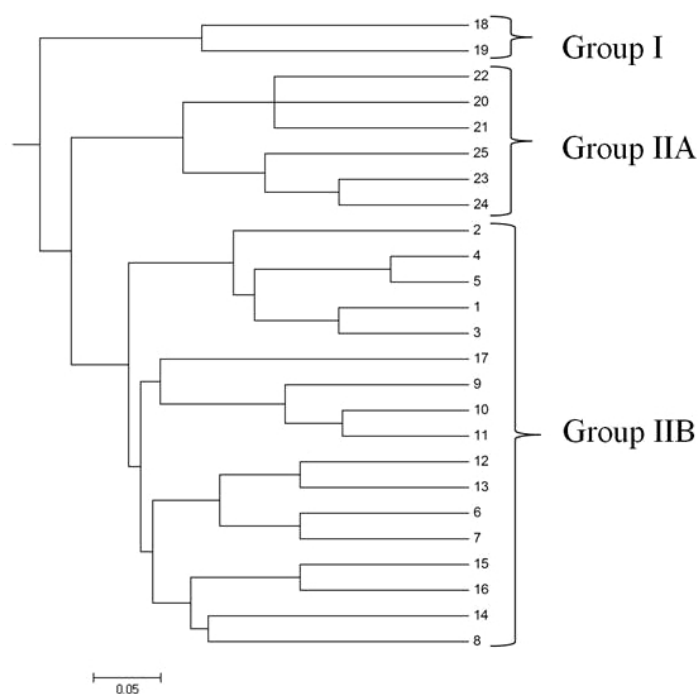


Figure 2. Dendrogram of the genetic similarity among 25 *Theobroma speciosum* individuals, obtained using the unweighted pair group method with arithmetic mean (UPGMA) and Nei's (1973) genetic distance as the measure of similarity.

The UPGMA analysis indicated plants 18 and 19 were the most dissimilar pair of individuals, while plants 4 and 5 were the most similar.

A dendrogram was also generated for *T. subincanum* and each individual was assigned into a major group using the UPGMA method (Figure 3). Group I consisted of 15 individuals and was divided into two subgroups (subgroup IA and IB). Group II contained 10 individuals

and was also divided into two subgroups: subgroup IIA with two individuals and subgroup IIB with eight individuals. According to the UPGMA analysis, individuals 14 and 17 were the most dissimilar and individuals 1 and 2 were the most similar.

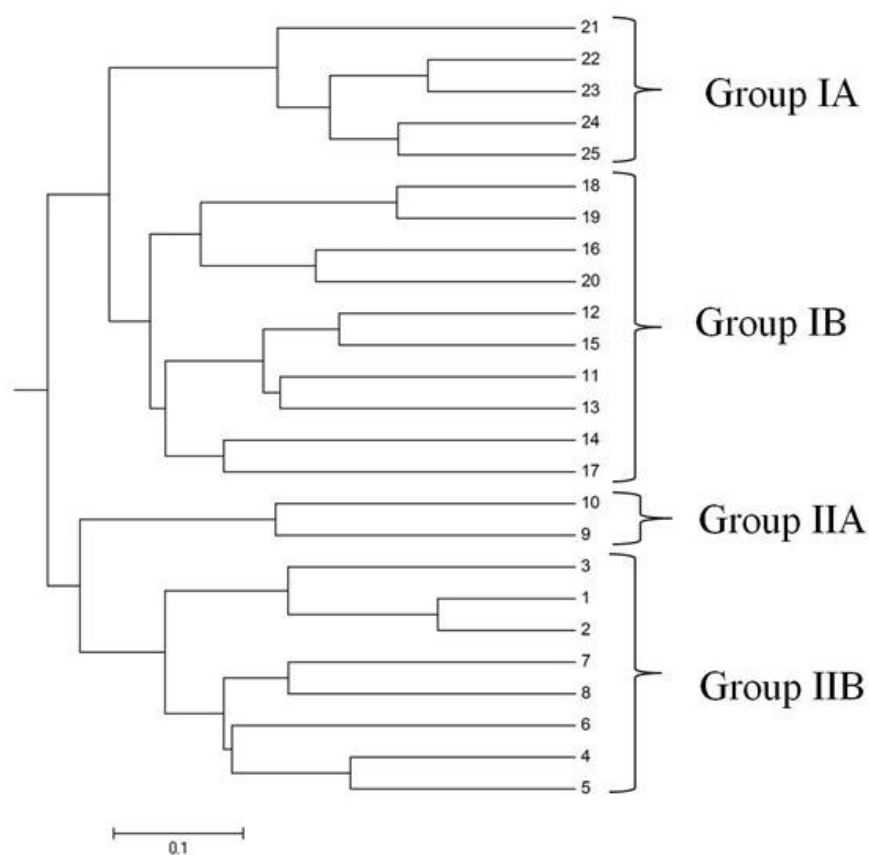


Figure 3. Dendrogram of the genetic similarity among 25 *Theobroma subincanum* individuals, obtained using the unweighted pair group method with arithmetic mean (UPGMA) and Nei's (1973) genetic distance as the measure of similarity.

The number of groups (K) in each species was also determined using the Structure program (Pritchard et al., 2000), which is based on Bayesian statistics. For *T. speciosum*, the best value for K was found for division of the sample into two groups ($K = 2$), which is supported by the clear distribution of individuals into two groups (Figure 4). Group 1 included 17 individuals and Group 2, eight individuals. In the grouping produced by the Structure program, Group 2 was consistent with our classification of Group I and IIA in the UPGMA dendrogram, and Group 1 included individuals classified as Group IIB by UPGMA.

Comparison of the outcomes of the Structure program (Bayesian analysis) and of UPGMA showed that some groups were consistent in both analyses, while others differed. Overall, the results produced by the Structure program suggested either further subdivisions or the aggregation of more genotypes.

For *T. subincanum*, the best value for K was found for division of the sample into three groups (K = 3): Group 1 contained eight individuals, Group 2 had 10 individuals, and Group 3 had seven individuals. Group 3 of the Structure program analysis included Groups IA and IB of the UPGMA dendrogram. Group 2 included individuals from Groups IB and IIA of the dendrogram, and Group 1 corresponded to Group IIB of the UPGMA analysis (Figure 5).

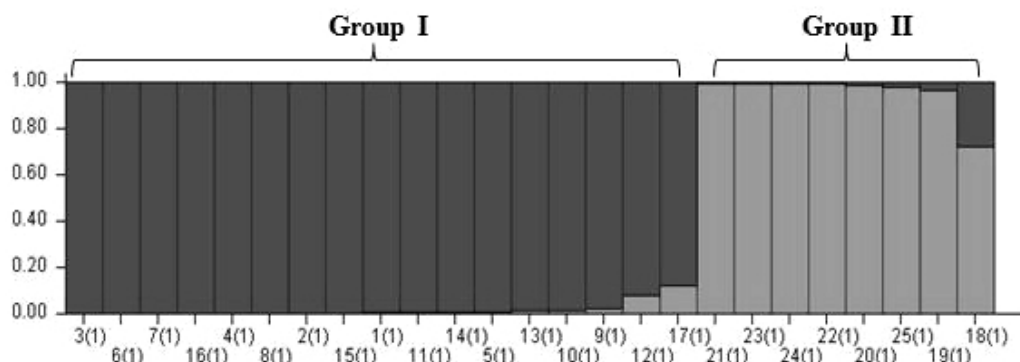


Figure 4. Distribution of the 25 *Theobroma speciosum* individuals into two groups based on molecular data from 13 SSR primers using the Structure program. Individuals are represented by vertical columns, and shaded according to their group (two groups, K = 2).

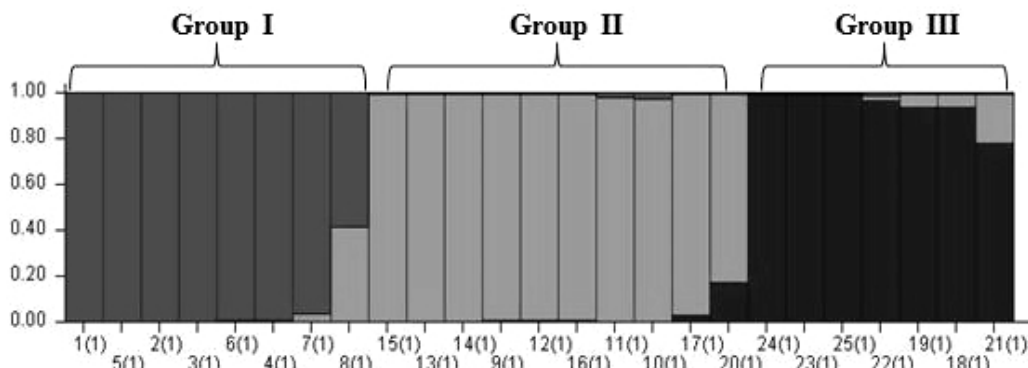


Figure 5. Distribution of the 25 *Theobroma subincanum* individuals into three groups based on molecular data from 13 SSR primers using the Structure program. Individuals are represented by vertical columns, and shaded according to their group (three groups, K = 3).

For both species, there were similarities between the dendrogram produced using UPGMA and the grouping generated by the Structure program; however, the two methods did not produce identical results. All individuals in both species sampled showed genetic diversity, as we did not find any pairs of individuals with similarity values of 0.

The results for *T. speciosum* suggest that individuals 18 and 19 of Group I, individual 20 from Group IIA, and individual 17 of Group IIB produced the greatest genetic distances within each group. These individuals are therefore potential seed sources for germplasm bank,

conservation, and survivorship studies. For *T. subincanum*, the individuals with the greatest potential as seed sources were numbers 21 from Group IA, 14 and 17 from Group IB, 9 and 10 from IIA, and 6 from Group IIB. These individuals had the largest genetic differences from the rest of the sample analyzed.

DISCUSSION

In this study, the total number of alleles detected in *T. speciosum* was 65, with an average of 5.0 per primer; the total number for *T. subincanum* was 87 alleles, with an average of 6.69 alleles per primer. In their analysis of 24 genotypes in cacao, Lanaud et al. (1999) found similar results, and recorded an average of 5.6 alleles per locus. Sereno et al. (2001), studying 94 cacao genotypes, identified 4.4 alleles per locus. By contrast, Mafra et al. (2011) found an average of 16.67 alleles per locus in their analysis of *T. cacao* populations in the states of Acre and Amazonas.

PIC values varied between 0.0 and 0.82 for *T. speciosum* and between 0.57 and 0.85 for *T. subincanum*. In an analysis of 43 plants from 12 different accessions of *Passiflora edulis*, Oliveira et al. (2005), obtained PIC values varying between 0 and 0.81; these results are similar to those found here for *T. speciosum*.

The average PIC values obtained in our analysis show the quality of the markers used. According to the classification outlined by Botstein et al. (1980), markers with PIC values higher than 0.5 are considered highly informative, values between 0.25 and 0.5 are considered moderately informative, and values below 0.25 are minimally informative. For *T. speciosum*, nine markers were highly informative, two were moderately informative, and two had little value. For *T. subincanum*, all the markers were highly informative.

CONCLUSIONS

With the exception of *mTcCIR11*, the loci used in this study showed evidence of polymorphism. For *T. speciosum*, the *mTcCIR4* and *mTcCIR8* loci had most alleles (8) and for *T. subincanum*, *mTcCIR17* had most alleles (10). The quality of the markers was confirmed by their PIC values, which averaged 0.55 for *T. speciosum* and 0.75 for *T. subincanum*.

Individuals 18 and 19 of *T. speciosum* showed the largest pairwise dissimilarity, whereas the most similar were individuals 4 and 5.

For *T. subincanum*, the most dissimilar pair was individuals 14 and 17, while the most similar were individuals 1 and 2.

The samples of *T. speciosum* and *T. subincanum* from the PnJu, Mato Grosso, Brazil showed extensive genetic diversity when analyzed using microsatellite markers. These plants therefore could be used for the creation of germplasm banks as well as for genetic improvement programs for both species.

ACKNOWLEDGMENTS

The authors wish to thank BIONORTE-MT (Projeto Conhecimento, Uso Sustentável e Bioprospecção da Biodiversidade na Amazônia Meridional, Processo: #554330/2010-5) contribution No. 13.

REFERENCES

- Barreto P, Barreto G and Verissimo T (2005). Pressão humana na floresta Amazônica brasileira. Governo do Estado do Pará, Belém.
- Bekessy SA, Allnutt TR, Premoli AC, Lara A, et al. (2002). Genetic variation in the vulnerable and endemic Monkey Puzzle tree, detected using RAPDs. *Heredity* 88: 243-249.
- Botstein D, White RL, Skolnick H and Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet. Res.* 32: 314-331.
- CDB (1992). Convenção sobre Diversidade Biológica. Ministério do Meio Ambiente, Brasília.
- Cuatrecasas J (1964). Cacao and its allies: a taxonomic revision of the genus *Theobroma*. Contribution from the United States National Herbarium, Washington, 35: 379-614.
- Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.
- Estopa RA, Souza AM, Moura COM, Botrel MCG, et al. (2006). Diversidade genética em populações naturais de candeia (*Eremanthus erythropappus* (DC) Macleish). *Sci. For.* 70: 97-106.
- Evano G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14: 2611-2620.
- Gaiotto FA, Grattapaglia D and Vencovsky R (2003). Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart.). *J. Hered.* 94: 399-406.
- Kumar S, Tamura K and Nei M (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* 5: 150-163. Available at [http://www.ncbi.nlm.nih.gov/pubmed/15260895]. Accessed July 26, 2013.
- Lanaud C, Risterucci AM, Pieretti I, Falque M, et al. (1999). Isolation and characterization of microsatellites in *Theobroma cacao* L. *Mol. Ecol.* 8: 2141-2143.
- Lemes JS, Müller KE, Giustina LD, Vieira FS, et al. (2010). Otimização da extração de DNA para quatro espécies do gênero *Theobroma* visando análises moleculares. In: Anais 3 Jornada científica da UNEMAT, Cáceres.
- Liu K and Muse S (2005). PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21: 2128-2129.
- Lowe A, Harris S and Ashton P (2005). Ecological genetics: design, analysis, and application. Oxford, Blackwell.
- Mafra GS, Gusmão MTA, Silva CRS and Albuquerque PSB (2011). Diversidade genética, endogamia e frequência alélica de populações de cacauero do Acre e Amazonas. In: Anais do 9º Seminário Anual de Iniciação Científica, Amazônia.
- Nei M (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U. S. A.* 70: 3321-3323.
- Oliveira NA and Amaral IL (2004). Florística e fitossociologia de uma floresta de vertente na Amazônia Central, Amazonas, Brasil. *Acta Amaz.* 34: 70-81.
- Oliveira EJ, Pádua JG, Zucchi MI, Camargo LEA, et al. (2005). Development and characterization of microsatellite markers from the yellow passion fruit (*Passiflora edulis* f. *flavicarpa*). *Mol. Ecol. Notes* 5: 331-333.
- PNJU (Parque Nacional do Juruena) (2013). Unidades de Conservação Brasileira-PNJU. Available at [http://www.parquenacionaldojuruena.com.br]. Accessed August 20, 2013.
- Pritchard J, Stephens M and Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Pritchard JK and Wen W (2004). Documentation for structure software. Version 2.1. Available at [http://pritch.bsd.uchicago.edu]. Accessed August 21, 2013.
- Sena MC, Oliveira MCS, Reginato LCA, Roese AD, et al. (2007). Fundamentos teórico-práticos e protocolos de extração e de amplificação de DNA por meio da técnica de reação em cadeia da polimerase. Embrapa Pecuária Sudeste, São Carlos.
- Sereno ML, Figueira A and Albuquerque PSB (2001). Estimación de la diversidad genética de poblaciones silvestres de *Theobroma cacao* L. Amazónico Brasileño, mediante microsatélites. In: Simpósio de Recursos Genéticos para a América Latina e Caribe, Londrina, 418-420.
- Sousa NR, Filho FJN and Souza AGC (2009). Caracterização, avaliação e documentação de recursos genéticos de espécies amazônicas. In: Domesticação e melhoramento: espécies amazônicas (Borém A, Lopes MTG and Clemente CR, eds.). Editora UFV, Viçosa.
- Souza DR, Souza AL, Leite HG and Yared JAG (2006). Análise estrutural em floresta ombrófila densa de terra firme não explorada, Amazônia Oriental. *Rev. Árvore* 30: 75-87.
- Young A and Boyle T (2000). Forest fragmentation. In: Forest conservation genetics: principles and practice (Young A, Boshier D and Boyle T, eds.). CABI, Wallingford, 123-134.