GMR

Genetic diversity of *Schizolobium parahyba* var. *amazonicum* (Huber ex. Ducke) Barneby, in a forest area in Brazil

A.L. Silva Júnior[1], L.C. Souza[1], A.G. Pereira[1], M.V.W. Caldeira[2] and F.D. Miranda[3]

[1]Programa de Pós-Graduação em Genética e Melhoramento,
Laboratório de Bioquímica e Biologia Molecular,
Centro de Ciências Agrárias e Engenharias,
Universidade Federal do Espírito Santo, Alegre, ES, Brasil
[2]Departamento de Ciências Florestais e da Madeira,
Universidade Federal do Espírito Santo, Jerônimo Monteiro, ES, Brasil
[3]Departamento de Biologia, Centro de Ciências Exatas Naturais e da Saúde,
Universidade Federal do Espírito Santo, Alegre, ES, Brasil

Corresponding author: A.L. Silva Júnior
E-mail: adelsonlemes@yahoo.com.br

Genet. Mol. Res. 16 (3): gmr16039774
Received July 10, 2017
Accepted August 25, 2017
Published September 21, 2017
DOI http://dx.doi.org/10.4238/gmr16039774

ABSTRACT. *Schizolobium parahyba* var. *amazonicum* (Fabaceae) is an arboreal species, endemic to the Amazon Rainforest, popularly known as paricá. It is used on a commercial scale in the timber sector, pulp and paper production, reclamation projects in degraded and landscaped areas. However, there is no availability of genetically improved material selected for the environmental conditions of the State of Espírito Santo, Brazil. In this sense, the present study aimed to characterize the genetic diversity in a population of *S. amazonicum*, established in a forest area in the southern region of the State of Espírito Santo, using inter-simple sequence repeat (ISSR) molecular markers. DNA samples from 171 individuals were analyzed using 11 ISSR primers, which generated 79

polymorphic bands in a total of 136 fragments (58%). The polymorphic information content performed for the ISSR markers revealed a mean of 0.37, classifying them as moderately informative. The number of loci found (N = 79) was greater than that established as the optimal number (N = 69) for the analyses. High genetic diversity was found with the parameters, genetic diversity of Nei (H_E = 0.375) and Shannon index (I = 0.554). The data demonstrated in the dendrogram, based on the UPGMA cluster analysis, corroborated by the Bayesian analysis performed by the STRUCTURE program, which indicated the formation of two distinct clusters (K = 2). One of the groups was formed with the majority of the individuals (153 genotypes) and the second with the minority (18 genotypes). The results revealed high genetic diversity in the population of *S. amazonicum* evaluated in the present study, determining the potential of the population to be used as an orchard for seed collection and production of seedlings with confirmed genetic variability.

Key words: Molecular markers; ISSR; Tree species; Seed orchard; Genetic variability

INTRODUCTION

The demand for wood resources in Brazil and the world is increasing due to the population growth and consequently the increase in consumption. However, the history of logging due to the advancement of anthropic activities has led to the reduction of plant cover (Peres et al., 2010). The growing demand from state and federal governments for the expansion of plantations and production has encouraged studies using native species of commercial interest to increase forest cover, also subsidizing, conservation strategies of the plant species involved (Higa and Silva, 2006; Freitas et al., 2013).

The State of Espírito Santo, Brazil, has an area of 246.441 thousand hectares of planted forest (Indústria Brasileira de Árvores, 2016) and the goal for this area is to be expanded to 621 thousand hectares of forest planted by 2025. For this desirable scenario of the Capixaba Silviculture, one of the viable alternatives is the increase in the area of planted forests, using non-traditional forest species in different edaphic environments of the State (Espírito Santo - SEAG, 2008). Such an initiative must take into account the physical, physiological, and genetic quality of the seeds and seedlings, optimizing the genetic gains for each environment and maintaining the genetic variability in the planted areas (Degen and Sebbenn, 2014).

As a non-traditional species for the State of Espírito Santo, which has great potential for acclimatization and commercial interest, it stands out the *Schizolobium parahyba* var. *amazonicum* (Huber ex. Ducke) Barneby, belonging to the Fabaceae family, popularly known as paricá (Barneby, 1996). It is an arboreal species, native to the Amazon Forest, deciduous, with pollination bees and small insects as its main vectors. It has a fast growth, allowing its exploitation between 5 and 9 years of age, being used on a commercial scale by timber companies in the north and northeast regions of Brazil, and can also be used in pulp and paper production, reclamation projects in areas degraded and landscapes (Hoffmann et al., 2011).

Despite the potential of the species for different purposes, there is no availability of genetically improved material, selected for the environmental conditions of southern Espírito

Santo. Information on population genetic diversity and genetic structure is important in germplasm pre-improvement and conservation programs and can be obtained with the use of molecular markers that allow the detection of the current variability directly at the DNA level (Cruz et al., 2011). The inter-simple sequence repeat (ISSR) molecular marker stands out in this context because it is a universal marker, not requiring previous information of the DNA sequence of the species under study, has a dominant character, being highly informative and with a high repeatability (Ng and Tan, 2015).

In this sense, the present study aimed to characterize the genetic diversity in an *S. amazonicum* population, established in a forest area planted in the southern region of the State of Espírito Santo, using ISSR molecular markers. Thus, it is expected to obtain data to evaluate the potential of the population to be used shortly as an orchard for seed collection and production of seedlings with confirmed genetic variability, providing a basis for commercial plantations.

MATERIAL AND METHODS

Population sampling and vegetable material

Young leaves with a good phytosanitary appearance were collected in 171 individuals selected from a planted population of *S. amazonicum*. The planting was carried out in June 2011, through seeds acquired from seedlings located in Dom Elizeu - PA, Brazil, not having any information on the number of matrices used in the seed collection and genetic basis.

The study area is located at Instituto Federal de Educação Ciência e Tecnologia do Espírito Santo, Campus de Alegre, ES, on the highway ES 482, km 47, district of Rive, Alegre, ES, Brazil. It is between the coordinates 20°46'23.3 of South latitude and 41°27'25.9 of West longitude, with an average altitude of 152 m (Figure 1).

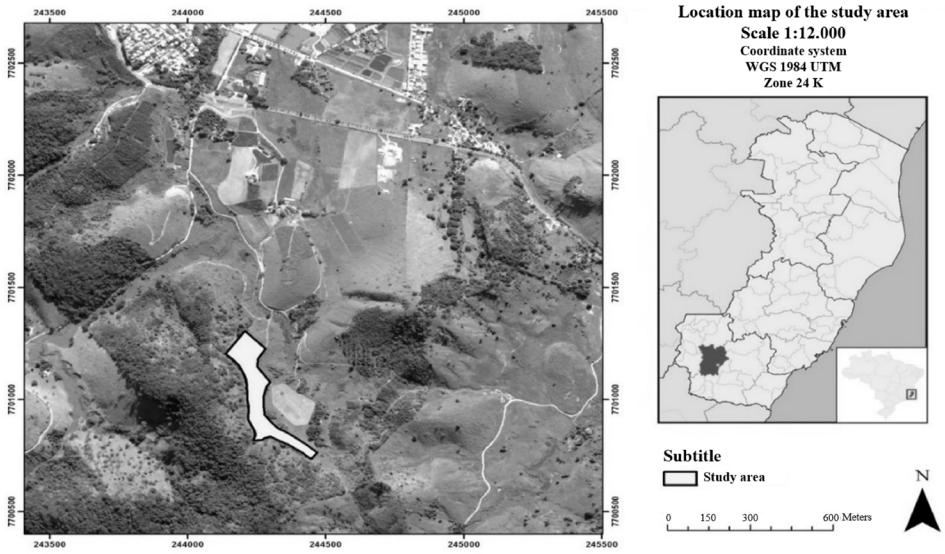

Figure 1. Map of location of the study area evidencing the area where the experiment was installed.

DNA extraction

Genomic DNA was isolated by the protocol of Doyle and Doyle (1990). The concentrations and purity of the samples were determined using the NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The parameter adopted to define the purity of the DNA samples was the ratio A_{260}/A_{280}, considered ideal when within the range of 1.8 to 2.0. Range less than 1.8 indicates protein contamination and above 2.0 indicates contamination with chloroform or phenol (Barbosa, 1998).

PCR amplification

Eleven ISSR primers produced by the University of British Columbia, Vancouver, Canada, were used for the genetic characterization of the 171 individuals of *S. amazonicum*. The amplification reactions were carried out at the Laboratory of Biochemistry and Molecular Biology, Universidade Federal do Espírito Santo, UFES (Brazil). The total reaction volume was 20 µL, containing a 1X buffer (10 mM Tris-HCl, pH 8.5, and 50 mM KCl), 2.5 mM $MgCl_2$, 0.25 mM of each dNTP, 0.2 µM primer, 1 U Taq DNA polymerase and about 50 ng genomic DNA. The following amplification conditions were used: initial denaturation at 94°C for 5 min, followed by 35 denaturation cycles (94°C for 45 s), annealing (52°C for 45 s) and elongation (72°C for 90 s), with a final extension at 72°C for 7 min. The amplification products were separated by electrophoresis on 2% agarose gel in 1X TBE for 5 h at 100 V.

Statistical analysis

The reading of the information contained on the gels allowed the generation of a binary matrix, from the presence (1) and absence (0) of bands in each locus. Subsequently, a descriptive analysis of the data, involving a total number of bands (TNB), number of polymorphic bands (NPB), the percentage of polymorphic bands (PPB) per primer, and the mean polymorphic information content (PIC), was performed. The PIC values for each ISSR locus were calculated as proposed by Roldán-Ruiz et al. (2000) for dominant markers, by the formula $PIC = 2fi\ (1 - fi)$, where fi = the frequency of the fragments present in the sample; $1 - fi$ = the frequency of absent fragments.

The optimal number of polymorphic loci needed to characterize genetic diversity in this study was evaluated by bootstrap analysis, being obtained the correlation values (r) and stress (E) (Kruskal, 1964). The matrix with the values of dissimilarity between the pairwise individuals was obtained with the arithmetic complement of the Jaccard coefficient (Jaccard, 1901). The individuals were grouped by the unweighted pair group method with arithmetic mean (UPGMA), where the calculation of the cut-off point was estimated as proposed by Mojema (1977), whose formula is described as $Pc = m + KSD$, where m = the mean of the distance values of the melting levels corresponding to the stages; K = 2.0; SD = standard deviation.

The cophenetic correlation coefficient (r) between the genetic dissimilarity matrix and the generated graphic image after the clustering was obtained to verify the consistency of the clustering. All these analyses were performed using the GENES program (Cruz, 2013). After obtaining these results, they were exported to the R program (R Development Core Team, 2016), to confection the graphic representation through a dendrogram.

To evaluate the genetic structure of the population, the Bayesian approach to the genotypic data was performed using the STRUCTURE 2.3 software (Pritchard et al., 2000), with 20 runs for each K value, where the number of established groups (K) was K = 1 to K = 4, with 7500 Monte Carlo interactions via Markov chains (MCMC) and an after burn-in length of 2500 interactions. The number of genetic groups was estimated by the *ad hoc* method ΔK proposed by Evanno et al. (2005) using the STRUCTURE HARVESTER software (Earl and Vonholdt, 2012). The genetic diversity parameters such as number of observed alleles (N_A), number of effective alleles (N_E), genetic diversity of Nei (H_E) (Nei, 1973), and Shannon index (I) (Shannon and Weaver, 1949), which classify the genetic diversity contained in the population, were calculated using the POPGENE program (Yeh and Boyle, 1997).

RESULTS AND DISCUSSION

Descriptive analyses

The set of 11 ISSR primers provided a total of 136 fragments in 171 subjects. Among these fragments, 79 were polymorphic (58% polymorphism), ranging from 5 (UBC 811, UBC 813, UBC 822, and UBC 855) to 13 (UBC 868) bands per primer, with a mean of 7.18 (Table 1). Based on the PIC, this in turn was described by Botstein et al. (1980) as a parameter of the quality of the marker in genetic studies, classifying it as satisfactory in polymorphic content when it presents a value superior to 0.5, medium informative between 0.25 and 0.50 and little informative when it has a value inferior to 0.25. According to Roldán-Ruiz et al. (2000), the calculation of PIC for dominant markers represents the probability of finding the marker in two states, absence or presence. For this study, the ISSR markers used can be classified as moderately informational, with an average value of 0.37 (Table 1).

Table 1. ISSR primers used in the present study including the total number of bands (TNB), number of polymorphic bands (NPB), the percentage of polymorphic bands (PPB) per primer, the size range of loci in base pairs (SBP) and polymorphism information content (PIC).

Primer	Sequence (5'- 3')	TNB	NPB	PPB (%)	SBP (max-min)	PIC
UBC 810	GAG AGA GAG AGA GAG AT	16	8	50.0	1090-410	0.34
UBC 811	GAG AGA GAG AGA GAG AC	10	5	50.0	1130-500	0.38
UBC 813	CTC TCT CTC TCT CTC TT	9	5	55.5	1150-700	0.46
UBC 814	CTC TCT CTC TCT CTC TA	13	9	69.2	1170-500	0.38
UBC 815	CTC TCT CTC TCT CTC TG	15	10	66.6	1190-570	0.33
UBC 822	TCT CTC TCT CTC TCT CA	8	5	62.5	1160-740	0.46
UBC 836	AGA GAG AGA GAG AGA GYA	9	6	66.6	1110-375	0.44
UBC 842	GAG AGA GAG AGA GAG AYG	12	6	50.0	1080-490	0.40
UBC 855	ACA CAC ACA CAC ACA CYT	10	5	50.0	1130-680	0.25
UBC 868	GAA GAA GAA GAA GAA GAA	20	13	65.0	1780-440	0.37
UBC 891	HVH TGT GTG TGT GTG TG	14	7	50.0	1080-500	0.33
Total	-	136	79	58%	-	0.37

*A = Adenine; T = Thymine; C = Cytosine; G = Guanine; H = (A, T or C); V = (A, C or G); and Y = (C or T).

In this study, we observed that the 79 polymorphic loci found were sufficient to determine the genetic diversity contained in the population, being superior to the minimum value of 69 loci obtained by the bootstrap analysis, which estimates the optimal number of ISSR loci necessary for the characterization of the genetical diversity. The correlation (r) found for the optimal value of loci was 0.962, and the stress value (E) was 0.047 (Figure 2). Values of E < 0.05 and correlation (r) closer to 1 indicate precision estimates (Kruskal, 1964).

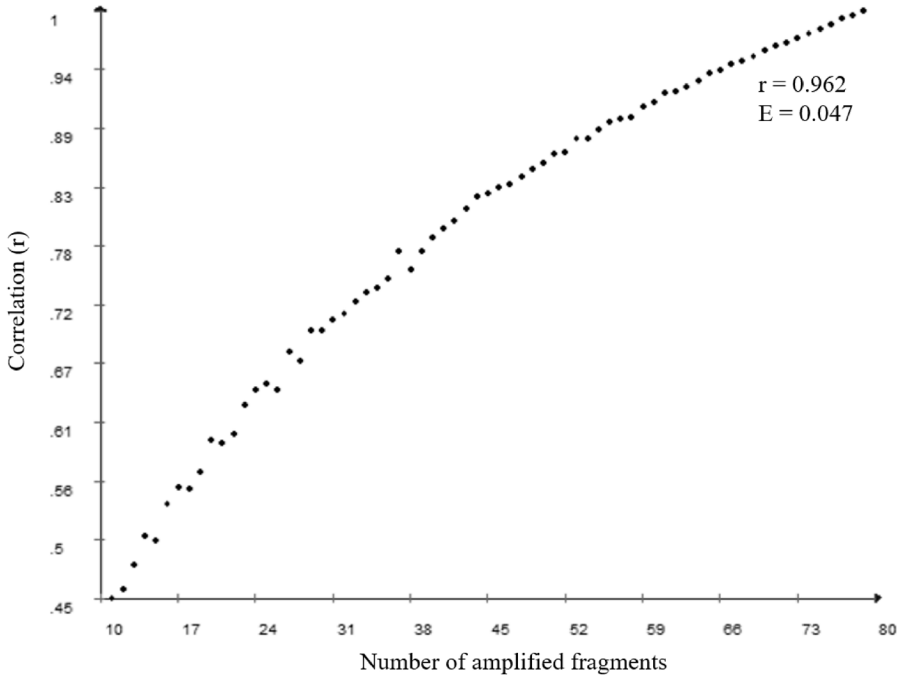

Figure 2. Estimates of the correlations (r) between the genetic similarities acquired by the bootstrap analysis and the optimal number of polymorphic fragments obtained for the 171 individuals of *Schizolobium amazonicum* using 11 ISSR primers.

Diversity and genetic structure

The N_A and the N_E revealed mean values of 2.00 and 1.65, respectively. Similar values were found in a study with a natural population of the tree species *Copernicia prunifera* using ISSR molecular markers, (N_A = 2.00 and N_E = 1.46) (Vieira et al., 2015).

H_E (0.375) and I (0.554) demonstrated high genetic diversity for *S. amazonicum*. According to Lewontin (1972), the values of I can vary from 0 to 1, and the closer to 1 the more diversified genotypically is the population.

In a study with the *S. parahyba* variety using random amplified polymorphic DNA (RAPD) markers, a value of 0.370 for the H_E was also obtained (Freire et al., 2007). Similar values for genetic diversity indexes were also found for other tree species in which similar methodologies were used. In studies of genetic diversity with *Prosopis cineraria* (Fabaceae) using ISSR markers, H_E = 0.301 and I = 0.438 were found (Sharma et al., 2011). High genetic diversity was also found for the species *Cornus mas* L., presenting H_E = 0.416 and I = 0.595 (Hassanpour et al., 2013).

The dissimilarity matrix obtained for the 171 individuals through the complementarity of the Jaccard coefficient based on the ISSR markers resulted in 29.241 dissimilarity indexes, with an average of 0.52, of which the 20 values of greater and less dissimilarity are presented in Table 2. The pair 110 x 122 had the lowest genetic distance value (0.0454), characterizing very similar individuals in genetic terms, while the 18 x 64 pair obtained the greatest genetic distance (0.8).

Table 2. Pairs of genotypes with higher and lower dissimilarity values (d) based on the complementarity of the Jaccard coefficient, calculated from 79 ISSR markers in 171 individuals of *Schizolobium amazonicum*, Espírito Santo, Brazil.

Order	Longer distance			Shorter distance		
	Individuals	Groupings	d	Individuals	Groupings	d
1	18 x 64	A x B	0.8000	110 x 122	A x A	0.0454
2	18 x 86	A x B	0.7857	63 x 110	A x A	0.0888
3	19 x 100	A x B	0.7857	63 x 122	A x A	0.1304
4	12 x 76	A x B	0.7846	152 x 153	A x A	0.2500
5	19 x 145	A x B	0.7666	101 x 157	A x A	0.2549
6	10 x 100	A x B	0.7636	32 x 98	A x A	0.2592
7	12 x 161	A x B	0.7627	50 x 54	A x A	0.2692
8	18 x 62	A x B	0.7627	127 x 132	A x A	0.2708
9	18 x 136	A x B	0.7592	69 x 113	A x A	0.2750
10	76 x 169	A x B	0.7555	4 x 8	B x B	0.2766
11	12 x 31	A x B	0.7541	90 x 157	A x A	0.2826
12	18 x 32	A x B	0.7538	98 x 131	A x A	0.2830
13	4 x 33	A x B	0.7500	123 x 131	A x A	0.2830
14	9 x 161	A x B	0.7500	2 x 6	B x B	0.2857
15	12 x 100	A x B	0.7500	160 x 166	A x A	0.2884
16	16 x 104	A x B	0.7500	97 x 133	A x A	0.2888
17	18 x 41	A x B	0.7500	145 x 157	A x A	0.2888
18	18 x 76	A x B	0.7500	21 x 23	A x A	0.2909
19	19 x 159	A x B	0.7500	3 x 89	A x A	0.2916
20	143 x 149	A x B	0.7500	55 x 57	A x A	0.2916

From the matrix of dissimilarity obtained between the individuals, a grouping by UPGMA was generated revealing two groups (Figure 3). Using the method of Mojema (1977) with k = 2.0 the recommended cut-off point of 92.24% was reached. Thus, one group was discriminated with the majority of the individuals (153 genotypes) and the second with the minority (18 genotypes). To confirm the relationship between the dissimilarity measures and the graphical measurements formed from the dendrogram, the cophenetic correlation coefficient (r) was obtained, which was favorable for this study, with a value of 68%.

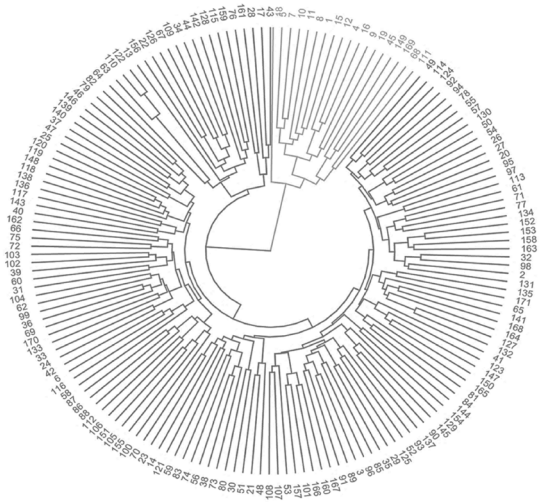

Figure 3. Dendrogram of genetic dissimilarity obtained by ISSR markers, using the Jaccard arithmetic complement based on the UPGMA method for 171 individuals of *Schizolobium amazonicum*, generating two genetic clusters throughout the population sample.

The Bayesian approach through the STRUCTURE program performed the analysis of the genotypic data and, according to the ΔK method described by Evanno et al. (2005), indicated that the most probable number of K clusters was defined as two (K = 2) (Figure 4). The data obtained by the Bayesian analysis in the STRUCTURE program followed the UPGMA method, similarly discriminating the individuals and separating them in the determined groups.

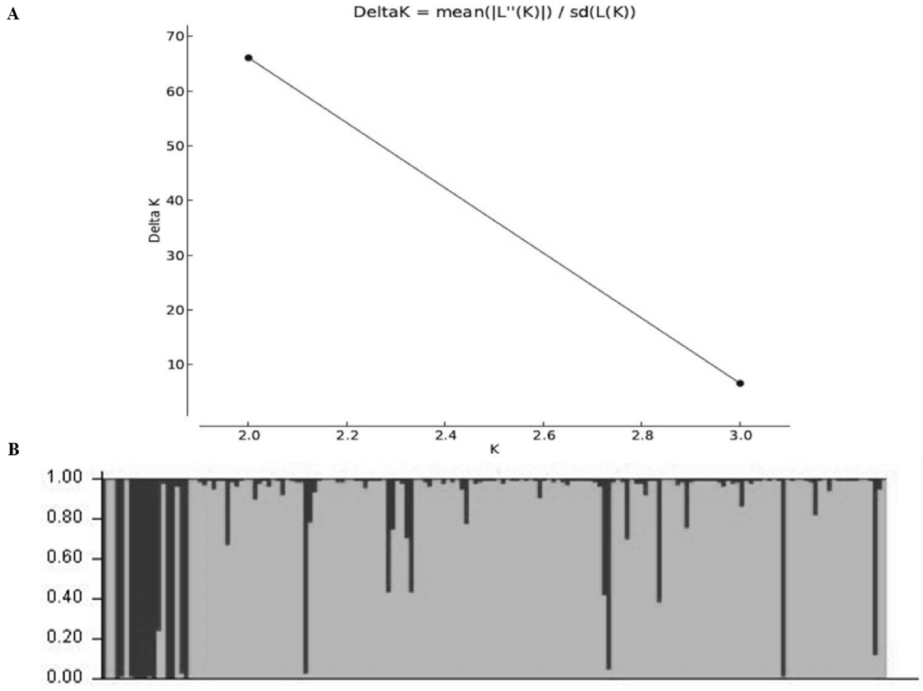

Figure 4. Bayesian approach using the STRUCTURE software for 171 individuals of *Schizolobium amazonicum*. **A**. Graph of the best value of ΔK for each value of K (best K = 2). **B.** Bar graph indicating the number of groups of the study population.

From the results obtained, it is important to point out that the genetic characterization was carried out in a planted forest, with only the information of the place of origin of the seeds that gave origin to the planting (Dom Elizeu - PA). Therefore, it is probable that the high genetic diversity and the consistent structure found were related to the establishment of the planting, being this one constituted by two distinct populations.

CONCLUSION

ISSR markers were efficient in the study of genetic diversity in *Schizolobium amazonicum*, determining the great potential of the population to be used as an orchard for seed collection and production of seedlings with confirmed genetic variability.

ACKNOWLEDGMENTS

The authors thank Fundação de Apoio à Pesquisa do Espírito Santo (FAPES) and Universidade Federal do Espírito Santo (UFES) for research support.

REFERENCES

Barbosa MM (1998). Quantificação e controle da qualidade do DNA genômico. In: Marcadores moleculares em plantas (Milach SCK, ed.). UFRGS, Porto Alegre.

Barneby RC (1996). Neotropical Fabales at NY: asides and oversights. *Brittonia* 48: 174-187. https://doi.org/10.2307/2807811

Botstein D, White RL, Skolnick M and Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.

Cruz CD (2013). GENES - a software package for analysis in experimental statistics and quantitative genetics. *Acta Sci. Agron.* 35: 271-276. https://doi.org/10.4025/actasciagron.v35i3.21251

Cruz CD, Ferreira FM and Pessoni LA (2011). Biometria aplicada ao estudo da diversidade genética. Suprema, Visconde do Rio Branco.

Degen B and Sebbenn AM (2014). Genetics and tropical forests. In: Trop. Forest. Handb (Pancel L and Köhl M, eds.). Heidelberg, Berlin. Avaiable at [http://doi.org/10.1007/978-3-642-41554-8_75-1].

Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Earl DA and Vonholdt BM (2012). Structure Harvester: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4: 359-361. https://doi.org/10.1007/s12686-011-9548-7

Espírito Santo - SEAG (2008). Plano estratégico de desenvolvimento da agricultura capixaba: novo pedeag: 2007-2025. Secretaria de Estado da Agricultura, Abastecimento, Aquicultura e Pesca, Vitória.

Evanno 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. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Freire JM, Piña-Rodrigues FCM, Lima ER, Sodré SRC, et al. (2007). Estrutura genética de populações de *Schizolobium parahyba* (Vell.) Blake (guapuruvu) por meio de marcadores RAPD. *Sci. For.* 74: 27-35.

Freitas EP, Moraes JFL, Filho AP and Storino M (2013). Indicadores ambientais para áreas de preservação permanente. *Rev. Bras. Eng. Agric. Ambient.* 17: 443-449. https://doi.org/10.1590/S1415-43662013000400013

Hassanpour H, Hamidoghli Y and Samizadeh H (2013). Estimation of genetic diversity in some Iranian cornelian cherries (*Cornus mas* L.) accessions using ISSR markers. *Biochem. Syst. Ecol.* 48: 257-262. https://doi.org/10.1016/j.bse.2013.01.002

Higa AR and Silva LD (2006). Pomares de Sementes de Espécies Nativas - Situação Atual. In: Pomar de Sementes de Espécies Florestais Nativas (Higa AR and Silva LD, eds.). Fundação de Pesquisas Florestais do Paraná, Curitiba.

Hoffmann RG, Silva GF, Chichorro JF, Ferreira RLC, et al. (2011). Caracterização dendrométrica de plantios de paricá (*Schizolobium amazonicum* Huber ex. Ducke) na região de Paragominas, PA. *Agraria* •••: 6.

Indústria Brasileira de Árvores (2016). Relatório IBÁ 2016. São Paulo.

Jaccard P (1901). Etude comparative de la distribuition florale dans une porion des Alpes et des Jura. *Bull. Soc. Vaud. Sci. Nat.* 37: 547-579.

Kruskal JB (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29: 1-27. https://doi.org/10.1007/BF02289565

Lewontin RC (1972). The apportionment of human diversity. *Evol. Biol.* 6: 381-398.

Mojema R (1977). Hierarchical grouping methods and stopping rules: an evaluation. *Comput. J.* 20: 359-363. https://doi.org/10.1093/comjnl/20.4.359

Nei M (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70: 3321-3323. https://doi.org/10.1073/pnas.70.12.3321

Ng WL and Tan SG (2015). Inter-Simple Sequence Repeat (ISSR) Markers. *ASM Sci. J.* 9: 30-39.

Peres CA, Gardner TA, Barlow J, Zuanon J, et al. (2010). Biodiversity conservation in human-modified Amazonian forest landscapes. *Biol. Conserv.* 143: 2314-2327. https://doi.org/10.1016/j.biocon.2010.01.021

Pritchard JK, Stephens M and Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

R Development Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Roldán-Ruiz I, Dendauw J, Bockstaele EV, Depicker A, et al. (2000). AFLP markers reveal high polymorphic rates in ryegrasses (Lolium spp.). *Mol. Breed.* 6: 125-134. https://doi.org/10.1023/A:1009680614564

Shannon CE and Weaver WA (1949). Mathematical model of communication. Urbana: University of Illinois Press.

Sharma SK, Kumar S, Rawat D, Kumaria S, et al. (2011). Genetic diversity and gene flow estimation in *Prosopis cineraria* (L.) Druce: A key stone tree species of Indian Thar Desert. *Biochem. Syst. Ecol.* 39: 9-13. https://doi.org/10.1016/j.bse.2010.12.018

Vieira FA, Sousa RF, Silva RAR, Fajardo CG, et al. (2015). Diversidade genética de *Copernicia prunifera* com o uso de marcadores moleculares ISSR. *Agraria* 10: 525-531. https://doi.org/10.5039/agraria.v10i4a5040

Yeh FC and Boyle TJB (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belg. J. Bot.* 129: 156-157.