




Nuclear and chloroplast microsatellites reveal high genetic diversity and structure in *Platonia insignis* Mart., an endangered species native to the Brazilian Amazon

Wellington F. Nascimento^{1*} , Gabriel Dequigiovanni² , Santiago L. F. Ramos³ , Caroline B. Garcia⁴ 
and Elizabeth A. Veasey⁴ 

Received: July 7, 2020

Accepted: November 19, 2020

ABSTRACT

Bacurizeiro (*Platonia insignis*) is a fruit of the Amazon that, due to anthropic actions, has been suffering serious damage to its genetic diversity. We analyzed the genetic diversity and structure of seven *bacurizeiro* populations distributed among four Brazilian states and two biomes, the Amazon and *Cerrado*, based on eight nuclear (ncSSR) and three chloroplast (cpSSR) microsatellite markers. Higher heterozygosity values were found for the Northern populations of Japurá and Marapanim, and the Northeast population of Chapadinha when using ncSSR. Higher diversity indices were also observed for these populations with cpSSR, although higher haplotypic diversity parameters were detected for the two Northern populations. Genetic structure analysis showed the formation of two well defined groups: I - populations from Maranhão and Piauí (*Cerrado*) and II - populations from Amazonas and Pará (Amazon). Positive and significant correlations were identified between genetic and geographical distances for both ncSSR ($r = 0.09$; $p = 0.0010$) and cpSSR ($r = 0.11$; $p = 0.0008$), in agreement with the genetic structure analysis. The high genetic structure among populations probably reflects the divergent natural and human selection pressures to which *bacurizeiros* are subjected in both the Amazon and *Cerrado* biomes, with higher diversity maintained in the Amazon populations.

Keywords: Amazonian fruits, *bacurizeiro*, conservation genetics, genetic diversity, molecular markers, haplotype network

Introduction

To reduce the impacts of genetic erosion on plant species, it is of fundamental importance to study the levels of diversity, distribution and genetic structure of their natural populations (Brown & Hodgkin 2015). Such information, in addition to others of ecological nature, may support

efficient strategies for the collection and conservation of species, whether *in situ* and/or *ex situ* (Whitlock *et al.* 2016).

Bacurizeiro, among others, is the popular name given to the species *Platonia insignis* (Clusiaceae) (Carvalho & Nascimento 2017), a semi-domesticated and native to the Amazon fruit tree (Clement 1999). This species is one of the few allogamous trees in the Amazon that reproduce both

¹ Centro de Ciências Agrárias e Ambientais, Universidade Federal do Maranhão, 65500-000, Chapadinha, MA, Brazil

² Centro Universitário Univel, 85806-080, Cascavel, PR, Brazil

³ Instituto de Ciências Exatas e Tecnologia, Universidade Federal do Amazonas, 69103-128, Itacoatiara, AM, Brazil

⁴ Departamento de Genética, Escola Superior de Agricultura 'Luiz de Queiróz', Universidade de São Paulo, 13418-900, Piracicaba, SP, Brazil

* Corresponding author: wellington.fn@ufma.br



by seeds and by sprouting from roots, and which presents sporophytic self-incompatibility (Maués & Venturieri 1996).

Due to intense anthropic actions in the regions where it occurs, the *bacurizeiro* has been suffering serious damage to its genetic diversity. In addition to Pará, a Brazilian state considered to be the center of origin and diversity of *bacurizeiro* (Cavalcante 2010), the species is found naturally in other states in the North, where its economic relevance is observed based on sustainable management and development (Alvarez & Potiguara 2013). A wide occurrence area of *bacurizeiro* is also found in the rocky relief further north of South America, called the Guiana Shield, which covers Guyana, French Guiana, Suriname, part of Venezuela and Colombia, and in Brazil, the states of Roraima, Amapá and the north of Amazonas and Pará (GBIF 2020). Probably, because this region has a pre-Cambrian geological formation, being one of the oldest formations on Earth, harboring different forms of life and biomes (Ferreira 2015), the Guiana Shield is more closely related to the possible *P. insignis* origin scenario. The *bacurizeiro* also occurs in the Northeast of Brazil, particularly in the Mid-North sub-region, formed by Maranhão and part of Piauí, where the transition forest between the Amazon Forest and the *Caatinga* biome, from the northeastern semiarid, is found (Carvalho & Nascimento 2017). In the transition forest, there is also the presence of the *Cerrado* biome, characterized by intense agricultural activity, where deforestation, long periods of drought and events of forest fires are quite common (Furley 1999; Pivello 2011).

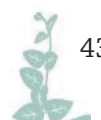
P. insignis fruits, called *bacuri*, are highly appreciated in the areas of occurrence of this fruit tree, both *in natura* and in the industry, for the production of sweets and ice cream (Souza *et al.* 2000). Although the pulp is the main product of *bacurizeiro*, an oil with anti-inflammatory properties is extracted from its seeds (Carvalho *et al.* 2003); and from the skin of the fruit an oil rich in palmitic acid and oleic acid is extracted (Villachica *et al.* 1996). Also, for being resistant its wood is used in the construction of boats and houses (Souza *et al.* 2007). Despite the great economic potential of *bacurizeiro*, there are no recent official statistical data related to the production and commercialization of the species, as observed for several other Amazonian fruits, such as the *pupunha* [*Bactris gasipaes* (Kunth)], *uxi* [*Endopleura uchi* (Huber)], *tucumã* (*Astrocaryum aculeatum* G. Mey) and several species of the genus *Oenocarpus*, known as *bacaba*. These species are classified as “invisible products from the Amazon”, but of great importance for the survival of local family farmers (Menezes 2002; Homma *et al.* 2018).

Despite the large offer of *bacuri* in the producing regions, few efforts have been made by the government and research institutions to learn more about the *bacurizeiro*, which justifies the scarcity of research aiming at the elaboration of conservation strategies, and, consequently, present a better use of the genetic variability of the species by current and

future generations (Lima 2007; Soares *et al.* 2013). The fact is that currently the accelerated process of genetic erosion caused by indiscriminate deforestation and growth of urban areas in areas where this species occur, has probably led to the loss of much of the genetic variability existing in the *bacurizeiro* (Santos *et al.* 2017).

Several molecular markers have been used to study the genetic diversity and genetic structure of populations. Among these, microsatellites (SSR-Simple Sequence Repeats), which are present in nuclear DNA (ncSSR) and in chloroplast DNA (cpSSR), have been extensively used for this purpose (Ellegren 2004). The ncSSRs are regions of DNA that have one to ten base pairs repeated in tandem, distributed throughout the genome of most eukaryotes; they are highly polymorphic, multi-allelic, codominant and have high reproducibility (Vieira *et al.* 2016). The cpSSRs are short sequences of mononucleotides that occur in non-coding regions of the chloroplast genome and that show intraspecific variation in the number of repetitions between individuals (Ebert & Peakall 2009). Although they are similar, unlike ncSSR, cpSSR are haploid and most of the time, do not perform the recombination process. These properties combined with uniparental inheritance tend to reduce the effective size of the chloroplast genome compared to the nuclear genome (Olmstead & Palmer 1994). This makes the variability of the chloroplast genome more vulnerable to evolutionary factors, such as genetic drift, thus being more effective in detecting population structure when compared to ncSSR (Petit *et al.* 2005).

Souza *et al.* (2013) when performing the molecular characterization of 72 accessions of *bacurizeiros* collected in the Maranhão and Piauí states, based on ISSR (Inter Simple Sequence Repeats), observed a high index of genetic diversity and differentiation between populations. Also based on ISSR, Pontes *et al.* (2017) and Pena *et al.* (2020) characterized accessions belonging to different locations in the Pará state and both observed greater genetic variation within the sampled populations. Only two studies have been developed so far based on microsatellite markers. Paraense *et al.* (2020) developed primers for the *bacurizeiro*, while Pena (2016) used these primers to characterize the germplasm bank of Embrapa Amazônia Oriental. It is worth mentioning that although there are two studies with ncSSR markers described for *P. insignis*, no information is available on cpSSR markers for this species. Also, there are no studies covering the entire area of occurrence of the species. Thus, using cpSSR and ncSSR markers, the objective of this study was to analyze populations of *bacurizeiro* in Brazil originating from the North and Northeast regions of Brazil, including two biomes (Amazon and *Cerrado*), which represents a broader sample when compared to previous studies, aiming to estimate the diversity and genetic structure of these populations, and propose appropriate strategies for the collection and conservation of the species.



Materials and methods

Sampling, DNA extraction and quantification

Young leaves were collected from adult individuals taken at random from seven populations: one population in the Amazonas state (AM), municipality of Japurá (03°10'01" S; 58°28'14" W) and one in the Pará state (PA), municipality of Marapanim (00°42'28" S; 47°41'52" W), both from the North region, Amazon biome; three populations in the Maranhão state (MA) distributed in the municipalities of Chapadinha (03°32'35" S; 43°54'57" W), Vargem Grande (03°32'35" S; 43°54'57" W) and Anapurus (03°71'16" S; 43°30'43" W), and two populations in the Piauí state (PI) distributed in the municipalities of Teresina (05°05'20" S; 42°48'07" W) and Barras (04°14'41" S; 42°17'40" W), from the Northeast region, *Cerrado* biome. The number of sampled plants varied from 4 to 16 individuals per population, in a total of 62 individuals (Tab. 1; Fig. 1). Due to the high capacity of clonal propagation of *bacurizeiro* and in order to represent significantly the genetic variability of the species, whenever possible, priority was given to collecting individuals at about 100 m away from each other.

DNA was extracted based on Doyle & Doyle (1990) with modifications. DNA was quantified based on the phage λ molecular size standards (Invitrogen) in different concentrations (50, 100 and 200 ng μL^{-1}) in a 1% (w/v) agarose gel stained with Gel Red.

Obtaining and analyzing ncSSR markers

Ten specific microsatellite primers developed by Paraense *et al.* (2020) were tested (Tab. S1 in supplementary material). The PCR amplifications were performed in the final volume of 13 μL containing: 1 μL of genomic DNA at 10 ng μL^{-1} ; 10 μL of SuperMix PCR (GeneDireX - MB200-P100) composed of Mg⁺⁺, dNTPs, *Taq* DNA polymerase; 0.5 μL of forward primer (10 pmol), 1 μL of reverse primer (10 pmol) and 0.5 μL of M13 primer (1.5 pmol). The amplification program used had an initial denaturation phase at 95 °C for 4 min, followed by 30 cycles under the following conditions: 30 s

for denaturation at 95 °C; specific annealing temperature for each primer for 1 min; and extension at 72 °C for 1 min. The final extension was performed at 72 °C for 10 min. The amplifications were purified with the Wizard® SV Gel kit and PCR Clean-Up System (Promega). Amplification quality and quantity were evaluated by electrophoresis in agarose 1% (w/v) gels stained with Gel Red by comparison with phage λ molecular size standards (Invitrogen).

The detection of polymorphism between individuals was carried out in an automatic sequencer performed at the Research Center on Human Genome and Stem Cells at the Universidade de São Paulo. The primers were marked with a specific fluorophore (6-FAM and HEX). The samples were then subjected to separation by capillary electrophoresis in an automatic sequencer ABI3500xl (Applied Biosystems). The determination of the sizes of the SSR fragments was performed with GeneMarker v.1.95 (Softgenetics 2021).

The following genetic parameters were estimated for each ncSSR locus and population: allele frequencies for each locus, number of alleles (*A*), allele richness (A_R), number of private alleles (A_p), number of multilocus genotypes (MLGs), observed (H_o) and expected (H_e) heterozygosities and Wright's fixation index (*f*). These parameters were obtained from hierfstat (Goudet 2005) and poppr (Kamvar *et al.* 2014) packages for R (R Development Core Team, 2018).

Bayesian analyzes were performed with Structure v.2.3.4 (Pritchard *et al.* 2000) to evaluate the genetic structure existing among *bacurizeiro* populations. Ten simulations were performed for each K, with a cluster number ranging from 1 to 10, considering the model of correlated allelic frequencies and ancestry admixture, 500,000 iterations of the Markov Monte Carlo Chain (MCMC) after an initial burn in period of 200,000 iterations. The most likely K number of clusters was determined by the method of Evanno *et al.* (2005) based on ΔK performed on Structure Harvester (Earl & vonHoldt 2012). Also, a Discriminant Analysis of Principal Components (DAPC) was performed with the *Adegenet* v.1.3.1 package (Jombart & Ahmed 2011) for R (R Development Core Team 2018). DAPC was performed using the machine learning method *K-means* (from the *find*.

Table 1. Genetic diversity parameters based on eight nuclear simple sequence repeats (ncSSR) loci for seven populations of *bacurizeiro* (*Platonia insignis*) collected in the Maranhão and Piauí states (Northeast region, *Cerrado* biome), and Amazonas and Pará states (North region, Amazon biome): number of individuals (*N*), average number of alleles per locus (*A*), allelic richness (A_R), number of private alleles (A_p), number of multilocus genotypes (MLGs), observed (H_o) and expected (H_e) heterozygosity and Wright's fixation index (*f*).

Populations	Biome	State	Municipality	<i>N</i>	<i>A</i>	A_R	A_p	MLGs	H_o	H_e	<i>f</i>
1	<i>Cerrado</i>	Maranhão	Vargem Grande	5	4.38	2.00	0	2	0.415	0.476	0.128
2	<i>Cerrado</i>	Maranhão	Chapadinha	16	12.75	4.88	2	4	0.632	0.559	-0.131*
3	<i>Cerrado</i>	Maranhão	Anapurus	5	4.88	2.62	1	1	0.529	0.536	0.013
4	<i>Cerrado</i>	Piauí	Teresina	7	6.75	2.38	1	2	0.543	0.528	-0.028
5	<i>Cerrado</i>	Piauí	Barras	4	3.88	2.12	0	1	0.233	0.481	0.516*
6	Amazon	Pará	Marapanim	16	14.88	8.88	4	8	0.885	0.620	-0.427*
7	Amazon	Amazonas	Japurá	9	8.62	5.25	3	6	0.875	0.665	-0.316*
			<i>Average</i>	8.9	8.02	4.02	1.57	3.43	0.587	0.552	-0.245*

* Significant with the confidence interval of 95%



Nuclear and chloroplast microsatellites reveal high genetic diversity and structure in *Platonia insignis* Mart., an endangered species native to the Brazilian Amazon

clusters function in *Adegenet*), with successive runs occurring to increase the number of clusters (K) and then transform the data using a Principal Component Analysis (PCA). For each model, a statistical measure of adequacy (Bayesian Information Criterion - BIC) was calculated to choose the best K (Jombart 2012). The best number of PCAs retained by means of the α -score function was defined. With this scatterplot, the membership probabilities based on the retained discriminant functions were generated.

The relationship between populations was assessed using a cluster analysis, based on the genetic distance of Cavalli-Sforza and Edwards (1967) calculated with the MSA v.4.05 program (Dieringer & Schlötterer 2003). The dendrogram using the Neighbor-Joining grouping method was obtained with Phylip v.3.6 (Felsenstein 2005). The stability of the groupings was assessed through a procedure of resampling with 1,000 bootstraps. Values higher than 70 in the nodes that join the groups indicate homogeneity among accessions.

The distribution of genetic variation, performed at three hierarchical levels: between regions (North and Northeast), between populations, and within populations, was evaluated based on the Analysis of Molecular Variance (AMOVA) with Arlequin v.3.5 (Excoffier & Lischer 2010). The AMOVA was also carried out within and between groups obtained from DAPC and Structure results. The correlation between genetic and geographic distance matrices, using Pearson's correlation (r), was evaluated by the Mantel test (Mantel 1967) with 9999 permutations and using NTSYS-pc software (Rohlf 2009).

Obtaining and analyzing cpSSR markers

Eight universal chloroplast microsatellite (cpSSR) primers described for non-coding regions of the dicots genomes were tested: ccmp 1, ccmp 2, ccmp 5 and ccmp 10 (Weising & Gardner 1999); and ccSSR 4, ccSSR 6, ccSSR 7 and ccSSR 9 (Chung & Staub 2003) (Tab. S2 in supplementary material). PCRs were performed in the final volume of 13 μ L containing: 1 μ L of genomic DNA at 10 ng μ L⁻¹; 10 μ L of SuperMix PCR (GeneDireX - MB200-P100) composed of Mg⁺⁺, dNTPs, *Taq* DNA polymerase; 0.5 μ L of the forward primer (10 pmol), 1 μ L of the reverse primer (10 pmol) and 0.5 μ L of the M13 primer (1.5 pmol). The amplification program used had an initial denaturation phase at 95 °C for 4 min, followed by 30 cycles under the following conditions: 30 s for denaturation at 95 °C; specific annealing temperature for each primer for 1 min; and extension at 72 °C for 1 min. The final extension was performed at 72 °C for 10 min. The amplifications were purified with the Wizard® SV Gel kit and PCR Clean-Up System (Promega). Gel image analysis, as well as the detection of polymorphism between individuals, were carried out the same way as described for ncSSR markers.

From a matrix of presence and absence the number of alleles (A) and Shannon-Weaver diversity index (I) were estimated. From the combination of cpSSR loci data, the total number of haplotypes (N_H), effective number of haplotypes (N_{EH}) and haplotypic diversity (H_{Ehap}) were estimated. Each unique allelic combination between the cpSSR loci was considered a haplotype, whereas the effective number of haplotypes was estimated based on the formula:

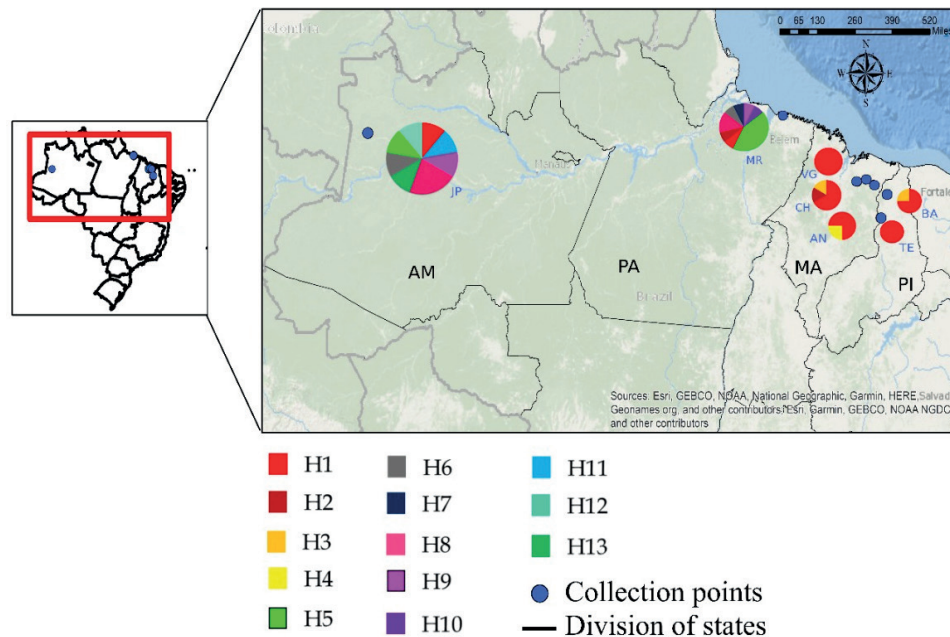


Figure 1. Map of Brazil with the collection locations of *Platonia insignis* and geographical distribution of chloroplast simple sequence repeats (cpSSR) haplotypes for each population, where each haplotype is identified by a different color indicated in the legend and the size of the circles is proportional to the number of haplotypes that compose them. The populations are originated from the municipalities of Vargem Grande (VG), Chapadinha (CH) and Anapurus (AN), all from Maranhão state (MA); Teresina (TE) and Barras (BA), from Piauí state (PI); Marapanim (MR), from Pará state; and Japurá (JP), from Amazonas state.

($\hat{n}_e = 1/\sum p_i^2$), where, p_i refers to the frequency of the i -th haplotype. The effective number of haplotypes is the number of equally frequent haplotypes that would produce the same observed homozygosity (Hartl & Clark 1997). All of these genetic parameters were obtained with GenALEX 6.5 (Peakall & Smouse 2012).

The distribution of haplotypic diversity was evaluated based on the Analysis of Molecular Variance (AMOVA) with Arlequin v.3.5 (Excoffier & Lischer 2010), with the same hierarchical distribution as for ncSSR. The correlation between genetic and geographic distance matrices, using Pearson's correlation (r), was evaluated by the Mantel test (1967) with 9999 permutations and using NTSYS-pc software (Rohlf 2009). The significance of the divergence measures was tested with 20,000 permutations. A haplotype network was built using the Median-Joining method (Bandelt *et al.* 1999) with the Network program (Forster *et al.* 2000).

Results

Genetic diversity and structure based on ncSSR

Of the 10 ncSSR loci tested, eight were 100% polymorphic (Tab. S3 in supplementary material). Loci PI 18 and PI 51 did not show good amplification or did not present clear profiles for the determination of alleles, and were, therefore, removed from the analysis. The number of alleles found per locus varied from 4 to 8 (average of 6 alleles/locus), with size ranging from 160 bp to 383 bp. The observed heterozygosity (H_o) ranged from 0.130 to 0.897, with an average of 0.656, while the expected heterozygosity (H_e) ranged from 0.520 to 0.776, with an average of 0.593. Wright's fixation index

(f), whose average was -0.102, ranged from -0.627 to 0.764. Five of the eight loci analyzed (PI 1, PI 6, PI 7, PI 27 and PI 49) resulted in negative fixation indexes, showing an excess of heterozygotes. With the exception of the PI 17 locus, all had a significant f value at 95 % confidence interval.

Estimates of genetic diversity per population (Tab. 1) showed an average number of alleles (A) of 8.02, ranging from 3.88 (Barras-PI) to 14.88 (Marapanim-PA). In addition to Marapanim-PA, the populations of Chapadinha-MA (12.75) and Japurá-AM (8.62) presented the highest number of alleles. The allelic richness (A_R) was also higher in the populations of Marapanim-PA, Japurá-AM and Chapadinha-MA, with values of 8.88, 5.25 and 4.88, respectively. The other populations showed allelic richness values equal or less than 2.62. Private alleles were observed in five populations, ranging from 1 (Anapurus-MA and Teresina-PI) to 4 (Marapanim-PA). As for the number of multilocus genotypes (MLGs), the highest values were found in the populations of Marapanim-PA, Japurá-AM and Chapadinha-MA, with 8, 6 and 4 multilocus genotypes, respectively.

The seven populations presented an average H_o of 0.587, with populations of Marapanim-PA, Japurá-AM and Chapadinha-MA showing the highest H_o values (0.885, 0.875 and 0.632, respectively) (Tab. 1). Barras-PI population had the lowest H_o (0.233). The average H_e was 0.552, with Japurá-AM, Marapanim-PA and Chapadinha-MA presenting the highest values, with 0.665, 0.620 and 0.559, respectively. Therefore, an excess of heterozygotes was observed in most of the studied populations, as shown in the negative values obtained for the Wright's fixation index (f) of Japurá-AM (-0.316), Marapanim-PA (-0.427), Chapadinha-MA (-0.131) and Teresina-PI (-0.028) populations. Except for the values obtained for the populations of Vargem Grande-

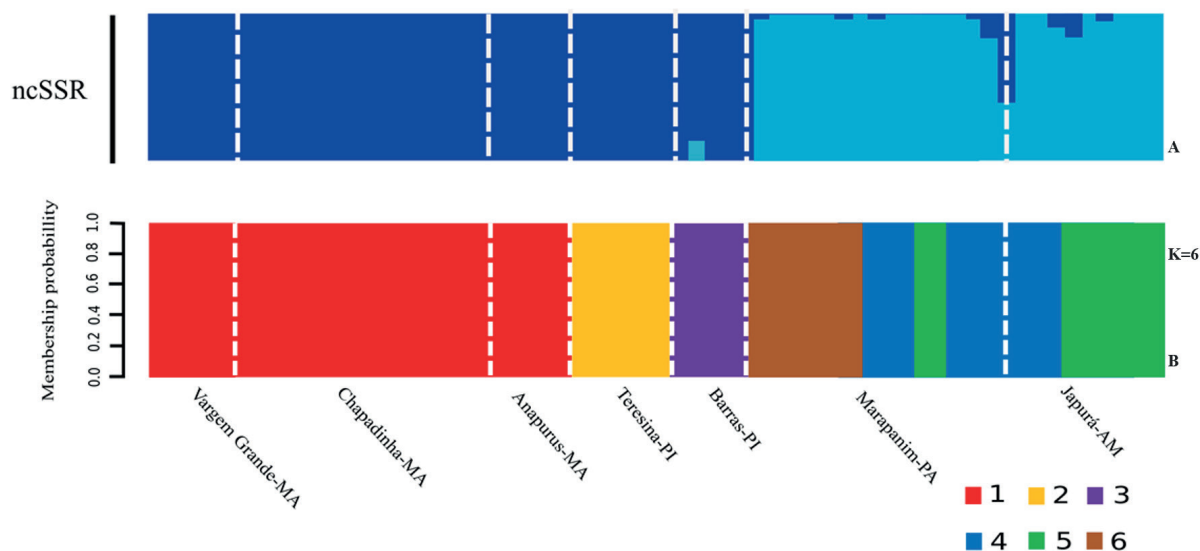


Figure 2. Genetic structure of *bacurizeiro* populations based on data from (A) nuclear simple sequence repeats (ncSSR) and (B) Discriminant Analysis of Principal Components (DAPC) obtained from ncSSR data. The dotted lines indicate the separation between the studied populations from the Maranhão (MA) and Piauí (PI) states (Northeast region); Pará (PA) and Amazonas (AM) states (North region).

**Nuclear and chloroplast microsatellites reveal high genetic diversity and structure in *Platonia insignis* Mart.,
an endangered species native to the Brazilian Amazon**

MA, Anapurus-MA and Teresina-PI, all *f* values obtained were significant at 95 % confidence interval.

The genetic structure analysis among *bacurizeiro* populations, carried out based on Bayesian analyzes, showed that the most likely *K* number determined by the method of Evanno *et al.* (2005) was equal to 2 (Fig. S1 in supplementary material), indicating that individuals from the seven populations were divided into two groups. The first group was formed by the populations of the Northeast region (Maranhão and Piauí, *Cerrado* Biome), while the second group was formed by the populations of the North region (Amazonas and Pará, Amazon biome) (Fig. 2A).

The analysis of molecular variance (AMOVA) between and within groups in the Bayesian analysis showed that most of the variation is distributed within groups (77 %), although significant differences are shown between groups (23 %, $F_{st} = 0.280$, $p < 0.001$), which also represents the differences between the two regions or the two biomes (Tab. 2). In agreement with the Bayesian analysis, the AMOVA showed significant differences between regions (20 %), and to a lesser extent between populations within regions (8%) (Tab. 2). However, most of the molecular variation was observed within the studied populations (72 %). The analysis showed that genetic variability is highly structured among the studied *bacurizeiro* populations ($F_{st} = 0.247$, $p < 0.001$).

In the dendrogram obtained from the ncSSR loci, the formation of two groups was also observed (Fig. 3), corroborating the genetic structure obtained from the Bayesian analysis. The first group was formed by individuals from the Maranhão and Piauí populations, and the second group was formed by individuals from Amazonas and Pará. When performing a more detailed analysis within groups, the first group was divided into three subgroups: subgroup 1, formed by individuals from Barras-PI; subgroup 2, formed by individuals from Teresina-PI; and subgroup 3, formed by individuals from the three populations of Maranhão. In the second group, four subgroups are observed: subgroup 1, formed by individuals from Marapanim-PA; subgroup 2 formed by individuals from Marapanim-PA and Japurá-AM; subgroup 3, formed by individuals only from Marapanim-

PA; and subgroup 4, formed by individuals from both populations, Marapanim-PA and Japurá-AM. Bootstrap values ranged between regions from 93 (group 2) to 95 (group 1) and from 55 to 89 between populations in each region (data not shown in the dendrogram).

The Discriminant Analysis of Principal Component (DAPC) performed using the *K-means* and *α-score* method to retain the best number of PCAs, considering at least 70 % probability of ancestry in one of the clusters ($q \geq 0.7$), detected six clusters (Fig. 2B), in agreement with both Bayesian and cluster analyses. There was a more accurate separation of populations according to their respective regions. The three populations of Maranhão formed a single genetic group. The populations of Piauí (Teresina and Barras) formed each one a distinct genetic group. The populations of the North region were classified into three groups, with the population of Pará (Marapanim) forming three distinct genetic groups, and two of these groups were also present in the population of Amazonas (Japurá). Similarly to the AMOVA conducted for the Bayesian analysis, when considering the distribution of the variation among and within DAPC groups, the AMOVA showed that most of the variation was concentrated within groups (78 %), although there was a significant variation between groups (22 %, $F_{st} = 0.288$, $p < 0.001$) (Tab. 2). Positive and significant correlation ($r = 0.09$; $p = 0.0010$) was identified between the genetic and geographical distances for the ncSSR data, demonstrating considerable spatial structuring of the genetic material in the sampled geographical area.

Diversity and haplotypic structure based on cpSSR

Only three out of eight universal chloroplast microsatellite loci (cpSSR) tested for *bacurizeiro* populations presented polymorphism: ccmp 1 and ccmp 2 (Weising & Gardner 1999), and ccSSR 9 (Chung & Staub 2003). The ccmp 1 and ccmp 2 loci had 3 alleles, while ccSSR 9 locus had 2 alleles (Tab. S4 in supplementary material). The average Shannon diversity index of the cpSSR loci estimated for the *bacurizeiro* populations was 0.330.

The analysis of cpSSR loci per population (Tab. 3) showed an average number of alleles of 1.82. The Marapanim-PA

Table 2. Molecular variance analysis (AMOVA) based on eight nuclear simple sequence repeats (ncSSR) loci for seven populations of *bacurizeiro* (*Platonia insignis*) collected in the Maranhão and Piauí states (Northeast region, *Cerrado* biome), and Amazonas and Pará states (North region, Amazon biome).

Variation sources	D.F.	S.S.	M.S.	P.V.	F Statistics
Between regions	1	63.691	63.691	20 %	$F_{st} = 0.247^*$
Between populations	6	34.072	6.814	8 %	
Within populations	116	287.921	4.806	72 %	
Total	123	385.684	75.311	100 %	
Between groups (Structure)	1	63.101	63.101	23 %	$F_{st} = 0.280^*$
Within groups (Structure)	122	322.954	64.983	77 %	
Total	123	386.054	128.084	100 %	
Between groups (DAPC)	5	97.424	19.385	22 %	$F_{st} = 0.288^*$
Within groups (DAPC)	118	209.083	3.372	78 %	
Total	123	306.507	22.757	100 %	

D.F.: Degree of freedom; S.S.: Sum of squares; M.S.: Mean square; P.V.: Percentage of variation; *Significant at $p < 0.001$ (999 permutation)



and Japurá-AM populations showed the highest average number of alleles values (2.7 and 2.3, respectively). For the Shannon diversity index (I), with an average value of 0.319, the populations of Japurá-AM (0.621), Chapadinha-MA (0.558) and Marapanim-PA (0.492) presented the greatest diversity. The populations of Vargem Grande-MA and Teresina-PI showed no diversity.

Thirteen different chloroplast haplotypes were observed among the seven populations analyzed with three cpSSR polymorphic loci, of which haplotype 1 (H1) showed the highest frequency (27 individuals) (Tab. S5 in supplementary material; Fig. 4). Haplotypes 2, 3, 5, 6, 8 and 9 presented frequencies between 2 and 7 individuals. The remaining haplotypes (4, 7, 10, 11, 12 and 13) occurred in only one individual. Regarding the number of haplotypes per population, despite having the same number of individuals evaluated, Marapanim-PA presented more haplotypes (8) than Chapadinha-MA (3) (Tab. 4; Fig. 1). The Japurá-AM population also presented 8 haplotypes. The populations of Anapurus-MA (H1 and H4) and Barras-PI (H1 and H3) presented two haplotypes, while the populations of Vargem Grande-MA and Teresina-PI presented only one haplotype (H1). Haplotype H4 was found only in the population of Anapurus-MA; H7 and H10 were exclusive to Marapanim-PA, while H11, H12 and H13 were exclusive to Japurá-AM (Fig. 1). Therefore, the haplotype network demonstrated the existence of an association between groups of haplotypes

and studied regions, and H1 was the only one shared among individuals from all studied populations (Fig. 4).

Regarding the effective number of haplotypes (N_{eH}), Japurá-AM and Marapanim-PA populations presented the highest values, 6.4 and 3.92, respectively (Tab. 4). The other populations showed much lower values, between 1.0 and 1.6. With the exception of Vargem Grande-MA and Teresina-PI, all other populations showed Shannon diversity indexes (I) greater than 0.5, especially the populations of Japurá-AM and Marapanim-PA, which presented values equivalent to 0.906 and 0.636, respectively. These two populations also showed greater haplotypic diversity (H_{Ehap}) (Japurá-AM, 0.964; Marapanim-PA, 0.802).

The AMOVA analysis for the cpSSR data showed significant differences between regions (61%), and to a much lower extent between populations within regions (1%). However, a considerable molecular variation was observed within the studied populations (38%). The analysis showed that genetic variability of cpSSR is highly structured among the studied *bacurizeiro* populations ($F_{st} = 0.372, p < 0.001$). Positive and significant correlation ($r = 0.11; p = 0.0008$) was identified between the genetic and geographical distances for the cpSSR data, demonstrating the spatial structure of the genetic material in the sampled geographical area, which is in accordance with the results presented above with ncSSR markers.

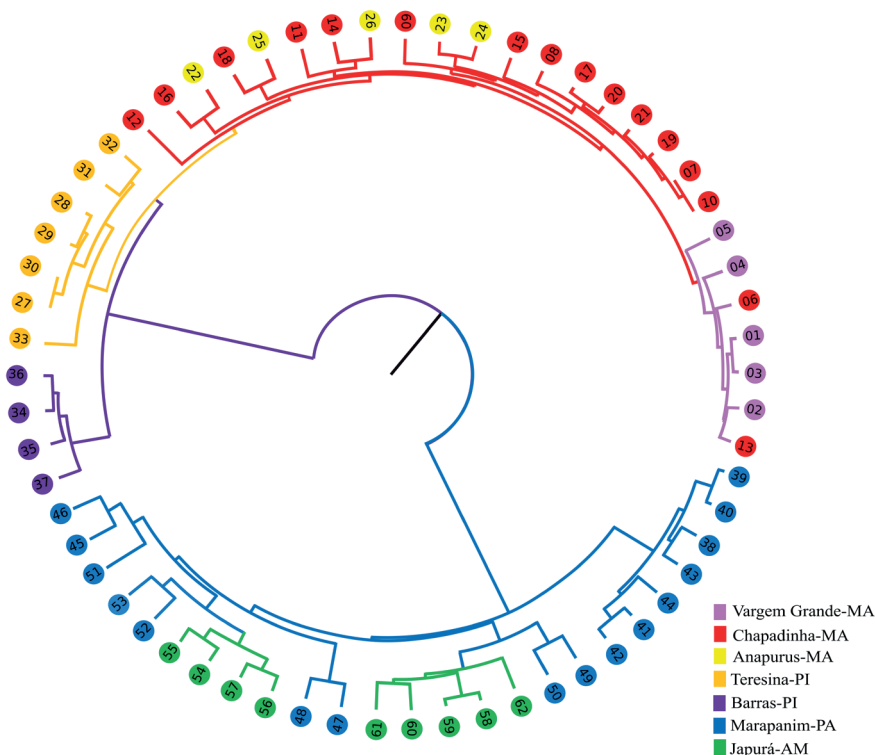


Figure 3. Dendrogram obtained based on eight nuclear simple sequence repeats (ncSSR) loci, demonstrating the genetic relationship between 62 individuals from seven Brazilian *bacurizeiro* populations, constructed by the Neighbor-Joining method and distances from Cavalli-Sforza and Edwards (1967). The numbers in the dendrogram indicate the accessions collected. Northeast region: 1-5 from Vargem Grande, 6-21 from Chapadinha, 22-26 from Anapurus, all from Maranhão states (MA); 27-33 from Teresina and 34-37 from Barras, all from Piauí states (PI). North region: 38-53 from Marapanim, Pará state (PA) and 54-62 from Japurá, Amazonas state (AM).

Discussion

Dynamics of genetic variability in *bacurizeiro* populations

In spite of presenting very efficient asexual reproduction from sprouts that appear spontaneously in the roots of adult plants, the *bacurizeiro* is a preferentially allogamous species

with most of its propagation occurring via seeds (Maués & Venturieri 1996). Such characteristics justify the fact the populations evaluated in this study showed moderate to high levels of diversity, with the two populations in the North region (Japurá and Marapanim) and one population in the Northeast (Chapadinha) having the highest levels of diversity for all parameters evaluated. These results demonstrate the greatest genetic diversity within the *bacurizeiro* populations

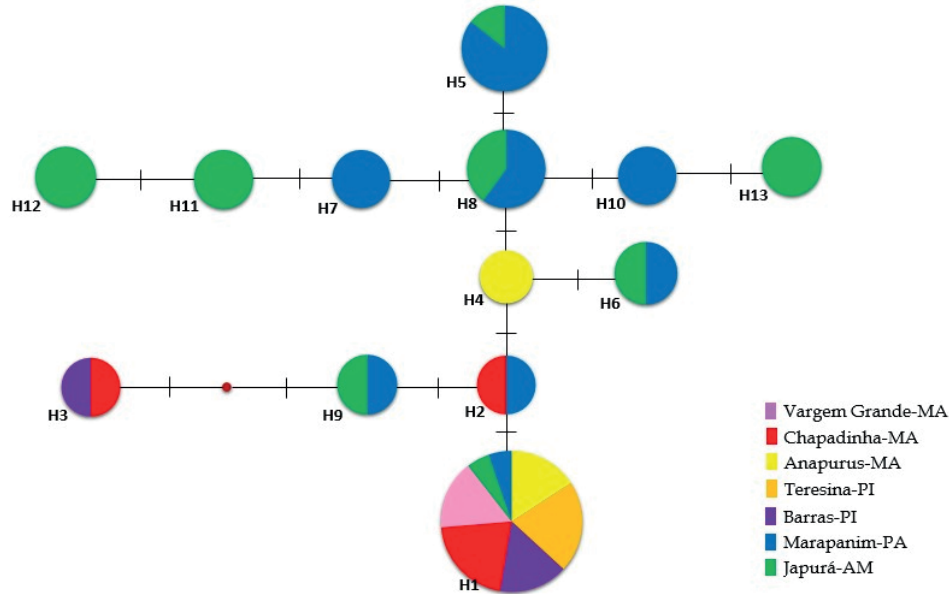


Figure 4. Haplotype network obtained from chloroplast simple sequence repeats (cpSSR) data observed in seven *bacurizeiro* populations. The haplotypes in pizza with more than one color indicate that individuals from different populations share the same haplotype. Each circle represents a haplotype and the size of the circles is proportional to the frequency of the haplotype. Each crossbar represents a mutation between the haplotypes.

Table 3. Genetic diversity parameters for the seven populations of *bacurizeiro* (*Platonia insignis*) studied based on three chloroplast simple sequence repeats (cpSSR) loci: A = average number of alleles; I = Shannon's index.

Population	Biome	States	Municipality	A	I
1	Cerrado	Maranhão	Vargem Grande	1.0	0.000
2	Cerrado	Maranhão	Chapadinha	2.0	0.558
3	Cerrado	Maranhão	Anapurus	2.0	0.249
4	Cerrado	Piauí	Teresina	1.0	0.000
5	Cerrado	Piauí	Barras	1.7	0.318
6	Amazon	Pará	Marapanim	2.7	0.492
7	Amazon	Amazonas	Japurá	2.3	0.621
Average				1.82	0.319

Table 4. Haplotypic diversity parameters for the *bacurizeiro* (*Platonia insignis*) populations studied based on three chloroplast simple sequence repeats (cpSSR): N = number of individuals, N_H = number of haplotypes, N_{eH} = effective number of haplotypes, I = Shannon index, H_{Ehap} = haplotypic diversity.

Population	Biome	State	Municipality	N	N_H	N_{eH}	I	H_{Ehap}
1	Cerrado	Maranhão	Vargem Grande	5	1	1.000	0.001	0.001
2	Cerrado	Maranhão	Chapadinha	16	3	1.412	0.566	0.318
3	Cerrado	Maranhão	Anapurus	5	2	1.471	0.500	0.400
4	Cerrado	Piauí	Teresina	7	1	1.000	0.020	0.003
5	Cerrado	Piauí	Barras	4	2	1.600	0.562	0.500
6	Amazon	Pará	Marapanim	16	8	3.920	0.636	0.802
7	Amazon	Amazonas	Japurá	9	8	6.400	0.906	0.964
Average				8.9	3.57	2.400	0.456	0.427



in the Amazon region, as also demonstrated by Clement (1999) and Cavalcante (2010). However, we must consider the fact that these populations were better sampled, together with the Northeast population of Chapadinha, Maranhão. Although the Northeast populations of *bacurizeiro* showed, in general, lower diversity indexes, an exception was the population of Chapadinha, Maranhão. This may be explained by the fact that this is the site of an old ecological *bacurizeiro* reserve (Homma *et al.* 2010), which may have concentrated higher diversity in this region.

Since *bacuri* is an allogamous species and presents self-incompatibility, an excess of heterozygotes would be expected, which is also maintained by clonal propagation. Therefore, these aspects may justify the excess of heterozygosity found in the majority of ncSSR loci analyzed in this study, as well as in a study of genetic divergence between *bacurizeiro* accessions of the germplasm bank of Embrapa Amazônia Oriental conducted by Pena (2016), based, among others, on the same microsatellite loci analyzed here. Five loci (PI 1, PI 6, PI 7, PI 27 and PI 49) showed negative fixation indexes (Tab. S3 in supplementary material), which may be related to the fact that when populations are cross-pollinated and present sporophytic self-incompatibility, that is, when the flowers are self-pollinated, but there is no growth of the pollen tube (Carvalho & Nascimento 2017; Homma *et al.* 2018), it enables the establishment of populations with higher number of heterozygotes. The Marapanim-PA, Japurá-AM, Chapadinha-MA and Teresina-PI populations showed an excess of heterozygotes (Tab. 1). On the other hand, the populations showing positive indexes of fixation, especially the population from Barras-PI, for a tree species such as *bacurizeiro*, may indicate a reduction in heterozygosity in a short period of time, which characterizes a *bottleneck* effect (Santos *et al.* 2017). According to Luikart *et al.* (1998), recent *bottleneck* events may have been caused by constant fires, which reduces the effective population and increases the possibility of inbreeding, especially when populations are isolated or genetically distant. The low number of individuals sampled in these populations is due to a very small number of plants, which may explain also the low diversity parameters.

The genetic structure analyses from both Bayesian (Fig. 2A) and cluster analysis (Fig. 3) based on ncSSR, showed the formation of two genetically distinct groups. The first group was formed by populations from the Northeast of Brazil (Maranhão and Piauí), located in the *Cerrado* biome, while the second group was formed by populations from the North of Brazil (Amazonas and Pará), from the Amazon biome. Although the DAPC analysis detected six clusters, where individuals were grouped according to their respective collection sites, with a more accurate separation of populations, this analysis also showed a clear separation between the two regions (North and Northeast) and biomes (Amazon and *Cerrado*). This genetic divergence between the two groups may be related to the pressure

of divergent selection that acts between populations that occur in the two biomes. It is worth mentioning that the SSR markers used here are neutral and mostly amplify regions that have non-coding DNA sequences, with these regions being subject to weaker selection pressure due to environmental conditions (Takrouni & Boussaid 2010). Thus, their variation does not necessarily reflect the pattern of variation in adaptive genes. On the other hand, the formation of two genetic groups in this study may be related to the fact that some ncSSR loci are associated with genes that are under different selection pressure due to the environmental conditions present in their respective biomes, and also possibly human selection.

Although originated from the Amazon forest (Clement 1999; Alvarez & Potiguara 2013), where there is great incidence of rains and the hot and humid climate predominates, the species has high phenotypic plasticity (Souza *et al.* 2000) and seems to be indifferent to these conditions, as it also resists pronounced water deficiencies, as well as high temperatures (Villachica *et al.* 1996), edaphoclimatic conditions prevalent in the collection sites of Maranhão and Piauí. Due to the presence of the species in northeastern Brazil, in areas with water deficit much more accentuated than in the naturally occurring areas in the Brazilian Amazon, it is possible for drought-tolerant ecotypes to occur in these areas. As to native *bacurizeiros* from the Maranhão state, they are clearly more tolerant to drought than plants native to the region of Belém in the Pará state (Carvalho & Nascimento 2017). Still, both in the Amazon and *Cerrado*, the presence of traditional communities, extractivists and *quilombolas* in disaggregation, was also observed in most of the places visited to collect *bacuri*. According to Homma *et al.* (2018), over a long time there was a continuous process of clearing of primary vegetation for the formation of gardens, logging and agricultural activities in these places. It is also likely that the Amerindians or *caboclos*, peoples responsible for the management of *bacuri* in the Amazon (FAO 1986; Clement 2000), selected genotypes with genetic and morphological characteristics different from those selected by the traditional communities present in the northeastern *Cerrado*.

In all three analyses made, Bayesian, cluster and DAPC, the AMOVA showed that most of the genetic variability is concentrated within the sampled populations, as well as within groups in the Bayesian and DAPC, although there was also considerable variation between the North and Northeast regions (20%), also observed in the Bayesian analysis as between groups (22%), and between groups in the DAPC analysis (23%). Therefore, we can conclude that the genetic variability of *bacurizeiro* populations is highly structured ($F_{st} = 0.247$, $p < 0.001$), also shown in both Bayesian and DAPC analyses ($F_{st} = 0.280$, $p < 0.001$; $F_{st} = 0.288$, $p < 0.001$, respectively), with a genetic difference between genotypes collected in the Amazon and *Cerrado* biomes. When characterizing the genetic diversity of 31



accessions of *P. insignis* collected in different locations in Pará based on ISSR, Pena *et al.* (2020) also observed greater variation within the sampled locations (98%), as well as the studies developed by Pontes *et al.* (2017) using 23 ISSR primers, also observing 98% genetic variation within the sampled locations (Soure and Salvaterra, municipalities of Pará). Studying ten populations in Maranhão and Piauí based on ISSR, Souza *et al.* (2013) observed that the distribution of genetic diversity was also greater within states (72.4%) and, consequently, within populations (71.8%). Oliveira and Silva (2008), evaluating the distribution of the variability of SSR markers in açai (*Euterpe oleracea* Mart.), a fruit native to the Amazon, also registered high variability within populations (68.9%), similarly to the results obtained in this study. The greatest amount of genetic variation within populations or locations is expected for allogamous species with self-incompatibility (Takrouni & Boussaid 2010; Nas *et al.* 2011), and may sometimes be related to the greater effort to collect samples within the locations and/or populations studied. The low differentiation found between some populations sampled in our study is, perhaps, an effect of the genetic and geographic proximity between them (Silva *et al.* 2009; Carvalho-Saraiva *et al.* 2014), especially those collected in Maranhão and Piauí.

First insights on the dynamics of haplotypic variation in bacurizeiro populations

The non-coding chloroplast DNA markers (cpDNA) are valuable resources for studies of phylogenetics and plant evolution. The uniparentally inherited nature of the cpDNA genome is particularly useful for detecting historical demographic processes, such as the extent of contractions, expansions and fragmentations (Hedréon *et al.* 2008). Of particular interest for population-level studies are chloroplast microsatellites due to the high amounts of polymorphisms. When genus and/or species specific cpSSR primers are not available, it is recommended to use of universal primers (Provan *et al.* 2001), as was the case in the present study.

Similarly, to the results obtained with ncSSR, cpSSR loci also showed greater variation in the populations of the Amazon, decreasing as the populations dispersed towards the northeastern *Cerrado*, most probably characterizing a *bottleneck* effect. This result is confirmed in Fig. 1, which shows greater diversity in terms of the number of haplotypes for populations in the North, in the Amazon, when compared to the Northeast, in the *Cerrado*. In addition, there are two exclusive haplotypes in the Northeast (H3 and H4), while there are nine that occur exclusively in populations in the North (H5 to H13). Marapanim-PA, Japurá-AM and Chapadinha-MA were the populations presenting higher values for the genetic diversity parameters, such as mean number of alleles and Shannon index. However, the two northern populations, Japurá-AM and Marapanim-PA, showed more haplotypes,

a higher effective number of haplotypes, as well as higher Shannon index and haplotypic diversity, exhibiting higher diversity for the Amazon populations.

The haplotype network demonstrated the existence of an association between groups of haplotypes and studied regions (Figs. 1 and 4). Based on these results, the presence of H1 in all studied populations indicates a shared common origin. As H4 was found only in the population of Anapurus-MA, H7 and H10 only in Marapanim-PA, and H11, H12 and H13 were exclusive to Japurá-AM, all geographically more distant populations, there may be, in general, little current gene flow among *bacurizeiro* populations (Qiu *et al.* 2013). However, the geographical proximity of some populations, such as Anapurus-MA, Vargem Grande-MA and Chapadinha-MA, seems to favor a greater gene flow between them, which was confirmed by the greater genetic similarity based on ncSSR loci. This greater genetic similarity may also be an effect of the expressive vegetative propagation and moderate dispersion of pollen and/or seeds of individuals from these populations. On the other hand, as *bacurizeiro* is an allogamous species, with self-incompatibility and greater variation within populations (results from ncSSR data), as well as other Amazonian tree species, including *cupuaçuzeiro* (*Theobroma grandiflorum* Schum.) (Alves & Figueira 2002) and *castanheira-do-Brasil* (*Bertholletia excelsa* H. B.) (Cavalcante 2008), the new clones growing from the same mother plant would be incompatible, which makes fertilization unfeasible and increases the dependence on pollen coming from other more distant genetically and geographically populations.

In *P. insignis*, contrary to what was obtained with the ncSSR data, the cpSSR markers showed most of the variation between regions (61%). This can be evidenced from the positive and highly significant correlation identified between the chloroplastidial genetic and geographical distances ($r = 0.11$; $p = 0.0008$), demonstrating the spatial structure of the genetic material in the sampled geographical area. On the other hand, high variation was also found within the analyzed *bacurizeiro* populations. Phumichai *et al.* (2015), when studying *seringueira* populations [*Hevea brasiliensis* (Willd. Ex Adr. De Juss.) Muell-Arg.], a perennial species native to the Amazon, based on cpSSR, observed greater genetic differentiation between populations (64.6%), although they also observed high variation within populations (35.3%).

The high genetic structure found for both cpSSR and ncSSR markers, demonstrated by the highly significant values of F_{st} , shows a low gene flow among *bacurizeiro* populations from the two regions or biomes, with the dispersion of pollens and seeds restricted to the nearest populations, probably due to the absence, limited presence or limited flight of dispersing agents; in addition to the fragmentation areas where the species occurs, which may also block pollen and seed dispersion (Maués & Venturieri 1996; Maués *et al.* 1996). This high level of genetic differentiation demonstrates that the *bacurizeiro* populations



of both biomes are in an intense process of divergence, which could eventually lead to a process of speciation.

There is no consensus regarding the pollinating agent of *bacurizeiro* flowers. It is known that bees enter their flowers to collect pollen, but only birds attracted by the nectar produced in abundance, pollinate them more successfully (Souza 2011). Among birds, Maués and Venturieri (1996) detected the performance of the Psittacidae (*papagaio-de-barriga-branca*, *marianinha-de-cabeça-amarela*, *periquito-da-asa-dourada* and *aratinga-de-bando*), Coerebidae (*sai-roxa*), Icteridae (*japiim-xexéu*) and Thraupidae (*pipira-vermelha*, *sanhaço-azul*, *sanhaço-do-coqueiro*) in the pollination of *bacurizeiros*. In the Maranhão state, Azambuja (2008) registered 22 species of birds visiting the *bacurizeiro* flowers, and of these, 74 % were Thraupidae and 12 % Psittacidae, the most abundant species being *Thraupis palmarum* (Wied-Neuwied) (38 %), *Nemosia pileata* (Wied-Neuwied) (20 %), *Diopsittaca nobilis* (L.) (9 %) and *Coereba flaveola* (L.) (8 %). Therefore, as the pollen grains are present in a viscous agglomerate, which makes it impossible for them to be dispersed by wind (Maués *et al.* 1996), the destruction of the biomes and surrounding regions, as well as the capture and sale of these birds, can constitute a serious risk to the *bacurizeiro* populations, restricting the gene flow between populations, increasing inbreeding and, consequently, the level of genetic structuration among them (Homma 2007).

Implications for conservation

Until the early 1980s, the species *P. insignis* had minimal risks of genetic erosion. However, in the early 1990s, the expansion of livestock and urban areas, as well as the cultivation of soybean [*Glycine max* L. (Merrill)], cowpea [*Vigna unguiculata* (L.) Walp.], and pineapple [*Ananas comosus* (L.) Merrill], prevented the natural regeneration of *bacurizeiros*, reduced the populations in their areas of occurrence and, consequently, placed their genetic heritage at risk (Nascimento *et al.* 2007; Souza 2011; Loch & Muniz 2016). On the other hand, there is a relevant occurrence of this species in the *Cerrado*, mainly in the transition zone between the North and Northeast regions of Brazil. The flora of this transition zone has species from adjacent biomes, which justifies the occurrence of *P. insignis*, native to the Amazon, but with significant dispersion in the *Cerrado* (Cerratinga 2020). In the most conserved areas of this region, groups of six to eight adult individuals were observed, about 40 m apart, as well as more than 200 young plants/ha, in more degraded areas, corroborating with Homma *et al.* (2010). According to these authors, in secondary vegetation on dry land, populations of *bacurizeiros* can occur in high density in a homogeneous way, suggesting some type of management in past times, as observed in the municipality of Chapadinha, Maranhão, site of an old ecological reserve of *bacurizeiro*.

Another predominant factor was also observed, the fires outbreaks, which are an imminent danger for the region where *bacurizeiros* occur (Gerude 2013). In the vegetation

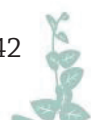
of the *Cerrado*, in periods of long drought, fire induces the detachment of the bark from the *bacurizeiro* trunk, releasing the resin, which intensifies the combustion and subsequent destruction of the trees (Homma *et al.* 2010; 2018).

Considering these factors that compromise the genetic variability of *bacurizeiro*, some strategies are recommended in order to avoid the reduction of populations and to diminish the effects of the genetic erosion in the species. Among these, implement measures to raise human awareness to avoid the formation of fire outbreaks; and planting genetically different seedlings brought from different locations for incorporation into agroforestry systems (Homma *et al.* 2010). Also, because the *bacurizeiro* is pollinated mainly by birds, it would be essential to conserve the fauna and the floristic resources that make up its environmental context, which would allow the success of the *bacurizeiro's* management, planting and conservation programs. It is worth mentioning that the remarkable capacity of clonal propagation of this species can determine that the area of one hectare is colonized by shoots of a single plant. In this case, the plants do not present genetic variation, which is a preponderant factor for the elaboration of collection strategies for *ex-situ* conservation of the species, being necessary to always prioritize the collection of very distant individuals, above 100 m, to try to represent significantly the genetic variability present in *P. insignis* (Homma *et al.* 2018).

Another important policy would be to create legislation to protect it. This aspect is an important alternative to promote the recovery of degraded areas and to restore Legal Reserve Areas (LRAs) and Permanent Preservation Areas (APPs). Banks and germplasm collections are also alternatives for the conservation of the genetic variability of a species. For Nass (2007), it is a way to reconcile the conservation of diversity with sustainable development. Conservation studies in a germplasm bank for *P. insignis* are rare, and only the efforts of Embrapa (Brazilian Agricultural Research Corporation) are known.

Results obtained in this study showed that the Japurá-AM, Marapanim-PA, and Chapadinha-MA are populations with greater genetic diversity. However, the two northern populations, Amazonas and Pará, showed greater haplotypic diversity. Also, despite the high genetic variability within the groups studied, the high genetic structure observed showing high differentiation among populations from the North and Northeast regions, possibly reflects the divergent selection pressure to which they have been subjected, as well as the low historical and recent gene flow between them. These phenomena, simultaneously, could lead to a process of speciation among populations of *bacurizeiros* occurring in their respective biomes over time.

In order to mitigate the effects of genetic drift, inbreeding and genetic erosion, we recommend the implementation of measures to raise awareness of the human being, in order to protect the areas where *bacurizeiros* occur, and fauna and flora resources that make up its environmental



context. In addition, considering the sporophytic self-incompatibility present in the *bacurizeiro*, it is necessary to carry out the planting of genetically different seedlings. It is worth mentioning that the high capacity for asexual reproduction of this species, makes it a priority to collect individuals that are geographically distant, in order to try to represent significantly the genetic variability of the species. The increase in the number of germplasm banks and collections would also be a good alternative for the *ex-situ* conservation of *bacurizeiros*, combining the conservation of diversity with the sustainable development of this important Brazilian Amazonian fruit, *P. insignis*.

Acknowledgements

We would like to thank the Universidade Federal do Maranhão and the Escola Superior de Agricultura “Luiz de Queiroz”/Universidade de São Paulo for providing the infrastructure to carry out the Project, to FAPESP (process no. 2019/04100-6) for financial support and to FAPEMA for a scholarship awarded to WFN (process no. BPD-08241/17). We also thank CNPq for the scholarship awarded to GD (Process n° 150297/2018-1) and EAV (Processes n° 301740/2016-0 and 309445/2020-5).

References

- Alvarez AS, Potiguara RCV. 2013. Caracterização anatômica foliar de espécimes de *Platonia insignis* Mart. (Clusiaceae) em diferentes períodos sazonais. *Bioscience Journal* 29: 562-569.
- Alves RM, Figueira A. 2002. Cupuassu (*Theobroma grandiflorum*) genetic resources and breeding in the Brazilian Amazon. *Ingenic Newsletter* 7: 25-32.
- Azambuja A. 2008. Interações entre *Platonia insignis* (Clusiaceae) e a avifauna visitante floral no cerrado do Maranhão. MSc Thesis, Universidade Estadual de Campinas, Campinas.
- Bandelt HJ, Forster P, Röhl A. 1999. Median-Joining network for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37-48.
- Brown AHD, Hodgkin T. 2015. Indicators of genetic diversity, genetic erosion, and genetic vulnerability for plant genetic resources. In: Ahuja S, Jain M. (eds.) *Genetic diversity and erosion in plants: indicators and prevention*. Switzerland, Springer. p. 25-23.
- Carvalho JEU, Nascimento WMO. 2017. *Platonia insignis*: Bacuri. In: Leggiadro R. Bacuri: *Platonia insignis*. Costa Rica, Instituto Interamericano de Cooperación para la Agricultura (IICA). p. 3-29.
- Carvalho JEU, Fontenelle DS, Müller CH. 2003. Propagação do bacurizeiro (*Platonia insignis* Mart.) por meio da raiz primária de sementes em início de germinação. Londrina, Informativo ABRATES. p. 1-5.
- Carvalho-Saraiva RV, Correia-Albuquerque PM, Girnos EC. 2014. Floral and vegetative morphometrics of three *Platonia insignis* Mart. (Clusiaceae) populations, a native tree from the Brazilian Amazon. *Plant Biosystems* 148: 666-674.
- Cavalcante MC. 2008. Visitantes florais e polinização da Castanha-do-Brasil (*Bertholletia excelsa*) em cultivo na Amazônia Central. MSc Thesis, Universidade Federal do Ceará, Fortaleza.
- Cavalcante PB. 2010. Frutas comestíveis na Amazônia. Museu Paraense Emílio Goeldi (Coleção Adolpho Ducke). <http://repositorio.museu-goeldi.br/bitstream/mgoeldi/901/1/P%20Avul%20n27%201974%20CAVALCANTE.pdf>.
- Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetics analysis: Models and estimation procedure. *American Journal of Human Genetics* 19: 233-257.
- Cerratinga. 2020. Bacuri: espécie da Amazônia e do Cerrado. <http://www.cerratinga.org.br/bacuri/>. 21 Sep. 2021.
- Chung SM, Staub JE. 2003. The development and evolution of consensus chloroplast primer pairs that possess highly variable sequence regions in a diverse array of plant taxa. *Theoretical and Applied Genetics* 107: 757-767.
- Clement CR. 1999. 1492 and the loss of Amazonian crop genetic resources. II. Crop biogeography at contact. *Economic Botany* 53: 203-216.
- Clemente CR. 2000. Bacuri: *Platonia insignis*. In: Clay JW, Sampaio PTB; Clemente CR. (eds.) *Biodiversidade amazônica: exemplos e estratégias de utilização*. Manaus, Ministério da Ciência e Tecnologia, Instituto Nacional de Pesquisa da Amazônia. p.159-166.
- Dieringer D, Schlötterer C. 2003. Microsatellite analyzer (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* 3: 167-169.
- Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.
- Earl DA, vonHoldt BM. 2012. Structure Harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359-361.
- Ebert D, Peakall R. 2009. Chloroplast simple sequence repeats (cpSSR): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Molecular Ecology Resources* 9: 673-690.
- Ellegren H. 2004. Microsatellite: simple sequences with complex Evolution. *Nature Reviews: Genetics* 5: 435-445.
- Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611-2620.
- Excoffier L, Lischer HEL. 2010. Arlequin suit ver. 3.5: A new series of programs to perform population genetics and analyses under Linux or Windows. *Molecular Ecology Resources* 10: 564-567.
- FAO - Food and Agriculture Organization of the United Nations. 1986. *Food and fruit-bearing forest species*. 3. Examples from Latin America. Roma, FAO. <https://agris.fao.org/agris-search/search.do?recordID=XF8659421>. 21 Sep. 2021.
- Felsenstein J. 2005. PHYLIP (Phylogeny Inference Package), version 3.6. <http://evolution.genetics.washington.edu/phylip.html>. 21 Sep. 2021.
- Ferreira RS. 2015. The Guiana Shield in a patrimonial perspective. *Revista CPC* 19: 136-163.
- Forster P, Bandelt HJ, Rohl A. 2000. NETWORK 3.1.1.0. <http://www.fluxus-engineering.com>. 21 Sep. 2021.
- Furley PA. 1999. The nature and diversity of neotropical savanna vegetation with particular reference to the Brazilian cerrados. *Global Ecology and Biogeography* 8: 223-241.
- GBIF - Global Biodiversity Information Facility. 2020. <https://www.gbif.org/>. 21 Sep. 2021.
- Gerude RG. 2013. Focos de queimadas em áreas protegidas do Maranhão entre 2008 e 2012. Foz do Iguacu, Instituto Nacional de Pesquisas Espaciais - INPE.
- Goudet J. 2005. Fstat version 1.2: A computer program to calculate F-statistics. *Journal of Heredity* 86: 485-486
- Hartl DL, Clark AG. 1997. *Principles of population genetics*. 3rd. edn. Sunderland, Sinauer Associates.
- Hedré M, Nordström S, Ståhlberg D. 2008. Polyploid evolution and plastid DNA variation in the *Dactylophiza incarnata/maculata* complex (Orchidaceae) in Scandinavia. *Molecular Ecology* 17: 5075-5091.
- Homma AKO. 2007. Manejando a planta e o homem: os bacurizeiros do Nordeste Paraense e da Ilha de Marajó. *Amazônia: Ciência & Desenvolvimento* 2: 119-135.
- Homma AKO, Menezes AJEA, Carvalho JEU, Matos GB. 2018. Manejo e plantio de bacurizeiros (*Platonia insignis* Mart.): a experiência no manejo e domesticação de um recurso da biodiversidade amazônica. *Inclusão Social* 12: 48-57.



- Homma AKO, Menezes AJEA, Carvalho JEU, Souto GC, Gibson CP. 2010. Manual de manejo de bacurizeiros. Belém, Embrapa Amazônia Oriental.
- Jombart T. 2012. A tutorial for Discriminant Analysis of Principal Components (DAPC) using adegenet 1. 3-4. Rvignette. London, Imperial College London, MRC Centre for Outbreak Analysis and Modelling. doi: 10.1038/72708
- Jombart T, Ahmed I. 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27: 3070-3071
- Kamvar ZN, Tabima JF, Grünwald NJ. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2: e281. doi: 10.7717/peerj.281
- Lima MC. 2007. Bacuri: (*Platonia insignis* Mart., Clusiaceae): Agrobiodiversidade. São Luis, Instituto Interamericano de Cooperação para a Agricultura.
- Loch VC, Muniz FH. 2016. Estrutura da vegetação de cerrado *stricto sensu* com extração do bacuri (*Platonia insignis* Mart.) em uma reserva extrativista, na região meio-norte do Brasil. *Revista de Biologia Neotropical* 13: 20-30.
- Luikart G, Sherwin WB, Steele BM, Allendorf FW. 1998. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology* 7: 963-974.
- Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 202-209.
- Maués MM, Venturieri GC. 1996. Ecologia da polinização do bacurizeiro (*Platonia Insignis* Mart.) Clusiaceae. Belém, Embrapa-CPATU.
- Maués MM, Venturieri GC, Souza LA, Nakamura J. 1996. Identificação e técnicas de criação de polinizadores de espécies vegetais de importância econômica no estado do Pará. In: Embrapa. (eds.) Centro de Pesquisa Agroflorestal da Amazônia Oriental (Belém, PA). Geração de tecnologia agroindustrial para o desenvolvimento do Trópico Úmido. Belém, Embrapa-CPATU.
- Menezes AJEA. 2002. Análise econômica da “produção invisível” nos estabelecimentos agrícolas familiares no Projeto de Assentamento Agroextrativista Praialta e Piranheira, Município de Nova Ipixuna, Pará. MSc Thesis, Universidade Federal do Pará, Belém.
- Nas MN, Bolek Y, Bardak A. 2011. Genetic diversity and phylogenetic relationships of *Prunus microcarpa* C.A. Mey. subsp. *Tortusa* analyzed by simple sequence repeats (SSRs). *Scientia Horticulturae* 127: 220-227.
- Nass LL. 2007. Recursos genéticos vegetais. Brasília, Embrapa Recursos Genéticos e Biotecnologia.
- Nascimento WMO, Carvalho JEU, Müller CH. 2007. Ocorrência e distribuição geográfica do bacurizeiro. *Revista Brasileira de Fruticultura* 29: 657-660.
- Oliveira MSP, Silva KJD. 2008. Diferenciação genética entre procedências de açaizeiro por marcadores RAPD e SSR. *Revista Brasileira de Fruticultura* 30: 438-443.
- Olmstead RG, Palmer JD. 1994. Chloroplast DNA systematics: a review of methods and data analysis. *American Journal of Botany* 81: 1205-1224.
- Paraense LCR, Pena DN, Darnet SH, Rodrigues SM, Menezes IC, Moura EF. 2020. First genomic microsatellite markers developed for *Platonia insignis* (Clusiaceae), a Brazilian fruit tree. *Molecular Biology Reports* 47: 2985-2989.
- Peakall R, Smouse PE. 2012. GenALEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28: 2537-2539.
- Pena DN. 2016. Divergência genética entre acessos de bacurizeiro (*Platonia insignis* Mart.) conservados na Embrapa Amazônia Oriental por meio de marcadores microsatélites. MSc Thesis, Universidade Federal Rural da Amazônia, Manaus.
- Pena DN, Moura EF, Rodrigues SM, Oliveira MSP, Sanches JP, Moura MF. 2020. Molecular characterization of a germplasm bank of *Platonia insignis* Mart.: a fruit tree. *Genetic Resources and Crop Evolution* 67: 411-420.
- Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG. 2005. Comparative organisation of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* 14: 689-701.
- Phumichai C, Phumichai T, Wongkaew A. 2015. Novel chloroplast microsatellite (cpSSR) markers for genetic diversity assessment of cultivated and wild *Hevea* rubber. *Plant Molecular Biology Reporter* 33: 1486-1498.
- Pivello VR. 2011. The use of fire in the cerrado and amazonian rainforests of Brazil: past and present. *Fire Ecology* 7: 24-39.
- Pontes LCG, Moura EF, Moura MF, Rodrigues SM, Oliveira MSP, Carvalho JEU, Therrier J. 2017. Molecular characterization of progenies of bacurizeiro (*Platonia insignis*) from Marajó Island, northeastern Amazon. *Acta Amazonica* 47: 293-300.
- Provan J, Powell W, Hollingsworth PM. 2001. Chloroplast microsatellites: New tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution* 16: 142-147.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Qiu Y, Liu Y, Kang M, Yi G, Huang H. 2013. Spatial and temporal population genetic variation and structure of *Nothotsuga longibracteata* (Pinaceae), a relic conifer species endemic to subtropical China. *Genetics and Molecular Biology* 36: 598-607.
- R Development Core Team. 2018. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org>. 21 Sep. 2021.
- Rohlf FJ. 2009. NTSYS-pc: numerical taxonomy and multivariate analysis system, version 2.2 (software). New York, Stony Brook.
- Santos EM, Monteiro Neto CAS, Nascimento CAJ, Muniz FH, Barros JRS. 2017. Genetic and population diversity of bacuri (*Platonia insignis* Mart.) in Chapada Limpa extractive reserve, Maranhão State, Brazil. *African Journal of Biotechnology* 16: 2317-2325.
- Silva RG, Chaves MCL, Arnhold E, Cruz CD. 2009. Repetibilidade e correlações fenotípicas de caracteres do fruto de bacuri no Estado do Maranhão. *Acta Scientiarum* 31: 587-591.
- Souares JDR, Dias GMG, Rodrigues FA, Pasqual M, Castro EM, Pio R. 2013. Description of *Platonia insignis*, a fruitful Amazon, by histochemistry test, leaf anatomy and flow cytometry. *Revista de Ciências Agrárias* 56: 37-43.
- Softgenetics. 2021. GeneMarker: the biologist friendly software. <https://softgenetics.com/GeneMarker.php>. 21 Set. 2021.
- Souza IGB. 2011. Caracterização morfológica e molecular do bacurizeiro (*Platonia insignis* Mart.). MSc Thesis, Universidade Federal do Piauí, Teresina.
- Souza IGB, Souza VAB, Lima PSC. 2013. Molecular characterization of *Platonia insignis* Mart. (“Bacurizeiro”) using inter simple sequence repeat (ISSR) markers. *Molecular Biology Reports* 40: 3835-3845.
- Souza VAB, Vasconcelos LFL, Araújo ECE. 2007. Recursos genéticos do bacurizeiro na região Meio-Norte do Brasil. In: Lima MC. (ed.) Bacuri: agrobiodiversidade. São Luis, Instituto Interamericano de Cooperação para a Agricultura.
- Souza VAB, Vasconcelos LFL, Araújo ECE, Alves RE. 2000. O bacurizeiro (*Platonia insignis* Mart.). Jaboticabal, Fundação de Apoio a Pesquisa, Ensino e Extensão - Funep.
- Takrouni MM, Boussaid M. 2010. Genetic diversity and population's structure in Tunisian strawberry tree (*Arbutus unedo* L.). *Scientia Horticulturae* 126: 330-337.
- Vieira MLC, Santini L, Diniz AL, Munhoz CF. 2016. Microsatellite markers: What they mean and why they are so useful. *Genetics and Molecular Biology* 39: 312-328.
- Villachica H, Carvalho JEU, Müller CH, Diaz SC, Almanza M. 1996. Frutales y hortalizas promissórias de La Amazonia. Lima, Tratado de Cooperación Amazônica.
- Weising K, Gardner RCA. 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* 42: 9-19.
- Whitlock R, Hipperson H, Thompson DBA, Butlin RK, Burke T. 2016. Consequences of in-situ strategies for the conservation of plant genetic diversity. *Biological Conservation* 203: 134-142.

