

Genetic structure analysis of *Mauritia flexuosa* natural population from the Lençóis Maranhenses region using microsatellite markers

Domingos Augusto Mendonça Castro¹, Tatiana Santos Costa², Ariane Santos Cardoso³, Helaine Christine Cancela Ramos⁴, Jorge Alberto López³, Leandro Eugenio Cardamone Diniz^{2*}

¹Instituto Federal de Educação Ciência e Tecnologia do Maranhão, Campus Barreirinhas, Rod. MA-225, km 04, Povoado Santa Cruz – 65590-000 – Barreirinhas, MA – Brasil.

²Embrapa Tabuleiros Costeiros, Aracaju, Av. Beira Mar, 3250 – 49025-040 – Aracaju, SE – Brasil.

³Universidade Tiradentes/Instituto de Tecnologia e Pesquisa, Av. Murilo Dantas, 300 – 49032-490 – Aracaju, SE – Brasil.

⁴Universidade Estadual do Norte Fluminense Darcy Ribeiro, Av. Alberto Lamego, 2000 – 28013-602 – Campos dos Goytacazes, RJ – Brasil.

*Corresponding author <leandro.diniz@embrapa.br>

Edited by: Leonardo Oliveira Medici

Received April 29, 2020

Accepted July 13, 2020

ABSTRACT: Buriti (*Mauritia flexuosa* L.f.), a palm tree native to South America and widely distributed in Brazil, displays significant ecological, economic and biotechnological importance. However, the disorderly extractivism and environmental degradation of its endemic areas are leading to reductions in, and/or extinction of the buriti palm tree, causing ecological imbalance with significant economic losses for rural communities and genetic diversity. Consequently, populational genetic diversity studies have become relevant as a strategy for conserving the species. Therefore, this study evaluated the genetic structure of 10 populations from the Lençóis Maranhenses region, in the state of Maranhão, Brazil, using microsatellite markers. Results indicated that eight pairs displayed a high level of polymorphism in the populations evaluated (98.5 %). The genetic diversity estimations allowed for identifying 220 alleles (average of 9.5 alleles/loci), and the heterozygosity averages observed (Ho) were lower than the heterozygosity expected (He) in the population (0.16 and 0.64) and loci levels (0.15 and 0.65), respectively. The Shannon Index (I) mean value of 1.36 indicated high diversity and genotypic richness in the populations evaluated, while the population index (FST) indicated a low value (0.05) and the fixation index pertaining to individuals indicated a high value (FIS = 0.79) exhibiting a moderate population distribution structure, and pointing to greater diversity between individuals. Based on these results, populations denominated as PA and PB presented a high genetic similarity (0.219), while populations denominated as PF and PJ exhibited more distant genetic characteristics (0.519). These results can be correlated based on prioritization of conservation of this non-domesticated species, influenced by environmental characteristics, suggesting that the genetic diversity found should be conserved in a germplasm bank, and subsequently exploited in breeding programs.

Keywords: Buriti, SSR, preservation, diversity, variability

Introduction

Buriti (*Mauritia flexuosa* L.f.) is an undomesticated palm tree widely spread across the Amazon Forest and Brazilian cerrado, playing a relevant economic and ecological role in these regions (Virapongse et al., 2017). Its fruit is used extensively in the production of candy and ice cream, while the oil is used in cooking and serves as raw material for the cosmetics industry (Freire et al., 2016; Milanez et al., 2016; Speranza et al., 2016). However, the agrobusiness expansion, unsustainable extractivism and tourism ventures are leading to a decrease in buriti populations, promoting an environmental imbalance, and affecting rural family income as well as losses in the populations' genetic diversity within this species (Ritter et al., 2017; Virapongse et al., 2017; Saura et al., 2018).

Genetic studies of populations, habitat fragmentation, and genetic diversity based on molecular markers reveal a loss of genetic diversity in native plant species both in the Amazon region (Gomes et al., 2011; Federman et al., 2012; 2014) and in the state of Mato Grosso (Rossi et al., 2014).

Therefore, data on *M. flexuosa* from different regions are relevant due to their habitat fragmentation,

natural isolation and genetic diversity of populations, considering that it is an undomesticated species with high agronomic and biotechnological potential (Freire et al., 2016; Milanez et al., 2016).

In this context, microsatellite markers (SSR – Simple Sequence Repeat) allow for efficient identification of polymorphisms to map genetically and populationally species of economic, ecological and biotechnological interest (Soldati et al., 2014; Neri et al., 2015).

In the Lençóis Maranhenses region, buriti is an important economic input as a source of income for rural communities due to the production of fiber and extractivism of fruit. However, this region is subject to significant anthropogenic action promoting environmental degradation that impacts the sustainability of its genetic diversity (Mendes et al., 2017). Based on the ecological and economic potential of *M. flexuosa*, studies are conducted as formulations of strategies for environmental preservation, genetic conservation, and sustainable use of this species. Thus, this study aimed to analyze the diversity and genetic structure of populations of *M. flexuosa* from the Lençóis Maranhenses (MA, Brazil) region using microsatellite markers.

Materials and Methods

Plant materials

One hundred plants from ten *M. flexuosa* populations were selected from the Lençóis Maranhenses region (Figure 1). Each population was distributed in an area of approximately 40 km² and their geographic coordinates are presented in Table 1.

DNA extraction

Genomic DNA was extracted from 200 mg of young leaves macerated in liquid nitrogen, using a modified CTAB

method described by Romano and Brasileiro (2003). DNA concentration was estimated by spectrophotometry and its integrity analyzed by electrophoresis in 1 % (w/v) agarose gel stained with ethidium bromide (Sambrook et al., 1989).

DNA amplification by PCR-SSR

The genetic diversity of *M. flexuosa* was analyzed through 13 pairs of SSR markers described by Federman et al. (2012), of which eight pairs were selected and used in intra- and interpopulation analyses. Each amplification reaction was performed under the conditions described by Federman et al. (2012) with adaptations suggested by the manufacturer's instructions. Briefly, a reactional

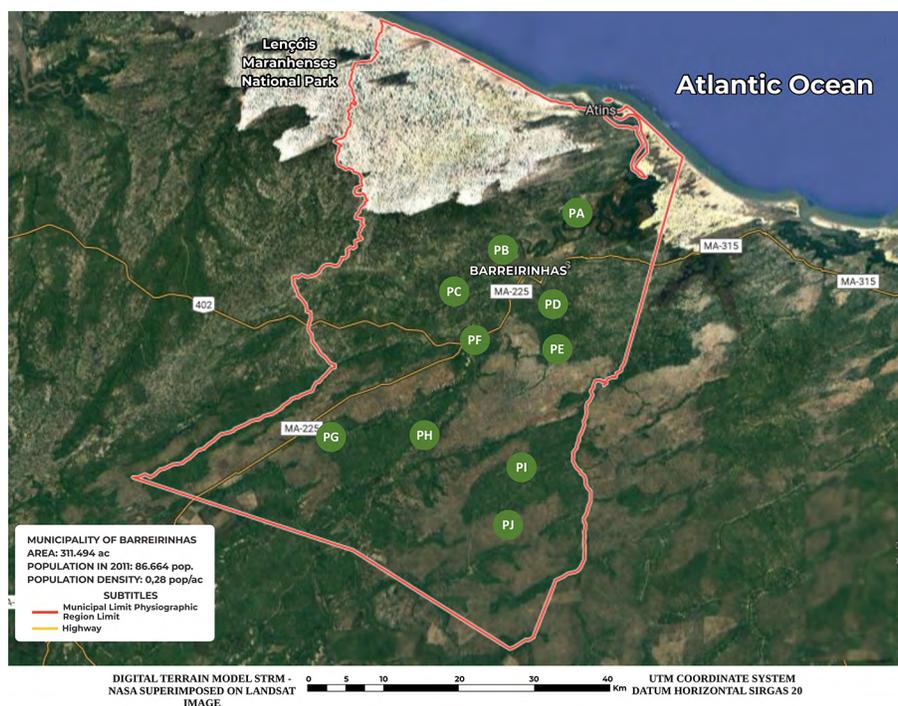


Figure 1 – Natural populations of *M. flexuosa* collected at Lençóis Maranhenses National Park, in the state of Maranhão, Brazil, indicating collected areas (adapted from CPRM – Geological Survey of Brazil, 2020). PA to PJ = populations studied.

Table 1 – Ten *Mauritia flexuosa* natural populations collected at Lençóis Maranhenses National Park, Brazil, indicating geographic coordinates. PA to PJ = populations studied.

Pop.	Communities	Latitude	Longitude	Altitude (m)	Area (km ²)
PA	Boa Vista, São Domingos, Mangaba, Vila Regina, Teles, Ilha, Ilha do Meio, Laranjeira	2°41'26" S	42°46'11" W	23	37.50
PB	Tapera, Bom Futuro, Camú, Baixinha, Santa Cruz, Vale Porto, Santa Rita, Amapá	2°43'23" S	42°50'38" W	14	56.25
PC	Bonito, Alto do Bonito, Sobradinho, Baixão do Júlio, Baixão dos Paulinos, Baixão dos Romualdo	2°46'13" S	42°55'33" W	17	31.25
PD	Ponta do Buriti, Forquilha, Estreito, Passagem do Canto	2°47'51" S	42°47'02" W	18	37.50
PE	Olho d'água dos Pereiras, Tamboril, Olho d'água, Bacabal, Vassoura, Fazendinha	2°52'13" S	42°46'40" W	26	37.50
PF	Buriti, Bandeira, Gonçalo	2°50'55" S	42°53'55" W	21	37.50
PG	Parque Náutico, Anajazal, Fome, Barra	2°57'34" S	43°04'22" W	99	37.50
PH	Jucaral dos canaveiras, Varas, Boi, Pequizeiro	2°58'22" S	42°58'23" W	59	37.50
PI	Palmeiras dos Reis, Gira Mundo, Casimiro, Cangote	3°00'25" S	42°51'15" W	74	37.50
PJ	São Roque, Mamede, Onça dos Peneiros	3°04'28" S	42°50'23" W	93	37.50

Pop. = Populations

volume of 20 μL containing: 1X of buffer sample (10X); 0.2 mM of each dNTP (2.5 mM); 1.5 mM of MgCl_2 (50 mM); 0.5 μM of primer (25 pmol); 2.5 units of Taq DNA Polymerase (5 U L^{-1}); approximately 30 ng of mold DNA, and Milli-Q was also used. Amplification followed in a thermocycler under conditions proposed by Federman et al. (2012). Amplicons were visualized in 1 % agarose gel stained with ethidium bromide (Sambrook et al., 1989) before submitting them to capillary electrophoresis using a dsDNA 905 Kit in a Fragment Analyzer System for fragment visualization.

Data Analysis

The electrophoresis profile of each SSR gel was transformed in a data matrix, elaborated from observed pattern bands, according to the peaks observed in the capillary electrophoresis. Data were used to estimate genetic diversity. For cluster analysis, genetic similarity assessment (Sgij) and Principal Coordinates Analysis (PCoA) were performed using the GenAlex software program v. 6.5 (Peakall and Smouse, 2012). The individual cluster, the correlation (r) and stress (E) values, the Polymorphic Information Content (PIC), the Heterozygosity value and the Shannon index were estimated by the UPGMA method, using the GENES software package (Cruz, 2008). The Molecular Variance Analysis (AMOVA) was calculated by the Euclidian distance matrix (EDM) to estimate the occurrence of inter- and intrapopulation variance composing two hierarchy levels, performed by GenAlex v. 6.5 (Peakall and Smouse, 2012).

For genetic structure analysis, the Bayesian method was used as applied by the STRUCTURE software program (Pritchard et al., 2000). Multi-loci genetic data were used to investigate the population and consider the presence of different populations and estimate their allele frequency. The Monte Carlo chain (MCMC) algorithm was implemented to verify the presence of species structure to allocate individuals into clusters according to their genotypes at multiple loci, and the logarithm $[\ln \text{Pr} (X|K)]$ was applied to calculate genotype (X) probability according to the number of clusters (K) subjacent the data under study (Ellegren and Galtier, 2016). This evaluation suggested the posterior probability of K being based on genotype to estimate the number of populations with a preserved structure (K), considering the most precise K for which $[\ln \text{Pr} (X|K)]$ is the maximum (Manel et al., 2005).

Results

It was possible to identify 220 alleles with 98.5 % of polymorphism, grouped into 23 loci exhibiting an average of 9.5 allele/loci. The resultant reliability was verified through stress values (0.044) and cophenetic correlation (0.985), whose stability had already been reached with amplification of 15 loci ($r = 0.9$) referring to the analysis of all *M. flexuosa* individuals (Figure 2).

The evaluation of the efficacy of the markers was also observed through the average value corresponding to the number of effective alleles per polymorphic locus (N_e). This value was 3.73, ranging from 3.23 (Mf 17F) to 4.22 (Mf 22F) for H_e values, with a variation between 0.55 (marker Mf 17F) and 0.72 (marker Mf 28F), with a mean of 0.65, and an H_o value between 0.05 (Mf 30F marker) and 0.22 (Mf 28F marker) (Table 2).

All populations presented an 'I' higher than 1.0, varying from 1.24 (PG) to 1.48 (PJ) (Table 3). The results demonstrated great diversity or genotypic variety within each population analyzed. The PB, PH and PJ populations exhibited a number of private alleles. F values were positive and different from zero in all populations, where the population with the lowest genetic variability was PF ($F = 0.78$, $H_o = 0.15$ and $\text{PIC} = 0.50$) and the population with the highest genetic variability was PB ($F = 0.71$, $H_o = 0.21$ and $\text{PIC} = 0.64$).

Using the statistical F coefficient (Wright, 1951), made it possible to estimate the allele frequency inter- and intra population in the ten *M. flexuosa* populations studied. The Population Fixation Index (FIT) was 0.800 ($p < 0.0001$): for populations (FST) 0.052 ($p < 0.0001$) and for individuals (FIS) 0.789 ($p < 0.0001$), indicating a high level of genetic differentiation in individuals and

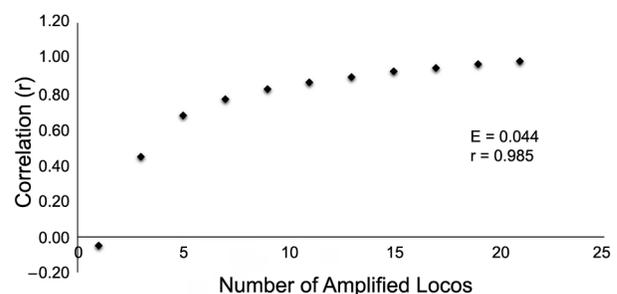


Figure 2 – Cophenetic correlation between the number of polymorphic fragments using SSR markers in ten natural populations of *M. flexuosa* located in the Lençóis Maranhenses region, Brazil. E = Stress value; r = Correlation.

Table 2 – Eight polymorphic SSR markers available for *M. flexuosa* used for the genetic diversity study in ten natural Buriti populations located in the Lençóis Maranhenses region.

SSR Markers	N	N_e	NL	H_e	H_o
Mf 17F	9.90	3.23	04	0.55	0.20
Mf 19F	9.33	3.53	03	0.57	0.18
Mf 22F	9.93	4.22	02	0.71	0.18
Mf 25F	10.00	3.80	03	0.70	0.15
Mf 28F	9.40	4.00	02	0.72	0.22
Mf 30F	10.00	3.51	02	0.68	0.05
Mf 31F	9.67	3.54	03	0.58	0.16
Mf 34F	10.00	4.05	04	0.71	0.12

N = Population size; N_e = Number of Effective Alleles; NL = Number of Loci; H_e = Expected Heterozygosity; H_o = Observed Heterozygosity.

a low level in populations (Yeh, 1999). Based on this coefficient, it was possible to obtain the AMOVA (Table 4) to estimate the highest genetic variety in individuals (95 %) and a value of 5 % for populations.

As regards the matrix generated from the Nei genetic dissimilarity coefficient (1978) (Table 5), it was possible to evaluate the genetic diversity in the ten populations under analysis. Results indicated that the PA and PB populations displayed more genetic similarity (0.219), while this similarity was lower in the PF and PJ populations (0.519). This dissimilarity coefficient also allowed for the clustering of these *M. flexuosa*

Table 3 – Estimated genetic variability between natural populations of *M. flexuosa* from the Lençóis Maranhenses region. PA to PJ = populations studied.

Pop.	I	He	Ho	F	PIC
PA	1.38	0.65	0.19	0.72	0.58
PB	1.43	0.68	0.21	0.71	0.64
PC	1.30	0.62	0.16	0.74	0.54
PD	1.36	0.65	0.17	0.75	0.55
PE	1.35	0.65	0.11	0.85	0.55
PF	1.34	0.65	0.15	0.78	0.50
PG	1.24	0.60	0.16	0.75	0.52
PH	1.38	0.66	0.16	0.79	0.53
PI	1.39	0.67	0.10	0.86	0.51
PJ	1.48	0.69	0.14	0.80	0.56

Pop. = Populations studied; I = Shannon Index; He = Expected Heterozygosity; Ho = Observed Heterozygosity; F = Fixation index; PIC = Polymorphism Information Content.

Table 4 – Analysis of Molecular Variance (AMOVA) from the SSR data carried out on 100 species from ten natural populations of *M. flexuosa* in the Lençóis Maranhenses area.

Variation source	dF	SS	AS	CV	V (%)	p
Individuals	190	1523.15	16.72	8.36	95	< 0.0001
Populations	9	216.45	24.05	0.45	5	< 0.0001
Total	199	1739.60	40.77	8.81	100	

dF = Degrees of freedom; SS = Square Sum; AS = Average Square; CV = Coefficient of Variance; V = Variance Percentage; p = Probability of presenting a component variation higher than those seen randomly. Probabilities were estimated from 1000 random permutations.

Table 5 – Nei genetic dissimilarity matrix (1978). PA to PJ = populations studied.

PA	PB	PC	PD	PE	PF	PG	PH	PI	PJ	
0.000										PA
0.219	0.000									PB
0.257	0.277	0.000								PC
0.254	0.256	0.276	0.000							PD
0.352	0.339	0.331	0.266	0.000						PE
0.314	0.396	0.398	0.309	0.330	0.000					PF
0.304	0.406	0.330	0.283	0.301	0.336	0.000				PG
0.323	0.407	0.407	0.265	0.420	0.378	0.242	0.000			PH
0.366	0.423	0.430	0.421	0.434	0.449	0.407	0.345	0.000		PI
0.443	0.463	0.499	0.477	0.498	0.519	0.392	0.419	0.369	0.000	PJ

populations into four distinct categories based on the Principal Coordinates Analysis on genetic similarity (Figure 3).

The results of the analysis of clustering populations according to the UPGMA method showed the formation of five main groups, which, in most cases, are in agreement with the dendrogram (Figure 4) and the distance matrix obtained. Group 5 exhibited the largest number of populations (PA, PB, PC and PD), and the other

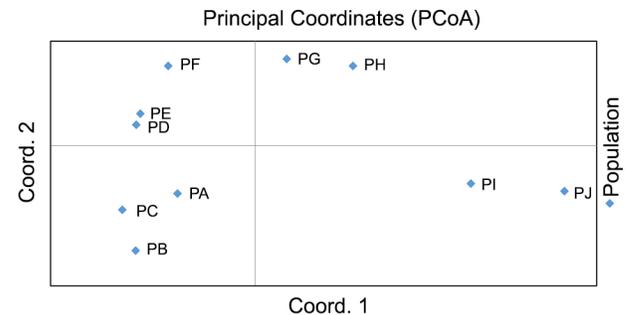


Figure 3 – Principal Coordinates Analysis (PCoA) of the genetic diversity of the ten natural populations of *M. flexuosa*, estimated on the basis of Nei genetic dissimilarity. PA to PJ = populations studied.

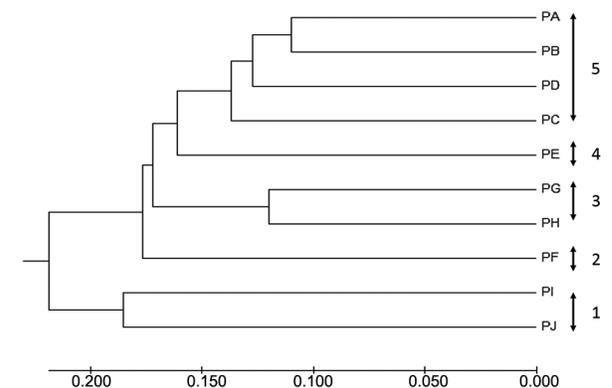


Figure 4 – Genetic dissimilarity dendrogram obtained by the UPGMA method, using the Nei genetic distance matrix, illustrating the genetic relationships between the ten natural Buriti population in the Lençóis Maranhenses region. PA to PJ = populations studied.

groups were allocated only the smaller populations of *M. flexuosa* (Figure 4). Thus, the populations of this species were easily separated using SSR markers, indicating genetic differentiation and correlations between the natural populations of this plant. The results show the formation of five main groups and, in most cases, are in agreement with the dendrogram (Figure 4) and the distance matrix obtained.

On the other hand, the cluster Bayesian analysis of the populations' genetics allowed for visualizing the formation of distinct groups pertaining to the allele distribution in each individual (Figure 5). The separation into four population groups was further supported by the AMOVA results, indicating considerable variability between individuals (Table 4).

Thus, it emerged that the PF, PH and PI populations displayed more homogeneity based on the individual allele distribution (with a higher predominance of a single colored bar, namely, one allele). The PI and PJ populations were analogous to each other in terms of allele distribution (with predominance of red bar) as well as the PH and PI populations (which displayed a predominance of blue bar). The remaining populations exhibited wide-ranging variability in their allele, attributable to a color mixture of bars. Thus, in these populations individuals presented a higher allele distribution.

Data from the analysis of the 100 individuals of *M. flexuosa* were used to elaborate a dissimilarity matrix based on the weighted average index, with a minimum value of 0.296 and maximum of 0.984 for the pairs formed between individuals PD7 × PH7 and PC10 × PJ7, respectively, from the different population groups (Table 6). From the genetic distance matrix, calculated using the GenAlix software program, a coordinate analysis graphic was elaborated to evaluate genetic diversity among individuals (Figure 6).

As regards the analysis by the UPGMA clustering method for all individuals were separated by similarity into five different groups (Figure 7). Thus, the individual PI10 and PI6 from group 1; PH7 and PD7 in group 2; PC4 and PH8 from group 3; PB1 and PF4 in group 4; and PC1 with PD3, and PF3 with PF7 from group 5 displayed more similarity between themselves within the group.

Discussion

Current knowledge regarding the genetic structure of the natural populations of *M. flexuosa* distributed in different regions of Brazil is based on analyses of certain molecular markers: AFLP (Gomes et al., 2011), ISSR (Rossi et al., 2014) and SSR (Federman et al., 2012; Menezes et al., 2012). However, studies of this species are still scarce, with no evaluation of the genetic structure regarding the distribution of natural buriti populations from Lençóis Maranhenses. Therefore, this study is the first one related to the evaluation of diversity, genetic structure and conservation status of *M. flexuosa* in this region. The relevance of this study is due to the marked environmental degradation of this region caused by advances in road infrastructure and the expansion of agribusiness and urban borders.

Specific SSR markers for this species were developed and applied to populations in other areas (Federman et al., 2012; Menezes et al., 2012; Oliveira et

Table 6 – Individual pairs with highest and lowest dissimilarity extreme values (1 – s) based on the weighted index, calculated from eight SSE markers in 100 *M. flexuosa* individuals. PA to PJ = populations studied.

Highest dissimilarity			Lowest dissimilarity		
Populations	Individuals	1 – s	Populations	Individuals	1 – s
PC × PJ	10 × 7	0.984	PD × PH	7 × 7	0.296
PD × PE	4 × 6	0.980	PB × PC	9 × 6	0.337
PE × PF	6 × 6	0.980	PD × PD	1 × 2	0.343
PE × PF	6 × 7	0.980	PI × PI	2 × 5	0.347
PE × PH	6 × 1	0.980	PD × PG	7 × 1	0.372
PE × PH	6 × 2	0.980	PC × PC	1 × 2	0.382
PD × PE	10 × 6	0.964	PF × PF	1 × 2	0.382
PE × PG	6 × 5	0.964	PD × PH	1 × 10	0.385
PB × PH	3 × 3	0.961	PC × PD	1 × 3	0.388
PE × PF	6 × 8	0.961	PA × PC	5 × 2	0.417
PE × PF	6 × 10	0.961	PC × PE	1 × 2	0.423
PE × PH	6 × 4	0.961	PC × PH	4 × 8	0.423
PF × PI	9 × 7	0.961	PA × PF	5 × 4	0.426
PF × PI	10 × 7	0.961	PA × PG	5 × 3	0.429
PA × PJ	7 × 7	0.955	PA × PH	9 × 8	0.429

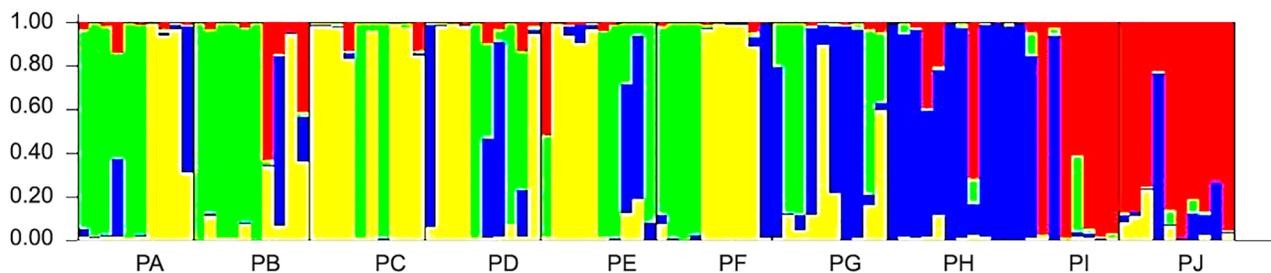


Figure 5 – Genetic structure analysis of the natural populations of *M. flexuosa* according to the Bayesian analysis of groupings using the STRUCTURE program for K = 4 with samples separated by population. Colors represent the four groups identified and the length of each color represents the genomic composition of each individual. PA to PJ = populations studied.

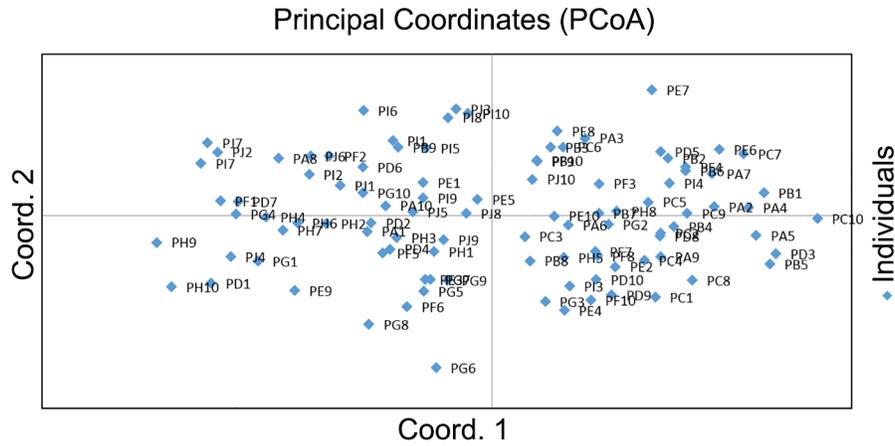


Figure 6 – Principal Coordinates Analysis (PCoA) of the genetic diversity of the ten natural populations of *M. flexuosa*, estimated on the basis of the average Euclidian distance. PA to PJ = populations studied.

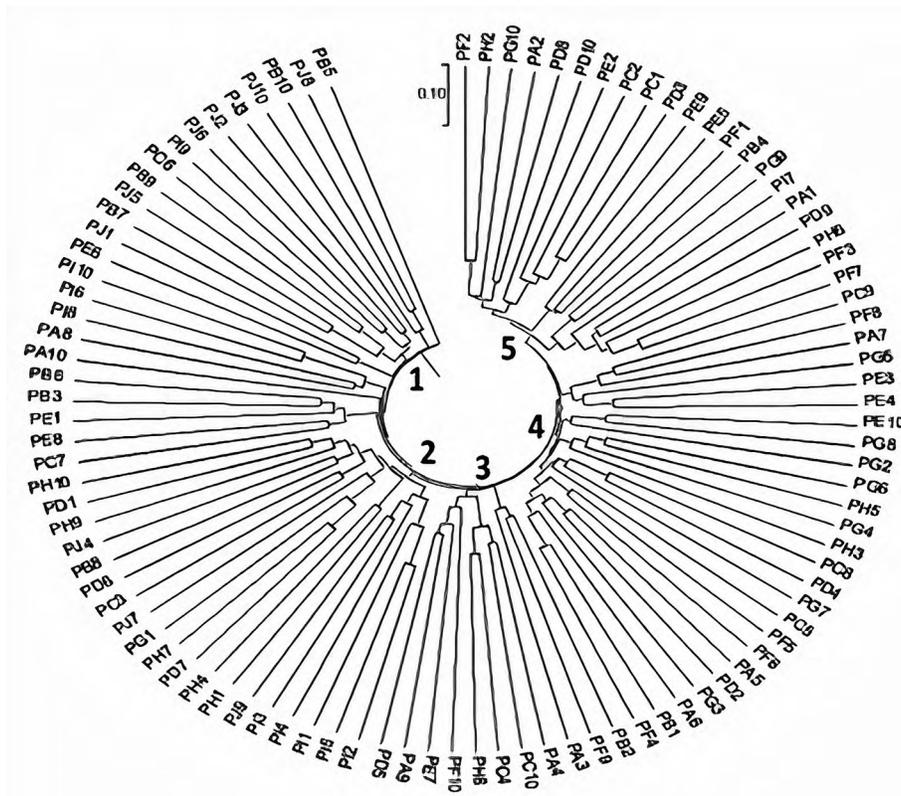


Figure 7 – Distribution of individuals in a Dendrogram of *M. flexuosa* considering the genetic dissimilarity matrix (weighted average index). PA to PJ = populations studied; 1 to 5 = Groups.

al., 2012; Federman et al., 2014). As regards this study, the specific markers used, as described by Federman et al. (2012), as well as the reliability of the results obtained from the co-phenotype correlation analyses ($E = 0.044$, $r = 0.985$), suggested optimal precision in the estimates due to the $E < 0.05$ (Kruskal, 1964). Although only eight SSR markers were used, the allele mean value (9.5 allele

per loci) was higher than those reported in the literature for the same species compared to Brazilian native populations (Federman et al., 2012) and other countries (Federman et al., 2014). These data indicate high rates of genetic diversity correlated to the high content of genetic information obtained from an average number of alleles. Nevertheless, there are also reports in the literature

which indicate values lower than those described in this work, using other SSR markers (Menezes et al., 2012).

N values higher than the N_e shown in Table 2 may indicate an unequal distribution of allele frequency among individuals, a rare allele presence, which occurs in frequencies < 0.05 in a population, as well as the presence of common allele with frequencies ≥ 0.05 (Whitlock et al., 2016). The identification of rare alleles is useful information in genetic conservation studies for singling populations that deserve special management (Kalinowski, 2004). However, for a better inference of the rare alleles in *M. flexuosa*, sample collections using a larger population sample are required to observe if there is variation in this allele frequency since this event is closely related to sample size (Marshall and Brown, 1975).

The values of H_e and H_o help in the genetic diversity evaluation since these parameters measure the homozygote and heterozygote allele proportion in analyzed loci. The values for H_e were higher than H_o in all the markers analyzed, indicating a high rate of homozygous alleles between the loci. Results with low heterozygote rates using SSR markers are common in studies on this species (Menezes et al., 2012; Federman et al., 2014), and also studies on other species (Ramos et al., 2012). However, reports have presented contradictory values, showing a slightly lower average for H_e than for H_o (Gomes et al., 2011). Higher H_e mean values than those of H_o indicate moderate genetic diversity in *M. flexuosa* populations. This is due to the excess of homozygotes as a result of marked endogamy in the reproduction of these populations. This inbreeding level may be associated with pollination, seed dispersal, inter- and intra-specific competition as well as the heterogeneity of the environment of each population selected, which directly affects the spatial distribution of individuals and, consequently, their population's genetic diversity (Freeland et al., 2011). These data are relevant to *M. flexuosa* exploration and domestication processes since this is a species exposed to ecological risk (Pickersgill, 2007).

As regards the PIC values (Table 3), the PB and PA populations exhibited more polymorphic information, indicating their high genetic variability. This parameter is applied to assessments of genetic polymorphism, since low values display low level genetic variation per loci, while PIC values ranging from 0.25 to 0.50 are considered merely informative (Botstein et al., 1980). This index showed that eight microsatellite markers efficiently evaluated genetic diversity in the *M. flexuosa* population. Thus, the PIC mean value of 0.548 denotes high variability in the ten populations analyzed, mainly in the PB population, displaying a higher polymorphism percentage (0.64). A number of reports have estimates of lower PIC values using SSR markers (Alves et al., 2013) or similar to those described in this study (Loiola et al., 2016), indicating the effectiveness of these markers for diversity analysis.

Furthermore, the Shannon Index (I) is also applied to determine diversity within a population to indicate genotypic richness through the certainty value of the genetic similarity estimate between individuals in each population (Holcomb et al., 1977). Consequently, I values ranging from 1.0 to 1.5 indicate moderate genetic diversity in individuals from a population, while a variation between 1.0 – 1.5 reveals high diversity (Shannon and Weaver, 1963). Therefore, I values ranging significantly from 1.24 (PG) to 1.48 (PJ) were observed in all populations analyzed (Table 3), which indicates moderate diversity or genotypic richness of the *M. flexuosa* populations from the Lençóis Maranhenses region. The I value determined in this study is higher compared to other studies using ISSR markers in natural populations of this species (Rossi et al., 2014). Thus, data indicate that natural populations from the Lençóis Maranhenses region exhibit more genotypic richness compared to those from the state of Mato Grosso. This result may be related to the extension of the Lençóis Maranhenses region and, consequently, the gene flow in the individuals of each population. Furthermore, the SSR marker specificity may contribute to careful evaluation in terms of the populations' genetic structures. Therefore, the results of this study indicate efficient discrimination in the genetic structure of the populations under analysis.

As regards the allele fixation index (F), data showed a decrease in genetic variability and an increase in homozygosity in the different subpopulations. Moreover, factors influence the genetic diversity level, affecting the allele frequency by randomized crossover as a result of evolutionary processes, such as effective size, mutations and natural selection. Thus, the estimate of the correlation of inter- and intrapopulation allele frequencies of the ten *M. flexuosa* populations analysed presented values of 0.052 and 0.789, respectively, using the statistical F coefficient (Wright, 1951). This result provided high F values (fixation index) for each population evaluated (average 0.77). This suggested a deviation in the Hardy-Weinberg balance equilibrium, due to homozygote excess, probably as a result of parental crossbreeding or auto fecundation among the individuals analyzed.

The endogamy coefficient (F) estimated for the *M. flexuosa* populations from Lençóis Maranhenses differs from the populations reported in other genetic studies. These studies evaluated the genetic variation of populations in the Amazonian river regions, which has impacted the dispersion and distribution of the species (Melo et al., 2018; Sander et al., 2018). Nonetheless, a comparison of climatic conditions and the discontinuity of the Lençóis Maranhenses habitat region showed marked differences as it is, geographically, a very isolated place and under intense anthropic degradation, factors which are relevant to the *M. flexuosa* population.

As for the *M. flexuosa* population assessed, the tendency to endogamy may have been the result of crossbreeding between parental individuals (full-sibs or half-sibs), since the auto fecundation possibility

was discarded, due to the cross-fertilization method of this species (Storti, 1993). Therefore, the high endogamy rate among individuals may be related to the geographic isolation of plant species, which decreases interpopulation genic flow. This geographical isolation is mainly due to extension of the dunes with heights of up to 60 m (Figure 1), and the soil quality, which hinders pollen and seed dispersion and seed germination.

The moderate intrapopulation genotypic diversity found can indicate a lower gene flow among the individuals. This feature must be carefully analyzed for conservation strategies since individuals and/or populations with divergent alleles are required for future crossing to increase genetic diversity among themselves. Therefore, it is a requirement that Germplasm Banks store accessions with lower F values (Alves et al., 2013; Loiola et al., 2016). Another parameter analyzed was the total fixation index with a low value of 0.052, considering all populations (FST) which is a common feature in natural *M. flexuosa* populations (Federman et al., 2012; Menezes et al., 2012; Federman et al., 2014; Rossi et al., 2014). This gene flow limitation may be associated with population environment conditions, such as swampy or dry and sandy regions, which can act as a flow barrier.

As regards genetic variability, the analysis by AMOVA, as depicted in Table 4, showed that there was more variability between individuals (95 %) than between populations (5 %). These results are in agreement with other studies of *M. flexuosa* in which the highest genetic diversity occurrence was observed at the intrapopulation level (Gomes et al., 2011; Rossi et al., 2014). These results may be related to the type of species analyzed since tree species with mixed breeding systems and efficient pollen and seed dispersal mechanisms exhibit greater intrapopulation genetic variation (Ellegren and Galtier, 2016). This is the result of a long distance allelic flow, which decreases the variation between populations, increasing the intrapopulation variation (Loveless and Hanrick, 1984).

Comparing the PCoA results to the geographic distance map (Figures 1 and 6, respectively), a correspondence was observed between the groups formed, suggesting a correlation between both the genetic and the geographic diversity of populations. The dendrogram also showed this correlation in the genetic similarity recorded on application of the Nei coefficient. Events correlating geographic and genetic distances reported previously indicated that low levels of genetic similarity among populations may be due to increased spatial fragmentation (Lemos et al., 2015). The spatial fragmentation in the Lençóis Maranhenses region is due to natural factors (extensive sandy bank areas) (Figure 1), and by anthropogenic actions (e.g. road construction, urban and agricultural advances) that hinder gene flow.

Genetic diversity analysis among individuals using the genetic dissimilarity matrix (Table 5) also allowed for estimating the presence of pairs of individuals with higher and lower genetic similarities. Based on this, no

crossing was determined between pairs of individuals belonging to the same population with high genetic diversity. This reinforces the connection between this species' population with the dispersal system and its geographic distribution.

Data from genetic diversity analysis between individuals are consistent with those observed with genetic structure analysis (Figure 5) and the principal coordinate analyses (Figure 6). These results emphasize the genetic diversity of the individuals analyzed and the formation of four genetic clusters associated with different genetics. Thus, geographically close populations tend to share alleles at the same rate. As previously described, this may be a consequence of the pollen dispersion of this species' population in the geographical environment analyzed and the fragmentation process of the Lençóis Maranhenses region, which may be affecting reproductive patterns and increasing the inbreeding rate as well as the loss of this species' reproductive capacity. Overall, the results evidenced the need to conserve the natural *M. flexuosa* populations in this region to avoid losing their genetic variability.

Conclusions

Current results using samples of the *M. flexuosa* species from the Lençóis Maranhenses region exhibited moderate genetic diversity. Nevertheless, the low FST value in populations indicated the occurrence of a decrease in gene flow in these individuals, which was reflected in high F values. Therefore, considering the accelerated degradation of its habitat, there is urgent pressure to preserve plant genetic diversity in *ex situ* conservation collections. Thereafter, data suggest the inclusion of *M. flexuosa* population genetic material from the Lençóis Maranhenses region in a Germplasm Bank for their conservation, aiming at application in breeding programs to preserve its genetic variability.

Acknowledgments

The authors extend thanks to Silvio Gomes dos Santos (Embrapa Coastal Tablelands) for all laboratory support and to the Foundation for Research and Technological Innovation Support of the state of Sergipe (FAPITEC/SE) and the Coordination for the Improvement of Higher Education Personnel (CAPES) federal agency for the fellowship granted to the second author.

Authors' Contributions

Conceptualization: Castro, D.A.M.; López, J.A.; Diniz, L.E.C. **Data acquisition:** Castro, D.A.M.; Costa, T.S.; Cardoso, A.S.; Ramos, H.C.C. **Data analysis:** Costa, T.S.; Ramos, H.C.C.; Diniz, L.E.C. **Design of methodology:** López, J.A.; Diniz, L.E.C. **Writing and editing:** Diniz, L.E.C.; Costa, T.S.; Castro, D.A.M.; López, J.A.

References

- Alves, R.M.; Silva, C.R.S.; Silva, M.S.C.; Silva, D.C.S.; Sebbenn, A.M. 2013. Genetic diversity in Brazilian Amazon collections of cupuassu tree germplasm [*Theobroma grandiflorum* (Willd. ex Spreng.) Schum]. *Revista Brasileira de Fruticultura* 35: 818-828 (in Portuguese, with abstract in English).
- Botstein, D.; White, R.L.; Skolnick, M.H.; Davis, R.W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics* 32: 314-331.
- Companhia de Pesquisa de Recursos Minerais [CPRM]. Map and Project Integrator = Integrador de Mapas e Projetos. CPRM Serviço Geológico do Brasil, Brasília, DF, Brazil. Available at: <https://geoportal.cprm.gov.br/novointegrador/> [Accessed Aug. 12, 2020] (in Portuguese).
- Cruz, C.D. 2008. Software Genes: Computational Application in Genetics and Statistics = Programa Genes. Aplicativo Computacional em Genética e Estatística. UFV, Viçosa, MG, Brazil (in Portuguese).
- Ellegren, H.; Galtier, N. 2016. Determinants of genetic diversity. *Nature Reviews Genetics* 17: 443-434.
- Federman, S.; Hyseni, C.; Clement, W.; Caccone, A. 2012. Isolation of 13 novel highly polymorphic microsatellite loci for the Amazonian palm *Mauritia flexuosa* (Arecaceae). *Conservation Genetics Resources* 4: 355-367.
- Federman, S.; Hyseni, C.; Clement, W.; Oatham, M.P.; Caccone, A. 2014. Habitat fragmentation and the genetic structure of the Amazonian palm *Mauritia flexuosa* L.f. (Arecaceae) on the island of Trinidad. *Conservation Genetics* 15: 355-362.
- Freeland, J.R.; Kirk, H.; Petersen, S. 2011. *Molecular Ecology*. 2ed. Wiley-Blackwell, Hoboken, NJ, USA.
- Freire, J.A.P.; Barros, K.B.N.T.; Lima, L.K.F.; Martins, J.M.; Araújo, Y.C.; Oliveira, G.L.S.; Aquino, J.A.; Ferreira, P.M.P. 2016. Phytochemistry profile, nutritional properties and pharmacological activities of *Mauritia flexuosa*. *Journal of Food Science* 81: 611-622.
- Gomes, L.R.P.; Lopes, M.T.G.; Bentes, J.L.S.; Barros, W.S.; Costa Neto, P.Q.; Contim, L.A.S. 2011. Genetic diversity in natural populations of Buriti (*Mauritia flexuosa* L.f.). *Crop Breeding and Applied Biotechnology* 11: 216-223.
- Holcomb, J.; Tolbert, D.M.; Jain, K. 1977. A diversity analysis of genetic resources in rice. *Euphytica* 26: 441-449.
- Kalinowski, S.T. 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics* 5: 539-543.
- Kruskal, J.B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29: 1-27.
- Lemos, R.P.M.; D'Oliveira, C.B.; Stefenon, V.M. 2015. Genetic structure and internal gene flow in populations of *Schinus molle* (Anacardiaceae) in the Brazilian Pampa. *Tree Genetics and Genomes* 11: 75-89.
- Loiola, C.M.; Azevedo, A.O.N.; Diniz, L.E.C.; Aragão, W.M.; Azevedo, C.D.; Santos, P.H.; Ramos, H.C.C.; Pereira, M.G.; Ramos, S.R.R. 2016. Genetic relationships among tall Coconut Palm (*Cocos nucifera* L.) accessions of the International Coconut Genebank for Latin America and the Caribbean (ICG-LAC), evaluated using microsatellite markers (SSRs). *PLoS One* 11: e0151309.
- Loveless, M.D.; Hamrick, J.L. 1984. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology, Evolution and Systematics* 15: 65-95.
- Manel, S.; Gaggiotti, O.E.; Waples, R.S. 2005. Assignment methods: matching biological questions with appropriate techniques. *Trends in Ecology & Evolution* 20: 136-142.
- Marshall, D.K.; Brown, A.H.D. 1975. Optimum sampling strategies in genetic conservation. p. 53-70. In: Frankel, O.H.; Hawkes G.J., eds. *Crop genetic resources for today and tomorrow*. Cambridge University Press, Cambridge, UK.
- Melo, W.A.; Freitas, C.G.; Bacon, C.D.; Collevatti, R.G. 2018. The road to evolutionary success: insights from the demographic history of an Amazonian palm. *Heredity* 121: 183-195.
- Menezes, E.V.; Souto, W.F.S.; Ciampi, A.Y.; Azevedo, V.C.R.; Valério, H.M.; Pimenta, M.A.S. 2012. Development and characterization of DNA microsatellite primers for buriti (*Mauritia flexuosa* L.f.). *Genetics and Molecular Research* 11: 4058-4062.
- Milanez, J.T.; Neves, L.C.; Silva, P.M.C.; Bastos, V.J.; Shahab, M.; Colombo, R.C.; Roberto, S.R. 2016. Pre-harvest studies of buriti (*Mauritia flexuosa* L. f.), a Brazilian native fruit, for the characterization of ideal harvest point and ripening stages. *Scientia Horticulturae* 202: 77-82.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.
- Neri, J.; Nazareno, A.G.; Wendt, T.; Palma-Silva, C. 2015. Development and characterization of microsatellite markers for *Vriesea simplex* (Bromeliaceae) and cross-amplification in other species of *Bromeliaceae*. *Biochemical Systematics and Ecology* 58: 34-37.
- Oliveira, V.B.; Yamada, L.T.; Fagg, C.W.; Brandão, M.G.L. 2012. Native foods from Brazilian biodiversity as a source of bioactive compounds. *Food Research International* 48: 170-179.
- Peakall, R.; Smouse, P.E. 2012. GENALEX 6: genetic analysis in Excel: population genetic software for teaching and research; an update. *Bioinformatics* 28: 2537-2539.
- Pickersgill, B. 2007. Domestication of plants in the Americas: insights from Mendelian and molecular genetics. *Annals of Botany* 100: 925-940.
- Pritchard, J.; Stephens, M.; Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Ramos, S.L.; Lopes, M.T.; Batista, J.S.; Formiga, K.M.; Silva, P.P.; Saulo-Machado, A.C.; Veasey, E.A. 2012. Microsatellite loci for tucumã of Amazonas (*Astrocaryum aculeatum*) and amplification in other Arecaceae. *American Journal of Botany* 99: e508-e510.
- Ritter, C.D.; McCratec, G.; Nilsson, R.H.; Fearnside, P.M.; Palme, U.; Antonelli, A. 2017. Environmental impact assessment in Brazilian Amazonia: challenges and prospects to assess Biodiversity. *Biological Conservation* 206: 161-168.
- Romano, E.; Brasileiro, A.C.M. 1999. Plants DNA extraction = Extração de DNA de plantas. *Biocientia* 9: 40-43 (in Portuguese).
- Rossi, S.F.; Rossi, A.A.B.; Dardengo, J.F.E.; Brauwiers, L.R.; Silva, M.L.; Sebbenn, A.M. 2014. Genetic diversity in natural populations of *Mauritia flexuosa* (Arecaceae) using ISSR markers. *Scientia Forestalis* 42: 631-639.

- Sambrook, J.; Fritsch, E.F.; Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual*. 2ed. Cold Spring Harbor Laboratory Press, New York, NY, USA.
- Sander, N.L.; Pérez-Zavala, F.; Da Silva, C.J.; Arruda, J.C.; Pulido, M.T.; Barelli, M.A.A.; Rossi, A.B.; Viana, A.P.; Boechat, M.S.B.; Bacon, C.D.; Cibrián-Jaramillo, A. 2018. Rivers shape population genetic structure in *Mauritia flexuosa* (Arecaceae). *Ecology and Evolution* 8: 6589-6598.
- Saura, S.; Bertzky, B.; Bastin, L.; Battistella, L.; Mandrici, A.; Dubois, G. 2018. Protected area connectivity: shortfalls in global targets and country-level priorities. *Biological Conservation* 219: 53-67.
- Shannon, C.E.; Weaver, W. 1963. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, IL, USA.
- Soldati, M.C.; Inza, M.V.; Fornes, L.; Zelener, N. 2014. Cross transferability of SSR markers to endangered *Cedrela* species that grow in Argentinean subtropical forests, as a valuable tool for population genetic studies. *Biochemical Systematics and Ecology* 53: 8-16.
- Speranza, P.; Falcão, A.O.; Macedo, J.A.; Silva, L.H.M.; Rodrigues, A.M.C.; Macedo, G.A. 2016. Amazonian Buriti oil: chemical characterization and antioxidant potential. *Grasas Aceites* 67: 1-9.
- Storti, E.F. 1993. Floral biology of *Mauritia flexuosa* Lin. Fil. In Manaus, AM, Brazil. *Acta Amazônica* 23: 371-381.
- Virapongse, A.; Endress, B.A.; Gilmore, M.P.; Horn, C.; Romulo, C. 2017. Ecology, livelihoods, and management of the *Mauritia flexuosa* palm in South America. *Global Ecology and Conservation* 10: 70-92.
- Whitlock, R.; Hipperson, H.; Thompson, D.B.A.; Butlin, R.K.; Burke, T. 2016. Consequences of *in situ* strategies for the conservation of plant genetic diversity. *Biological Conservation* 203: 134-142.
- Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics* 15: 323-354.
- Yeh, F.C.; Yang, R.; Boyle, T. 1999. POPGENE: Microsoft Windows based freeware for population genetic analysis: version 1.31. Open Science, Burlington, MA, USA.