Species delimitation and hybrid identification of *Acrocomia aculeata* and *A. totai* by genetic population approach

Brenda Díaz[1], Maria Zucchi[2], Alessandro Alves-Pereira[1], Joaquim Azevedo-Filho[2], Mariana Sanitá[2], and Carlos Colombo[2]

[1]State University of Campinas
[2]Instituto Agronomico

October 9, 2020

Abstract

To the Neotropical genus Acrocomia (Arecaceae) is attributed eight species with a wide distribution in America. *A. aculeata* and *A. totai* are the most important species because of their high economic potential for oil production. However, there is no consensus in their classification as different taxons and their distinctiveness is particularly challenging due to morphological similarities with large plasticity of the traits. In addition, there is doubt about the occurrence of interspecific hybrids between both species. In this study, we applied a genetic population approach to assessing the genetic boundaries, diversity and to identify interspecific hybrids of A. *aculeata* and *A. totai*. Thirteen loci of simple sequence repeat (SSR) were employed to analyze twelve populations representing a wide distribution of species, from Minas Gerais, Brazil to Formosa, Argentina. Based on the Bayesian analysis (STRUCTURE and NewHybrids) and Discriminant Analysis of Principal Components (DAPC), our study supports the recognition of *A. aculeata* and *A. totai* as two species and the estimates of genetic parameters revealed more genetic diversity in *A. totai* (H_E=0.551) than in *A. aculeata* (H_E=0.466). We obtained evidence of hybridization between the species and that admixed individuals were assigned as F2 hybrids. In conclusion, this study showed the usefulness of microsatellite markers to elucidate the genetic boundaries of *A. aculeata* and *A. totai*, supporting their classification as different species and increase our knowledge about genetic diversity at the level of populations and species. The results are essentials to establish strategies for the adequate management, conservation, and domestication of both species.

Species delimitation and hybrid identification of *Acrocomia aculeata* and*A. totai* by genetic population approach

Brenda G. Díaz[1], Maria I. Zucchi[1,2], Alessandro Alves-Pereira[1], Joaquim A. Azevedo-Filho[2], Mariana Sanita[2]Carlos A. Colombo[2*]

[1] State University of Campinas UNICAMP, Institute of Biology, Campinas-SP, Brazil.

Brenda G Diaz e-mail: diaz.brenda01@gmail.com ORCID: 0000-0001-7787-852

Alessandro Alves-Pereira e-mail: ale.alves.pereira@gmail.com, ORCID: 0000-0002-3012-6355

[2] Instituto Agronomico do estado de Sao Paulo-IAC, Centro de Pesquisa de Recursos Geneticos Vegetais, Campinas-SP, Brazil.

Maira I Zucchi e-mail: mizucchi@gmail.com ORCID: 0000-0002-4863-1843

Joaquim A. Azevedo-Filho e-mail: joaquimadelino66@gmail.com

Mariana Sanita e-mail: mari.san.03@hotmail.com, ORCID: 0000-0003-2937-966X

Carlos A. Colombo e-mail carlos.colombo@sp.gov.br ORCID: 0000-0003-3269-6068

* Corresponding author:

E-mail: carlos.colombo@sp.gov.br (CAC)

Abstract

To the Neotropical genus Acrocomia (Arecaceae) is attributed eight species with a wide distribution in America. *A. aculeata* and*A. totai* are the most important species because of their high economic potential for oil production. However, there is no consensus in their classification as different taxons and their distinctiveness is particularly challenging due to morphological similarities with large plasticity of the traits. In addition, there is doubt about the occurrence of interspecific hybrids between both species. In this study, we applied a genetic population approach to assessing the genetic boundaries, diversity and to identify interspecific hybrids of *A. aculeata* and *A. totai* . Thirteen loci of simple sequence repeat (SSR) were employed to analyze twelve populations representing a wide distribution of species, from Minas Gerais, Brazil to Formosa, Argentina. Based on the Bayesian analysis (STRUCTURE and NewHybrids) and Discriminant Analysis of Principal Components (DAPC), our study supports the recognition of *A. aculeata* and *A. totai* as two species and the estimates of genetic parameters revealed more genetic diversity in *A. totai* (H_E=0.551) than in *A. aculeata*(H_E=0.466). We obtained evidence of hybridization between the species and that admixed individuals were assigned as F2 hybrids. In conclusion, this study showed the usefulness of microsatellite markers to elucidate the genetic boundaries of *A. aculeata* and *A. totai* , supporting their classification as different species and increase our knowledge about genetic diversity at the level of populations and species. The results are essentials to establish strategies for the adequate management, conservation, and domestication of both species.

Keywords: Macauba palm, Microsatellites markers, Domestication, Conservation, Genetic Resources

1. Introduction

The genus Acrocomia belongs to the Arecaceae family, and although it is widely studied, its systematics are controversial and have implications for the undefinition of its species number. According to Henderson et al. (1995), more than 35 species and/or synonyms are recognized by several authors within the genus. However, Henderson et al. (1995) attribute only two species to the genus, *A. aculeata* (Jacq.) Lodd. ex Mart. and *A. hassleri* (Barb. Rodr.) W.J. Hahn. The first species is large, arboreal, and widely distributed throughout Central and South America, and the second is herbaceous, small, and restricted to areas of the Cerrado of Brazil and part of Paraguay. However, the most accepted classification today is Lorenzi (2010), who recognizes seven species for the genus: *A. aculeata* (Jacq.) Lodd. ex Mart.,*A. intumescens* Drude, *A. totai* Mart., *A. crispa*(Kunth) C.F. Baker ex Becc., of tree size, and *A. hassleri, A. glaucescens* Lorenzi and *A. emensis* , of small size. In addition, the tree size species *A. media* , endemic from Puerto Rico and Virgin Islands, is also recognized by site The Plant List (2020).

A. aculeata and *A. totai* are the species of greatest economic interest, primarily to produce vegetable oil, food and feed (Lorenzi et al., 2006). *A. aculeata* has a wide geographical distribution, occurring in tropical and subtropical America from Mexico and the Antilles to Argentina, except for Peru and Ecuador (Henderson et al. 1995; Scariot et al. 1995). *A. aculeata* is the most common palm in Brazil, being found in the states of Para, Maranhao, Ceara, Minas Gerais, Goias, Mato Grosso, Mato Grosso do Sul, Sao Paulo, Parana, Santa Catarina and Rio Grande do Sul, with higher occurrence being observed in Cerrado regions (Scariot et al., 1991; 1995). *A. totai* occurs only in South America, northeastern Argentina, eastern Bolivia, and Paraguay. In Brazil, this species is restricted to the state of Mato Grosso do Sul (Markley, 1956; Lorenzi, et al., 2010; Rodriguez and Aschero, 2005).

Although *A. aculeata* and *A. totai* are classified in different taxa, in many reports, they are treated as a single species. This discrepancy may be due to the great vegetative similarity and because both have great plasticity of morphological attributes, exhibit broad variation of characteristics adopted for their differentiation or the absence of intraspecific morphological patterns (Crocomo and Melo, 1996; Vianna et al., 2017a). Both species are perennial with a single cylindrical stem and can reach 10 to 15 meters in height (Lorenzi 2006;

Scariot et al., 1991). Among the morphological characteristics used to differentiate these species are leaf sheath remnants and spines,*A. aculeata* has leaf sheath remnants, whereas *A. totai*does not. In relation to spines, *A. aculeata* has higher density and longer spines than *A. totai* . The fruit in both species is globose, drupe type, varying in size between species, with diameters ranging from 3.5 and 5.0 cm in *A. aculeata* and from 2.5 to 3.5 cm in *A. totai* . (Lorenzi, 2010; Vianna, et al., 2017a, Silva, 2017).

Studies based on morphological characteristics (Lorenzi, 2010, Silva, 20017), leaf anatomy (Vianna et al., 2017b) and biometric and physicochemical characteristics of the fruit (Machado et al., 2015) support the hypothesis that *A. aculeata* and *A. totai* are from different taxonomic groups. However, some morphological characters may reveal low capacity for unambiguous separation of species, especially when there is continuous variation and overlap in the morphology adopted to differentiate them (Minder and Widmer, 2008). In this case, it is necessary to use alternatives to elucidate a dubious taxonomic relationship between species.

According to Carlos Colombo (Campinas Agronomic Institute, personal communication), intermediate phenotypes between *A. aculeata* and*A. totai* have been observed in areas of sympatry of these species on the border between the Brazilian states of Sao Paulo and Mato Grosso do Sul, suggesting the occurrence of interspecific hybridization and, consequently, further blurring the distinction between *A. aculeata* and *A. totai* .

A. aculeata and *A. totai* has been identified as an important source for vegetable oil production due to the large amount of oil rich in oleic acid and lauric acid present in its fruits and the by products generated from oil extraction with added value and being of great demand from the food, cosmetic and energy industries (Colombo et al., 2018). In *A. aculeata* , the estimated yield of pulp-derived oil is approximately 5,000 liters.ha^{-1}, equaling that of palm oil and notably surpassing soybeans, which produce only 500 liters.ha^{-1} (Coimbra and Jorge, 2011; Cesar et al., 2015).

Due to the advantages of this species, studies mainly in *A. aculeata* have intensified in recent years. However, given the significant after-effects that incorrect species recognition may have in actual and future studies, it is necessary to establish a correct taxonomic delimitation of *A. aculeata* and *A. totai* . For breeding purposes, the elucidation of the genetic relationships between*A. aculeata* and *A. totai* is crucial for defining specific breeding strategies. Moreover, proof of interspecific hybridization can represent an important opportunity to increase genetic diversity and identify genotypes with complementary characteristics of agronomic interest, as well as representing an important source of variation for adaptation to new environments (Lewontin, 1966).

Microsatellite molecular markers or single sequence repeats (SSRs) are widely used in plant genetics studies because they have important qualities, such as high polymorphism or multiallelism, being codominant, reproducible and transferable between related species (Mason, 2015). These markers are also widely used in closely related taxa phylogeny studies and hybrid identification (Dobrovolskayaa, et al., 2015; Vieira et al., 2016) and have been successfully used to delimit different palm species: *Bactris gasipaes* Kunth (Couvreur et al., 2006),*Phoenix atlantic* (Henderson et al., 2006), and *Euterpe edulis* Mart. (Gaiotto et al., 2003).

In Acrocomia, microsatellites have been used to assess the diversity and genetic structure of natural populations and most studies have been carried out on *A. aculeata* (Abreu et al., 2012; Lanes, et al., 2015; Mengistu 2015; Silva, 2017; Oliveira et al. 2012). To date, only one study addressing the genetic diversity of *A. aculeata* and*A. totai* has been reported (Lima et al., 2020). Thus, in the present study we used a population genetic approach for elucidate the boundaries between *A. aculeata* and *A. totai* and compare genetic variation within and between populations and species using genomic Simple Sequence Repeats (SSR) and Expressed Sequence Tags (EST) SSR.

For the purpose of the study, the concept of species metapopulation lineage was adopted, which identifies species as metapopulation lineages that evolved separately, but that did not necessarily acquire contingent species properties, that is, phenolic distinction, reproductive isolation, monophilia and divergence ecological (Queiroz, 2007).

2. Materials and methods

2.1 Species and population sampling

For the present study, 175 individuals from 12 natural populations were analyzed in a geographical gradient from Formosa, Argentina to Minas Gerais, Brazil (Table 1). Based on the taxonomic classification of Lorenzi et al. (2010), six populations from the states of Minas Gerais and Sao Paulo were considered to be *A. aculeata* , and three populations from Mato Grosso do Sul and one from Argentina were considered to be as *A. totai.* In addition, we included two populations located in the areas where both species converge (simultaneous occurrence) and considered possible interspecific hybrids between *A. aculeata* and *A. totai* due to their morphological characteristics (Figure 1), as reported by Silva (2017).

Table 1. Description of *Acrocomia sp.* populations sampled for genetic analyses and their geographic parameters.

Figure 1. Map of geographic distribution of *Acrocomia spp.* sampled populations.

2.2 DNA extraction and microsatellite genotyping

Total genomic DNA was isolated from leaf material using the protocol of Doyle & Doyle (1990), and the DNA quality and quantity were evaluated on a 1% agarose gel and a NanoVue Plus spectrophotometer (GE Healthcare). The study was performed using 13 microsatellite markers, five from genomic regions or gSSR (Nucci et al., 2008) and eight from expressed sequence tags EST-SSRs (Bazzo, 2018) (Table 2). Polymerase chain reaction (PCR) amplifications were performed in 15 µL of total volume containing 20 ng of DNA, 2.0 µL of each primer forward and reverse at 5 µM, 3 µL of Hot Start PCR Master Mix (2X) and 9 µL of ultrapure water. The PCR reaction was conducted in a Bio-Rad thermocycler (model T100) under the following conditions: initial denaturation at 94 °C for 2 minutes followed by 30 cycles at 94 °C for 1 minute, annealing at 55-58 °C (depending on primer, Table 2) for 1 min, elongation at 72 °C for 1 min and final extension at 72 °C for 10 min. Amplification products were separated by capillary electrophoresis on an automated 96-Capillary Fragment Analyzer CE system (Advanced Analytical Technologies, Ames, IA, USA) using the DNF-905 Reagent Kit (Advanced Analytical Technologies, Ames, IA, USA).

2.3 Data analysis

Estimates of genetic diversity parameters as number of alleles (Na), the effective number of alleles (Ne), Shanon diversity index (I), number of private alleles, observed (Ho) and expected (He) heterozygosity, and Wright's F-statistics (Wright, 1922) were estimated using GenAlEx 6.2 software (Peakall and Smouse, 2006). The allelic richness (Ar) was estimated on the R platform (R Core Team, 2018) with the PopGenKit package (Paquette, 2012).

Genetic structure was inferred by Discriminant Analysis of Principal Components (DAPC) according to Jombart et al. (2010) and using the Adegenet statistical package (Jombart & Ahmed, 2011) from R platform (R Core Team, 2018). To perform the DAPC, each population sampled was considered a distinct genetic group. In addition, we used the Bayesian method in STRUCTURE 2.3.4 Program (Pritchard et al., 2000) to infer the genetic structure. The assumed parameters were an admixture model with correlated allele frequencies and without prior species information.

The number of K clusters was set from 1 to 15. For each number of K clusters considered, 10 independent runs were performed with a burn-in period of 100,000 iterations followed by 250,000 replicates of the Markov Monte Carlo chain (MCMC). To determine the most probable number of K clusters, the data were interpreted by the online tool STRUCTURE HARVESTER (Earl and Von Holdt, 2012), which uses Evanno's method (Evanno et al., 2005). To identify pure species and putative interspecific hybrids, the NewHybrids program version 1.1 beta was used (Anderson and Thompson, 2002). The calculation of the probability of each individual being a F1, F2, backcross or a parental species was performed with 10 independent MCMC runs with 500,000 steps and a burn-in period of 100,000 iterations.

3. Results

3.1 Genetic variability in gSSR and EST-SSR microsatellite loci

In the present study, we used genomic microsatellite (gSSR) and expressed region (EST-SSR) molecular markers. The genetic variability obtained from both markers shows that gSSRs are more variable than EST-SSRs in terms of number of polymorphic loci, number of alleles per locus and heterozygosity. The five gSSR markers detected 53 alleles with a mean of 10.6, ranging from six (Aacu 45) to 14 (Aacu26) alleles. EST-SSR loci detected 48 alleles with a mean of 6 alleles per locus and an allelic variation from two (EST-SSR 278) to 11 alleles per locus (EST-SSR 64). The PIC values of gSSR ranged from 0.567 to 0. 861 with a mean of 0.752, while the PIC values with the EST-SSR varying from 0.105 to 0.792 with a mean of 0.421. The mean of private alleles was higher in gSSR than detected with EST-SSR, with 2.60 versus 1.63 private alleles, respectively. The mean genetic variation revealed by the gSSR was higher (H_E=0.777) than that obtained with microsatellites derived from EST-SSR (H_E=0.455) (Table 2).

3. 2 Inter and intraspecific genetic diversity

The inter- and intraspecific genetic diversity analysis was conducted with a total of 13 microsatellite markers (5 gSSR and 8 EST-SSR). The species *A. totai* presented a mean of 6.69 alleles (Na) and 3.28 effective alleles (Ne) per locus, higher allelic diversity values than those obtained for *A. aculeata* , with an average of 5.38 alleles (Na) and 2.32 effective alleles (Ne) per locus. A higher number of private alleles was found in *A. totai* (16 alleles) than in*A. aculeata* (9 alleles), while in populations with hybrid plants, only two private alleles were detected (Table 3).

The mean of expected heterozygosity (H_E) was higher than that observed (H_O) in both species and populations with hybrid plants, indicating an excess of homozygotes and positive values for the fixation index. *A. totai* and populations with hybrid plants exhibited similar values of genetic diversity (H_O), 0.287 and 0.331, respectively, which were higher than those found for *A. aculeata* (H_O = 0.208) (Table 3).

Table 2. Characteristics of the five gSSR and eight EST-SSR primers and summary statistics of the genetic diversity at each locus.

Table 3. Genetic diversity estimates for Acrocomia taxa using 13 microsatellite loci.

At the population level, the number of alleles per locus (Na) ranged from 2.15 to 5.0, and effective alleles (Ne) ranged from 1.97 to 2.92 in the population of Itapira (6) and Corumba (11), respectively. The genetic diversity evaluated was higher in all populations of *A. totai* and in populations with hybrid plants than in the population of*A. aculeata* when estimated by the number of alleles (Na and Ne), Shanon index (I) and expected and observed heterozygosity (H_E and H_O). Brauna and Luz populations presented the highest and lowest values of genetic diversity, as evidenced by H_O=0.347 and 0.158, respectively (Table 4).

The average values of fixation index (f) found for all populations were positive relatively and high, indicating significant deviation of Hardy–Weinberg equilibrium and deficiency of heterozygosity (Table 4).

Table 4. Genetic diversity estimates at the population level using 13 microsatellite loci.

Genetic differentiation among study taxa was considered moderate according to the F_{ST} value obtained for the pair-wise comparisons (total F_{ST} = 0.105, P <0.001). The highest genetic differentiation was found between *A. aculeata* and*A. totai* (F_{ST} = 0.09). Lower genetic differentiation was observed for the comparison between the population with hybrid plants and both species, presenting values of F_{ST} = 0.079 for hybrids x *A. aculeata* and F_{ST} = 0.059 for hybrids x *A. totai* (Table 5). The estimation of genetic differentiation at the intrapopulation level was high (total F_{ST} = 0.267 P <0.001), ranging from 0.05 between Rifaina and Itapira populations to 0.80 between Luz and Argentina (Table 6).

Table 5 . Matrix of pairwise F_{ST} values (below diagonal) among*A . aculeata A. totai* and Hybrids, based on 13 microsatellites loci.

Table 6. Matrix of pairwise F_{ST} values (below diagonal) among populations based on 13 microsatellites.

3.3 Delimitation of species *A. aculeata* and *A. totai.*

The most likely number of groups (ΔK) revealed by the Bayesian analysis using the STRUCTURE software indicated the formation of two genetically distinct groups (K = 2) (Figure 2). Most individuals of the two genetic groups presented high homogeneity without uncertain assignments. Cluster red grouped most individuals from populations ascribed as *A. aculeata,* and cluster green grouped individuals from populations ascribed as *A. totai* according to the classification by Lorenzi et al. (2010). Subsequently, a new Bayesian analysis for the two main genetic groups revealed was performed; however, no genetic substructure was found (data not shown).

The clustering of individuals in the DAPC was in agreement with the Bayesian results of STRUCTURE and also identified the formation of two well define groups conformed by the *A. aculeata* (Cluster red) populations and by population of *A. totai* (Cluster green), corroborating the clear separation between species (Figure 3). However, as in the Bayesian analyses in STRUCTURE, it was not possible to observe population-level structuring within each species. The first two major components explained 44.3% of the total genetic variation, 34.8% for the first axis and 9.5% for the second axis. The hypothesis of the occurrence of two genetically distinct species was also confirmed by the NewHybrids program, wherein *A. aculeata* and *A. totai* were clearly recognized as a parental species (Figure 2.)

For the three analyses performed (Bayesians and DAPC, Figures 2 and 3), the populations of Braúna and Fusquinha were signed to the group of*A. totai* (Cluster red). This outcome was unexpected, given that these populations are currently considered *A. aculeata* based on its location in Sao Paulo state and the morphology according to Lorenzi classification (2010).

Figure 2. Results of Bayesian clustering assignment with the software's a) STRUCTURE (K=2) and b) NewHybrids. Individuals are represented by vertical bars where the color represents the posterior probability of assignment to each group.

Figure 3. Discriminant analysis of principal components (DAPC) among populations of *A. aculeata* and *A. totai* based on the microsatellite data set. Only the two-first axes showing the two higher discriminant eigenvalues are presented.

3.4 Interspecific hybridization

The results of the three analyses (NewHybrids, STRUCTURE and DAPC) for the identification of interspecific hybrids were congruent. The Bayesian analyses using NewHybrids software confirmed the existence of a hybrid genotype. In the population of Campo Grande, 3 of the 10 plants analyzed (30.0%) and in Braúna, 2 of the 13 plants analyzed (15.3%) were identified as F2 generation hybrids using a value of q=0.5 (Figure 2). Through analyses in the STRUCTURE, the five individuals classified as hybrids by the NewHybrids also showed signs of admixture (IQ <75) between *A. aculeata* and *A. totai* (Figure 2) and were also detected by DAPC analysis, as shown by the intermediate position of these individuals between the two main groups (Figure 3).

4. Discussion

4.1 Genetic variability in genomic gSSRs and EST-SSRs

Microsatellite markers have been widely used for plant genotyping because they are highly informative, codominant, multiallelic, reproducible, and transferable between related species (Mason, 2015). In palm trees, these markers have been used to delimit closely related species (Couvreur et al., 2006; Henderson et al. 2006; Pintaud et al., 2010). In the present study, microsatellite markers of expressed regions (EST-SSR) and genomic regions (gSSR) were adopted to elucidate the taxonomic relationship between the most economically important Acrocomia species (*A. aculeata* and *A. totai*) and to identify possible interspecific hybrids. In the genus Acrocomia, several genetic studies performed with microsatellite markers were conducted with at most 8 gSSR loci developed for *A. aculeata* by Nucci et al. (2008), as in Lanes et al. (2015), Lanes et al. (2016), Mengistu et al. (2016), Neiva et al. (2016), Araújo et al. (2017), Coelho et al., (2018) and Lima et al. (2020). Thus, this work represents the first study using both EST-SSR and gSSR markers in Acrocomia. Because these markers access different regions of the genome, the genetic information generated by both markers

tends to be complementary, and their combination provides more accurate information on the distribution of genetic diversity and population genetic parameters (Kalinowski 2002).

In the present study, both gSSR and EST-SSR primers developed for *A. aculeata* showed good amplification profiles and high polymorphisms at both the population and species levels, thereby confirming their transferability to *A. totai* , as previously mentioned by Bazzo et al. (2018), for EST-SSR and Lima et al. (2020), for gSRR. Comparing the two types of markers, the five gSSR loci presented a higher average number of alleles per locus, number of private alleles, and polymorphism index than the eight EST-SSR loci (Table 2).

These results are congruent with those reported in other plant species (Hu et al., 2011; Song et al., 2012; Meyer et al., 2017) and can be attributed to the location of EST-SSRs in transcribed regions of the genome and therefore be less subject to variation (Ellis and Burke 2007; Varshney et al. 2005). Considering gSSR and EST-SSR together, the average number of alleles detected in this study, the polymorphism index and diversity estimates (H_E and H_O) (Table 2) were higher than those found by Lanes et al. (2016) and Mengistu et al (2016) and lower than those reported by Lanes et al. (2015) and Araujo et al. (2017) in studies using only gSSR developed for *A. aculeata* . The differences in the estimates obtained in comparison with the works cited may be due to the botanical material analyzed, as well as the number and nature of the markers used.

Although EST-SSRs have lower levels of polymorphism compared to gSSR, they still have a high amount of polymorphism, demonstrating that the use of both types of microsatellite markers represents a valuable tool for genetic and evolutionary studies in *A. aculeata* and related species.

4.2 Inter and intraspecific genetic diversity

For the management of genetic resources through exploration or conservation, it is desirable to understand the magnitude and structure of the genetic diversity of the species. Most studies on genetic diversity in Acrocomia have been conducted with *A. aculeata*(Abreu et al., 2012; Oliveira et al. 2012; Lanes, et al., 2015; Mengistu 2016; Silva 2013; Araujo et al., 2017).

Studies on the genetic diversity of accessions of populations in Mato Grosso do Sul (Lanes at al. 2015; Mengsitu 2016) and in the border region between the states of São Paulo and Mato Grosso do Sul (Abreu et al., 2012; Coelho et al., 2018) considered *A. aculeata* as study material. A single study was reported on the genetic diversity of*A. totai* through molecular markers with samples only from Brazil (Lima et al., 2020). Thus, our study presents results on the genetic diversity of *A. aculeata* and *A. totai* based on population data from a wide gradient of geographical distribution of both species, including samples of *A. totai* from another country of occurrence.

Genetic-population estimates obtained from microsatellite data were similar between the two species. However, greater genetic diversity was evidenced in *A. totai* , which presented a higher number of alleles per locus, allelic richness and higher values in heterozygosities than*A. aculeata* (Tables 2). Higher values of genetic diversity were also found in *A. totai* at the population level. Our results are in line with previous studies conducted by Lima et al (2020), which also identified higher levels of genetic diversity in *A. totai* . Likewise, Lanes et al. (2015) and Mengsitu et al. (2016) reported a higher number of alleles per locus, a greater number of specific alleles and a higher percentage of polymorphism in populations from Mato Grosso do Sul when compared to populations from São Paulo and Minas Gerais, considering that populations from Mato Grosso do Sul were attributed by the authors to the species *A. aculeata* . Similarly, Coelho et al. (2018) obtained higher values of heterozygosity observed in populations located in the border of the states of São Paulo and Mato Grosso do Sul than in the population located in the central region of the state of São Paulo.

The populations of Ibituruna, MG (3) and Corumbá, MS (11), belonging to*A. aculeata* and *A. totai* , respectively, presented higher genetic diversity estimated by the heterozygosity and Shannon diversity index and may therefore represent an important source of *in situ*genetic diversity of these species.

The obtained data also reveal high values for the fixation index (f), indicating heterozygote deficiency in the populations of both species, with the f values of *A. aculeata* being higher than those of*A. totai* (Tables 3 and 4). The lower values in the genetic diversity parameters and the high fixation rates found for both species

may be due to the reproductive system of these species. In A. *aculeata* , several studies have reported the predominance of mixed reproduction systems, with a preference for allogamy but with a high rate of self-fertilization (Scariot et al. 1991; Nucci et al., 2008; Abreu et al., 2012) and even apomixis (Brito, 2013), which may favor the reduction of heterozygotes. Currently, there is no knowledge about the*A. totai* reproductive system, but based on the results obtained in the present work, we suggest that this species has a reproductive system with a higher allogamy rate than *A. aculeata* . Another hypothesis suggests that differences in the values of genetic diversity index and fixation index between *A. aculeata* and *A. totai*species are associated with different domestication processes.

In other words, in *A. totai,* the main products of anthropic interest are the yellowish flesh of the fruit for flour production and the oil extracted from the seed, both characteristics with little variation in the species (Sanjinez-Argandoña et al., 2011; Ciconini, et al 2013; Conceição et al., 2015). On the other hand, in *A. aculeata,* the domestication honored plants with higher oil content in the pulp, a characteristic that, according to several authors (Conceição et al., 2015; dos Reis et al., 2017), presents a large variation, which would have caused a positive selection for this characteristic with loss of variability of the individuals with lower pulp oil content.

High values of genetic diversity were obtained for the population coming from the municipality of Braúna (SP) and that, based on the Bayesian analysis, their individuals were considered putative interspecific hybrids between *A. aculeata* and *A. totai* . It is widely known that interspecific hybridization in plants may be responsible for increased levels of genetic diversity (López-Caamal et al., 2014), according to studies in several plant genera that have reported increased genetic diversity as a result of interspecific hybridization, such as *Ulmus* spp. (Zalapa et al., 2009), *Quercus* spp. (González-Rodríguez et al., 2005; Tovar-Sánchez et al., 2008) and*Phoenix* spp. (González-Pérez et al., 2004).

4.3 Delimitation of species *A. aculeata* and *A. totai*

Species recognition and delimitation represent strategic information for any biological discipline (Queiroz, 2007; Hey, 2006). However, obtaining this information can be a difficult task due to such events as hybridization, introgression, recent divergence, and low levels of morphological and genetic differentiation between species (Schönswetter et al. 2009; Lega et al. 2012; Slovák et al. 2012). In the genus Acrocomia, the classification of species is mainly based on their geographical distribution and morphology. In the case of the species of greater economic interest, *A. aculeata* and *A. totai* both have great morphological similarity and high polymorphism for many phenotypic characters, which has hindered their taxonomic classification and generated controversial, which is why many studies with botanical material of areas of occurrence typically of *A. totai* have considered the study species to include *A. aculeata* , as in the works of Gauto et al. (2011), working with material from Paraguay, Lescano et al. (2015) and Ciconini et al. (2013) with collections from Campo Grande (state of Mato Grosso do Sul) and Traesel et al. (2014) with collections from Dourados (state of Mato Grosso do Sul), or Dos Santos et al. (2013), who named individuals collected in Paranavaí, Paraná with *A. aculeata* subsp. *totai* .

In the present work we use a population genetic approach by means of molecular markers to elucidate the species boundaries between *A. aculeata* and *A. totai* and to identify the occurrence of interspecific hybrids between them. Our results were based on the analysis of a large number of microsatellite loci and expressive population sampling representing a wide geographical distribution and adopting different statistical analysis approaches. Thus, it was possible to highlight the differentiation of *A. aculeata* and*A. totai* species, which was strongly supported by the formation of two clusters that were well defined by both Bayesian (STRUCTURE) and principal component discriminant analysis (DAPC), as shown in Figures 2 and 3.

Corroborating this result, the analysis performed in the NewHybrids program also indicated the existence of two "pure" species (Figure 2), supporting the classification of *A. aculeata* and *A. totai*as distinct species, consistent with Lorenzi (2010) and Vianna et al. (2017a) based on morphological characters, Vianna et al. (2017b) based on leaf anatomy data and Lima et al (2020) based on molecular data generated from SSR marker.

Although not treated as different species, but based on their geographic location, other studies have provided evidence of differentiation of*A. aculeata* and *A. totai* from molecular data. Lanes et al. (2015) analyzed the genetic diversity of individuals from Acrocomia from different Brazilian localities with microsatellite markers and identified the formation of two main groups, one formed by individuals from the Pantanal region of Mato Grosso do Sul with high genetic differentiation according to the F_{ST} index. Similarly, the information generated by the ITS region of 67 *A. aculeata*genotypes from the study by Silva et al. (2017) identified the formation of four haplotypes, two of them shared by genotypes of São Paulo and Minas Gerais, and a unique haplotype of genotypes collected in Mato Grosso do Sul near the Paraguay border, also suggesting the occurrence of different species of Acrocomia.

Some phenotypic characteristics have been used to distinguish species from the genus Acrocomia (Lorenzi 2010). In the case of *A. aculeata* and *A. totai* , fruit characteristics are highly informative for this purpose, with *A. aculeata* presenting fruits between 3.5 and 5.0 cM, while *A. totai* has smaller fruits between 2.5 and 3.5 cM (Lorenzi 2010, Vianna et al., 2017a and Silva 2017). Machato et al. (2015) compared the fruit size of the Brazilian municipalities of Contagem (state of Minas Gerais) and Umuarama (state of Paraná); therefore, *A. aculeata* and *A. totai* had mean diameters of 5.03 and 3.42 cM, respectively. In addition to the morphological differences, there is also variation in the oil content of the fruits. Fruits of *A. totai* from populations of the Pantanal region of the state of Mato Grosso do Sul showed oil content in the pulp ranging from 26 to 33% (Vianna et al. 2015, Hiane et al. 2005), while fruits of *A. aculeata* collected in the states of São Paulo and Minas Gerais presented higher oil content, ranging from 30 to 76% (Conceição et al., 2015).

The detection of a high number of private alleles in *A. aculeata*and *A. totai* species (Table 2) indicates probable independent evolution of both species and the emergence of their own adaptation mechanisms to their respective environments. A higher number of private alleles in *A. totai* than *A. aculeata* can be explained by the diversification of habitats in which the species occurs from temporarily flooded wetland and seasonal floodplains, as well as in dry areas (Herrera-MacBryde et al., 2000; Lorenzi, 2010), in addition to being the only species of the genus Acrocomia to occur in temperate zones (Markley, 1956). In the case of *A. aculeata* , this species occurs preferentially in cerrado areas (Lorenzi, 2010) in large and continuous population arrangements.

Currently, one of the parameters adopted to define the classification of Acrocomia species is the geographical distribution of the species, whereas *A. totai* was not observed in the state of São Paulo (Lorenzi, 2010). However, our results reveal that the populations of Braúna and Fusquinha, previously considered *A. aculeata* , due to their geographical location in the state of São Paulo, were grouped between *A. totai* samples from Argentina and the state of. Mato Grosso do Sul (Figures 2 and 3). These results are identical to those obtained by Abreu et al. (2012) in a study on the structuring of genetic diversity of *A. aculeata* samples from different origins of the states of São Paulo and Minas Gerais. In this work, samples from the locality of Piquerobi (region near the Fusquinha collections of our study) were genetically distanced from the samples from other locations, possibly because they were representative of *A. totai* and not of*A. aculeata* , as suggested at that time.

The origin of the genus Acrocomia is unknown. Pintaud et al. (2008) suggest as a possible center of origin the northern region of South America, given that the progressive shrinkage of the humid tropical rainforest in South America during the Eocene resulted in drier and more open environments that favored the process of diversification of Acrocomia. The oldest fossil records of *A. aculeata* found in Santarém, Pará, northern Brazil, dating from 11,200 years (Morcote-Ríos and Bernal, 2001) suggest that the center of origin of the genus Acrocomia may be the northern region of Brazil. This hypothesis is also be supported by the finding that with the exception of *A. crispa*and *A. media* , the other species of the genus occur in Brazil, with *A. emensis* , *A. glaucescens* and *A. intumescens*being endemic from Brazil (Lorenzi et al (2010; Henderson, 1995).

Based on these data, it can be assumed that the beginning of domestication of the species *Acrocomia aculeata* occurred in northern South America and was subsequently exposed to distinct patterns of local anthropic interest. The dispersal of human groups towards the western side of South America, particularly in the

countries of Bolivia, Paraguay and northern Argentina and part of the western region of Brazil bordering these countries, could explain the diffusion and emergence of a new botanical group represented by the species *A. totai* . Thus,*A. aculeata* would have taken other paths, spreading through central Brazil from Pará to the north of Paraná state (Lorenzi et al., 2010). In this case, these species would have gone through a process of allopatric speciation, geographical isolation, limiting contact and gene flow between them.

Another argument to argue for a possible genetic differentiation between*A. aculeata* and *A. totai* species associated with anthropic use and, consequently, by different domestication processes and different selective pressures due to the preferences of various indigenous groups are based on historical and ethnographic accounts. which documented different popular uses of *A. aculeata* (Welch et al., 2013, Nascimento et al., 2010; Balée, 2013; Hill et al., 1984) and*A. totai* (Steward, 1946; Patiño 2002).

The hypothesis of dispersion of Acrocomia spp is based on reports by several authors. In South America, both *A. aculeata* and *A. totai* have been strongly associated with human dispersions (Seemann, 1856; Janzen, 1983; Kahn and Moussa 1995; Piperno and Pearsall 1998; Morcote-Rios and Bernal 2001). Specifically, Barbosa Rodrigues (1891) reported that Bocayauba (*A. totai*) accompanied North-South Indian migrations and was always associated with the interests of indigenous tribes.

More recently, the dispersal of *A. aculeata* and *A. totai*seeds was observed to be performed by cattle herds, which are considered contemporary dispersers (Scariot, 1998, Donatti et al., 2011). In the specific case of *A. aculeata* , the expansion of the agricultural frontier in the Brazilian Cerrado in search of new areas for cattle pasture may have favored events of rapid and wide dispersal of the species.

4.4 Interspecific hybridization

Another objective of the study was to investigate the occurrence of hybrids between *A. aculeata* and *A. totai* . This hypothesis arose from visual field observations of individuals with intermediate morphology between both species, as well as from previous reports describing possible natural hybridization between these species (Abreu et al., 2012).

In Arecaceae, interspecific hybridization has been reported in several genera: *Attalea* (Henderson et al. 1995), *Calyptrogyne*H.Wendl. (Henderson 2005), *Caryota* L. (Hahn and Sytsma, 1999),*Copernicia* Mart. (Henderson et al. 1995), *Desmoncus* Mart. (Henderson et al. 1995), *Hyospathe* Mart. (Henderson 2004),*Phoenix* L. (Gonzalez-Perez et al. 2004; Pintaud et al., 2010),*Syagrus* Mart. (Henderson et al. 1995; Ramírez-Rodríguez et al., 2011). However, in the genus Acrocomia, there are no reports of studies on natural hybridization among its species to date.

Thus, based on genetic data and Bayesian analysis performed in the NewHybrids program, our results suggest the occurrence of interspecific hybrids in the populations of Braúna and Campo Grande, which have been attributed to generation class F2 (Figure 2). This result is also supported by the Bayesian analyses in STRUCTURE and DAPC (Figures 2 and 3).

Furthermore, we add that the flowering period of *A. aculeata* and*A. totai* species overlap in the sympathetic region of these species, favoring the possible occurrence of natural hybridization (Salis 2009; Scariot et al., 1995; Lorenzi 2006). Therefore, we are reporting unprecedented evidence of the occurrence of natural hybrids between species of the genus Acrocomia through molecular data. A previous study, based on the SSR marker, verified connectivity and admixture in a population of *A. totai* located geographically close to the population of *A. aculeata* (Lima et al., 2020).

Morphological data from previous studies also suggest a possible interspecific hybridization between *A. aculeata* and *A. totai* by revealing intermediate values for fruit characteristics. Ciconinni et al. (2013), studying fruits collected in Campo Grande (MS) and Sanjinez-Aragandoña and Machado (2011) in fruits collected in Presidente Epitácio, reported averages of 3.3 cm and 3.6 cm for vertical diameter and 3.4 cm and 3.1 cm for horizontal diameter, respectively.

Interspecific hybridization is more frequent in closely related species or those that are not sufficiently di-

vergent to develop complete reproductive isolation mechanisms (Taylor et al. 2015). Hybridization may be promoted by secondary contact after allopatric divergence, as has been reported in several studies (Petitet al. 2002; Fussi et al. 2010). Hybridization between *A. aculeata* and *A. totai* may have been the result of secondary contact after more recent species expansion across South America, as corroborated by the close genetic relationship between species, as low F_{ST} values (0.09) (Table 5), and maintenance of interspecific reproductive compatibility and formation of viable hybrids. The presence of only F2 hybrids and the low number of private alleles present in populations chosen to represent interspecific hybrids (Table 3) suggest that cross-species hybridization is recent without sufficient time for backcrossing or to accumulate mutations over many generations.

Currently, *A. aculeata* can be widely found throughout the state of Sao Paulo. However, according to the report of Hoehen (1944), the floristic surveys conducted in the nineteenth century by Saint-Hilaire do not cite the presence of the species *A. aculeata* in the state of Sao Paulo. However, the state of São Paulo presented a rapid process of occupation of its territory, primarily from the late nineteenth and early twentieth centuries, with the opening of several rail routes to the north, west and southwest of the state, reaching the border with the state of Mato Grosso do Sul (Oliveira and Marquis, 2002), where the species *A. totai* is observed in large massifs. Therefore, it can be assumed that the emergence of natural hybrids between *A. aculeata* and *A. totai* may have been facilitated by the dispersal of *A. aculeata* in the state of São Paulo during this state colonization process, enabling the sympatric coexistence of both species.

Finally, we emphasize that the interspecific hybrids identified between*A. aculeata* and *A totai* are at low frequency. This low number of hybrids detected may be due to the reduced number of plants analyzed in the Braúna (13 plants) and Campo Grande (10 plants) populations, suggesting that additional studies sampling more populations and plants by populations and using a multidisciplinary approach are needed to corroborate our results.

5. Conclusions

Our study represents a significant step towards understanding the systematics of the genus Acrocomia. The results show that *A. aculeata* and *A. totai* are genetically distinct species with strong evidence of possible interspecific hybridization. The occurrence of interspecific hybrids represents an opportunity to increase genetic diversity both for the appearance of genotypes with complementary characteristics of agronomic interest, as well as for representing an important source of variation for adaptation to new environments, especially in function of current climate changes. Additionally, this study is the first to assess the genetic diversity of *A. totai*from another country besides Brazil and to show that the genetic diversity in *A. totai* is superior to that found in *A. aculeata* . Our results contribute to the choice of better strategies for in situ and ex situ management of germplasm, as well as to guide selection criteria for the purposes of genetic conservation or domestication of both species.

Acknowledgements

We are grateful to Diego Wassner for kindly providing *A. totai*samples from different population of Argentina.

Data accessibility

The authors have deposited the microsatellite genotyping data at the Open Science Framework available at https://doi.org/10.17605/OSF.IO/C4YE2

References

Abreu AG, Priolli RHG, Azevedo-Filho JA, Nucci SM, Zucchi MI., Coelho RM, Colombo CA (2012) The genetic structure and mating system of*Acrocomia aculeata* (Areaceae) Genet Mol Biol 35:119-121. https://doi.org/10.1590/S1415-47572012005000002

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217–1229.

Araújo MRG, Melo J, Afrânio F, Menezes EV, Brandão MM, Cota LG, Oliveira DA, Royo VA, Vieira FA (2017) Fine-scale spatial genetic structure and gene flow in *Acrocomia aculeata* (Arecaceae): Analysis in an overlapping generation. Biochem Syst Ecol 71:47-154. https://doi.org/10.1016/j.bse.2017.02.005

Barbosa Rodrigues J (1891) Plantas novas cultivadas no Jardim Botânico do Río de Janeiro. Descriptas, classificadas e desenhadas. Rio de Janeiro, Brazil.

Balée W (2013) Cultural Forests of the Amazon: A Historical Ecology of People and their Landscape. Tuscaloosa, AL: University of Alabama Press.

Bazzo BR (2018) Macaubest: a comprehensive transcriptome of macaúba palm (*Acrocomia aculeata*). Thesis, University of Campinas-UNICAMP. Bazzo BR, Carvalho LM de, Carazzolle MF, Guimarães PGA, Colombo CA (2018) Development of novel EST-SSR markers in the macaúba palm (*Acrocomia aculeata*) using transcriptome sequencing and cross-species transferability in Arecaceae species. BMC Plant Biol 18: 2-10 https://doi.org/10.1186/s12870-018-1509-9

Brito AC (2013) Biologia reprodutiva de Macaúba: Floração, Polinizadores, Frutificação e Conservação de Pólen. Thesis, Universidade Federal de Viçosa

César AS, Almeida FA, Souza RP, Silva GC, Atabani AE (2015) The prospects of using *Acrocomia aculeata* (macaúba) a non-edible biodiesel feedstock in Brazil. Renew Sustain Energy Rev 49: 1213-220. https://doi.org/10.1016/j.rser.2015.04.125

Ciconini G, Favaro SP, Roscoe R, Miranda CHB, Tapeti CF, Miyahira MAM, Bearari L, Galvani F, Borsato AV, Colnago LA, Naka MH, (2013) Biometry and oil contents of *Acrocomia aculeata* fruits from the Cerrados and Pantanal biomes in Mato Grosso do Sul, Brazil. Ind Crops Prod 45: 208–214. https://doi.org/10.1016/j.indcrop.2012.12.008

Coimbra MC, Jorge N (2012) Fatty acids and bioactive compounds of the pulps and kernels of Brazilian palm species, guariroba (*Syagrus oleracea*), jerivá (*Syagrus romanzoffiana*) and macaúba (*Acrocomia aculeata*). J Sci Food Agric 92:679–684. https://doi.org/10.1002/jsfa.4630

Coelho NHP, Tambarussi EV, Aguiar BI, Roque RH, Portela RM, Braga RC, Sanson Da Silva RAR, Ferraz EM, Moreno MA, Kageyama PY, Gandara FB (2018) Understanding genetic diversity, spatial genetic structure, and mating system through microsatellite markers for the conservation and sustainable use of *Acrocomia aculeata* (Jacq.) Lodd. Ex Mart. Cons Genet 19:879-891. https://doi.org/10.1007/s10592-018-1061-z

Colombo CA, Berton LHC, Diaz BG, Ferrari RA (2018) Macauba: a promising tropical palm for the production of vegetable oil. OCL 25:1-9. https://doi.org/10.1051/ocl/2017038

Conceição LDHCS, Antoniassi R, Junqueira NTV, Braga MF, Faria-Machado AF, Rogério JB, Duarte ID, Bizzo HR (2015) Genetic diversity of macauba from natural populations of Brazil. BMC Res Notes 8:1-9. https://doi.org/10.1186/s13104-015-1335-1

Crocomo OJ, Melo M (1996) Acrocomia Species (Macauba Palm). In: Bajaj Y.P.S. (eds) Trees IV. Biotechnology in Agriculture and Forestry, Springer, Berlin, Heidelberg.

Couvreur TLP, Billotte N, Risterucci AM, Lara C, Vigouroux Y, Ludeña B, Pham JL, Pintaud JC (2006) Close genetic proximity between cultivated and wild *Bactris gasipaes* Kunth revealed by microsatellite markers in Western Ecuador. Genet Resour Crop Evol 53:361-1373. https://doi.org/10.1007/s10722-005-5033-z

Da Silva PH da (2017) Caracterização de *Acrocomia aculeata* e*Acrocomia totai* por meio de descritores agromorfológicos. Dissertation, Instituto Agronômico de Campinas

Dobrovolskayaa OB, Pontc C, Orlova YL, Salsec J (2016) Development of New SSR Markers for Homoeologous WFZP Gene Loci Based on the Study of the Structure and Location of Microsatellites in Gene Rich Regions of Chromosomes 2AS, 2BS, and 2DS in Bread Wheat. Russ J Genet: Appl Res 6:330-337. https://doi.org/10.1134/S2079059716030023

Donatti CI, Guimaraes PR, Galetti M, Pizo MA, Marquitti FMD, Dirzo R (2011) Analysis of a hyperdiverse seed dispersal network: modularity and underlying mechanisms. Ecol Letters 14:773-81. https://doi.org/10.1111/j.1461-0248.2011.01639.x

Dos Reis EF, Pinto JFN, Assunção HF da, Silva DFP da (2017) Genetic diversity of macaúba fruits from 35 municipalities of the state of Goiás, Brazi. Pesq Agropec Bras 52:277-282. https://doi.org/10.1590/s0100-204x2017000400008

Dos Santos JVF, Machado W, Lira FF, Takahashi LSS, Guimarães MF, Lea AC (2013) Caracterização Biométrica de Frutos Diferentes de Macaúba. BBR- Biochem Biotech Reports Especial number 2:14-16. http://dx.doi.org/10.5433/2316-5200.2013v2n3espp14

Earl DA, VonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Cons Genet Resour 4: 359–361. https://doi.org/10.1007/s12686-011-9548-7

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity 99:125-132. https://doi.org/10.1038/sj.hdy.6801001

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fussi B, Lexer C, Heinze B (2010) Phylogeography of *Populus alba*(L.) and *Populus tremula* (L.) in Central Europe: secondary contact and hybridization during recolonisation from disconnected refugia. Tree Genet Genomes 6:439-450. https://doi.org/10.1007/s11295-009-0262-5

Gaiotto FA, Grattapaglia D, Venkovsky R (2003) Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart.). J Hered 94:399-406. https://doi.org/10.1093/jhered/esg087

Gauto I, Spichiger RE, Stauffe FW (2011) Diversity, distribution and conservation status assessment of Paraguayan palms (Arecaceae). Biodivers Conserv 20:2705-2728. https://doi.org/10.1007/s10531-011-0100-6

González-Peréz MA, Caujapé-Castells J, Sosa PA (2004) Allozyme variation and structure of the Canarian endemic palm tree *Phoenix canariensis* (Arecaceae): implications for conservation. Heredity 93:307-315. https://doi.org/10.1038/sj.hdy.6800507

González-Rodríguez A, Arías DM, Oyama K (2005) Genetic variation and differentiation of populations within the *Quercus affinis – Quercus laurina* (*Fagaceae*) complex analyzed with RAPD markers. Can J Bot 83:155-162. https://doi.org/10.1139/b04-162

Hahn WJ, Sytsma KJ (1999) Molecular systematics and biogeography of the Southeast Asian genus Caryota (Palmae). Syst Bot 24:558-580.

Henderson A, Galeano G, Bernal R (1995) Field Guide to the Palms of the Americas New Jersey: Princepton University, New Jersey, USA.

Henderson SA, Billotte N, Pintaud JC (2006) Genetic isolation of Cape Verde Islands *Phoenix atlantic* a (Arecaceae) revealed by microsatellite markers. Cons Genet 7:213-223. https://doi.org/10.1007/s10592-006-9128-7

Henderson AJ (2004) A multivariate analysis of Hyospathe (Palmae). Am J Bot 91:953- 965. https://doi.org/10.3732/ajb.91.6.953

Henderson AJ (2005) A multivariate study of Calyptrogyne (Palmae). Syst Bot 30:60-83. https://doi.org/10.1600/0363644053661913

Herrera-MacBryde O, Dallmeier F, MacBryde B, Comiskey JA, Miranda C (2000) Biodiversidad, Conservación y Manejo en la Región de la Reserva de la Biosfera Estación Biológica del Beni, Bolivia. SI/MAB no. 4, Smithsonian Institution, Washington, D.C.

Hey J (2006) On the failure of modern species concepts. Trends Ecol Evol 21:447-450. https://doi.org/10.1016/j.tree.2006.05.011

Hiane PA, Ramos Filho MM, Ramos MIL, Macedo MLR (2005) Bocaiúva,*Acrocomia aculeata* (Jacq.) Lodd., pulp and kernel oils: characterization and fatty acid composition. Braz J Food Technol 8:256-259.

Hill K, Hawkes K, Hurtado M, Kaplan H (1984) Seasonal variance in the diet of Aché hunter-gatherers in eastern Paraguay. Hum Ecol 12:101-135. https://doi.org/10.1007/BF01531269

Hoehne FC (1944) Relatório Annual do Instituto de Botânica. São Paulo, Secretaria de Agricultura do Estado. São Paulo, Brazil

Hu J, Wang L, Li J (2011) Comparison of genomic SSR and EST-SSR markers for estimating genetic diversity in cucumber. Biol. Plantarum 55:577-580.

Janzen D (1983) *Acrocomia vinifera* (Coyol). in D. H. Janzen (ed.), Costa Rican natural history. Univ. of Chicago Press, Chicago, pp 184-185.

Jombart T, Ahmed I (2011) Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070-3071. https://doi.org/10.1093/bioinformatics/btr521

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. BMC Genetics 11:94. https://doi.org/10.1186/1471-2156-11-94

Kahn F, Moussa F (1995) Les migrations de palmiers provoquées par l'homme en Amazonie et à sa périphérie-Un premier constat. Biogeographica 71:161-177.

Kalinowski ST (2002) How many alleles per locus should be used to estimate genetic distances? Heredity 88:62-65. https://doi.org/10.1038/sj.hdy.6800009

Lanes ECM, Motoike YS, Kuki NK, Nick C, Freitas DR (2015) Molecular Characterization and Population Structure of the Macaw Palm,*Acrocomia aculeata* (Arecaceae), Ex Situ Germplasm Collection Using Microsatellites Markers. J Hered 106:102-112. https://doi.org/10.1093/jhered/esu073

Lanes, E. C. M., Motoike, S. Y., Kuki, K. N., Resende, M. D. V., Caixeta, E. T., 2016. Mating system and genetic composition of the macaw palm (*Acrocomia aculeata*): implications for breeding and genetic conservation programs. J Hered 107, 527-536. https://doi.org/10.1093/jhered/esw038

Lega M, Foir S, Prosser F, Bertolli A, Li M, Varotto C (2012) Application of the unified species concept reveals distinct lineages for disjunct endemics of the *Brassica repanda* (Brassicaceae) complex. Biol. J Linn Soc 106:482-497. https://doi.org/10.1111/j.1095-8312.2012.01887.x

Lescano CH, Oliveira IP, Silva LR, Baldivia DS, Sanjinez-Argandoña EJ, Arruda EJ, Moraes ICF, Lima FF (2015) Nutrients content, characterization and oil extraction from *Acrocomia aculeata*(Jacq.) Lodd. Fruits. Afr J Food Sci 9:113-119. https://doi.org/10.5897/AJFS2014.1212

Lewontin RC, Birch LC (1966). Hybridization as a source of variation for adaptation to new environments. Evolution, 315-336.

Lima NE de, Meerow AW, Manfrin MH (2020) Genetic structure of two Acrocomia ecotypes (Arecaceae) across Brazilian savannas and seasonally dry forests. Tree Genet Genomes16:56. https://doi.org/10.1007/s11295-020-01446-y

López-Caamal A, Tovar-Sánchez E (2014) Genetic, morphological, and chemical patterns of plant hybridization. Rev Chil Hist Nat 87:16. https://doi.org/10.1186/s40693-014-0016-0

Lorenzi GMAC (2006) *Acrocomia aculeata* (Jacq.) Lodd. ex Mart. - Arecaceae: bases para o extrativismo sustentável. Thesis, Universidade Federal do Paraná

Lorenzi H, Noblick L, Kahn F, Ferreira E (2010) Flora Brasileira: Arecaceae: Palmeiras. 1st ed. Nova Odessa: Plantarum.

Machado W, Guimarães MF, Lira FF, Santos JVF, Takahashi LSA, Leal A, Coelho GTCP (2015) Evaluation of two fruit ecotypes (*totai* and*sclerocarpa*) of macaúba (*Acrocomia aculeata*). Ind Crops Prod 63:287–293. https://doi.org/10.1016/j.indcrop.2014.11.002

Mason AS (2015) SSR Genotyping. In: Batley J. Plant Genotyping. Methods in Molecular Biology (*Methods and Protocols*). Humana Press, New York, NY, pp 77-89. https://doi.org/10.1007/978-1-4939-1966-6_6

Markley KS (1956) Mobocayá or Paraguay coco palm: an important source of oil. Econ Bot 10:3-32,

Mengistu FG, Motoike SY, Cruz CD (2016) Molecular characterization and genetic diversity of the macaw palm ex situ germplasm collection revealed by microsatellite markers. Diversity 8:20. https://doi.org/10.3390/d8040020

Meyer L, Causse R, Pernin F, Scalone R, Bailly G, Chauvel B, Le Corre V (2017) New gSSR and EST-SSR markers reveal high genetic diversity in the invasive plant *Ambrosia artemisiifolia* L. and can be transferred to other invasive Ambrosia species. PLoS ONE 12, e0176197. https://doi.org/10.1371/journal.pone.0176197

Minder AM, Widmer A(2008) A population genomic analysis of species boundaries: neutral processes, adaptive divergence and introgression between two hybridizing plant species. Mol Ecol 17:1552-1563. https://doi.org/10.1111/j.1365-294X.2008.03709.x

Morcote-Rios G, Bernal R (2001) Remains of Palms (Palmae) at Archaeological Sites in the New World: A Review. The Bot Rev 67:309-350.

Nascimento AR (2010) Riqueza e etnobotanica de palmeiras no Territorio Indigena Kraho. Tocantins. Brasil. Floresta 40:209-220. http://dx.doi.org/10.5380/rf.v40i1.17112

Neiva DS, Melo Junior AF, Oliveira DA, Royo VA, Brandao MM, Menezes EV (2016) *Acrocomia emensis* (Arecaceae) genetic structure and diversity using SSR molecular markers. Genet Mol Res 15:1-11. http://dx.doi.org/10.4238/gmr.15017785

Nucci SM, Azevedo-Filho JA, Colombo CA, Priolli RH, Zucchi MI (2008) Development and characterization of microsatellites markers from the macaw. Mol Ecol Resour 8:224–226. https://doi.org/10.1111/j.1471-8286.2007.01932.x

Oliveira DA, Melo JAF, Brandao MM, Rodrigues LA, Menezes EV, Ferreira PR (2012) Genetic diversity in populations of *Acrocomia aculeata*(Arecaceae) in the northern region of Minas Gerais, Brazil. Genet Mol Res 11:531-538. http://dx.doi.org/10.4238/2012.March.8.1

Oliveira PS, Marquis RJ (2002) The Cerrados of Brazil: Ecology and Natural History of a Neotropical Savanna. Columbia University Press, New York

Paquette SR (2012) PopGenKit: Useful Functions for (batch) file conversion and data resampling in microsatellite datasets. R package version 1.0.

Patino RVM (2002) Historia y dispersion de los frutales nativos del Neotropico. CIAT No. 326. Cali, Colombia

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288-295.

Petit RJ, Csaikl M, Bordacs S, Burg K, Coart E, Cottrell J, van Dam B, Deans JD, Dumolin-Lapegue S, Fineschi S, Finkelday R, Gillies A, Gla I, Goicoechea PG, Jensen JS, Konig A, Lowe AJ, Madsen SF, Matyas G, Munro RC, Olalde M, Pemonge MH, Popescu F, Slade D, Tabbener H, Taurchini D, de Vries

SMG, Ziegenhagen B, Kremer A (2002) Chloroplast DNA variation in European white oaks: phylogeography and patterns of diversity based on data from over 2600 populations. For Ecol Manage 156:5-26. https://doi.org/10.1016/S0378-1127(01)00645-4

Pintaud JC, Zehdi S, Couvreur T, Barrow S, Henderson S, Aberlenc-Bertossi F, Tregear J, Billotte N (2010) Species Delimitation in the Genus Phoenix (Arecaceae) Based on SSR Markers, with Emphasis on the Identity of the Date Palm (*Phoenix dactylifera*). Seberg O, Petersen G, Barfod A, Davis J. (Eds.): Diversity, phylogeny and evolution in the Monocotyledons. Aarhus University Press, Denmark, pp 267-286.

Pintaud JC, Galeano G, Balslev H, Bernal R, Borchsenius F, Ferreira E, Granville JJ, Mejia K, Millan B, Moraes M, Noblick L, Stauffer FW, Kahn F (2008) Las palmeras de America del Sur: diversidad, distribucion e historia evolutiva. Rev Peru Biol 15:7-29.

Piperno D, Pearsall D (1998) The origins of agriculture in the lowland Neotropics. Academic Press, San Diego.

Prichard JK, Stephens M, Donnelly P (2000) Inference of population structure from multilocus genotype data. Genetics 155:945–959.

Queiroz K de (2007) Species concepts and species delimitation. Syst Biol 56:879-886. https://doi.org/10.1080/10635150701701083

R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical. Computing, Vienna, Austria. URL http://www.R-project.org/

Ramirez-Rodriguez R, Tovar-Sanchez E, Jimenez RJ, Vega FK, Rodriguez V (2011) Introgressive hybridization between *Brahea dulcis* and*Brahea nitida* (Arecaceae) in Mexico: evidence from morphological and PCR–RAPD patterns. Botany 89:545-557. https://doi.org/10.1139/b11-043

Rodriguez MF, Aschero CA (2005) *Acrocomia chunta* (Arecaceae) raw material for cord making in the Argentinean Puna. J Archaeol Sci 32:1534-1542.

Salis SM, Mattos PP (2009) Floracao e Frutificacao da Bocaiuva (*Acrocomia aculeata)* e do Caranda (*Copernicia alba*) no Pantanal. Corumba: Embrapa Pantanal. Comunicado Tecnico 78.

Sanjinez-Argandona EJ, Machado CCA (2011) Caracterizacao biometrica, fisica e quimica de frutos da palmeira bocaiuva *Acrocomia aculeata* (Jacq) Lodd. Rev Bras Frutic 33:1023-1028. https://doi.org/10.1590/S0100-29452011000300040

Scariot AO (1998) Seed dispersal and predation of the palm*Acrocomia aculeata* . Principes 42:5-8.

Scariot AO, Lleras E, Hay JD (1995) Flowering and fruiting phenologies of the palm *Acrocomia aculeata* : patterns and consequences. Biotropica 27:168-173. https://doi.org/10.2307/2388992

Scariot AO, Lleras E, Hay DJ (1991) Reproductive biology of the palm*Acrocomia aculeata* in Central Brazil. Biotropica 23:12-22. https://doi.org/10.2307/2388683

Schonswetter P, Solstad H, Garcia PE, Elven R (2009) A combined molecular and morphological approach to the taxonomically intricate European mountain plant *Papaver alpinum* s.l. (Papaveraceae)-taxa or informal phylogeographical groups? Taxon 58:1326-1343. https://doi.org/10.1002/tax.584020

Seemann B (1856) Popular history of the palms and their allies. Lovell Reeve, 5, Henrietta Street, Covent Garden.

Silva LCC, Lemos RC, Carvalho CGI, Good-god PIV, De Oliveira LO, Costa MDL, Piovesan ND, Moreira MA (2017) Genetic diversity and structure of macaw palm: Implications for genetic variability sampling. Rev Arvore 41:1-9. http://dx.doi.org/10.1590/1806-90882017000500008

Slovak M, Kucera J, Marhold K, Zozomova-Lihova J (2012) The morphological and genetic variation in the polymorphic species*Picris hieracioides* (Compositae, Lactuceae) in Europe strongly contrasts with traditional taxonomical concepts. Syst Bot 37:258-278. https://doi.org/10.1600/036364412X616828

Song YP, Jiang XB, Zhang M, Wang ZL, Bo WH, An XM, Zhang DQ, Zhang ZY (2012) Differences of EST-SSR and genomic-SSR markers in assessing genetic diversity in poplar. For Stud China 14:1-7. doi:10.1007/s11632-012-0106-5. https://doi.org/10.1007/s11632-012-0106-5

Steward J H (1946) Handbook of South American Indians. Vol. 1 The marginal tribes. Washington, USA.

Traesel GK, De Souza JC, De Barros AL, Souz MA, Schmitz WO, Muzzi RM, Arena AC (2014) Acute and subacute (28 days) oral toxicity assessment of the oil extracted from *Acrocomia aculeata* pulp in rats. Food Chem Toxic 74:320–325. https://doi.org/10.1016/j.fct.2014.10.026

Tovar-Sanchez E, Mussali-Galante P, Esteban-Jimenez R, Pinero D, Arias DM, Dorado O, Oyama K (2008) Chloroplast DNA polymorphism reveals geographic structure and introgression in the *Quercus crassipes* x*Quercus crassifolia* hybrid complex in Mexico. Botany 86:228-239. https://doi.org/10.1139/B07-128

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48-55. https://doi.org/10.1016/j.tibtech.2004.11.005

Vianna SA, Berton LHC, Pott A, Guerreiro SMC, Colombo CA (2017a). Biometric Characterization of Fruits and Morphoanatomy of the Mesocarp of Acrocomia Species (Arecaceae). Int J Biol 9:78-92.

Vianna SA, Carmelo-Guerreiro SM, Noblick LR, Colombo CA (2017b) Leaf anatomy of Acrocomia (Arecaceae): an additional contribution to the taxonomic resolution of a genus with great economic potential. Plant Syst Evol 303:233–248. https://doi.org/10.1007/s00606-016-1369-4

Vianna SA, Hiane PA, Jordao MN, Da Costa EC, Pott A (2015) Physical and nutritional assessment of fruits of *Acrocomia aculeata* (Jacq.) Lodd ex Mart. (Arecaceae) based on pulp color. Ver Biodivers Neotrop 5:89-95.

Vieira MLC, Santini L, Diniz AL, Munhoz CF (2016) Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol 39: 312-328.

Welch J, Ventura Santo R, Flowers NM, Coimbra Jr CEA (2013) Na primeira margem do rio: territorio e ecologia do povo Xavante de Wedeze. Museu do Indio-FUNAI. Rio de Janeiro, Brazil.

Wright S (1922) Coefficients of inbreeding and relationship. Am Nat 56:330-338.

Zalapa JE, Brunet J, Guries RP (2009) Patterns of hybridization and introgression between invasive *Ulmus pumila* (Ulmaceae) and native*U. rubra* . Amer J Bot 96:1116-1128. https://doi.org/10.3732/ajb.0800334

Table 1. Description of *Acrocomia sp.* populations sampled for genetic analyses and their geographic parameters.

ID	Population	Estate*	Country	No. Plants	Geographical coordinates	Geographical coordinates
1	Luz	MG	Brazil	16	-19.773972 W	-45.864639 S
2	Patos de Minas	MG	Brazil	15	-18.591389 W	-46.455111 S
3	Ibituruna	MG	Brazil	12	-21.343028 W	-44.739736 S
4	Rifaina	MG	Brazil	12	-19.986278 W	-47.508472 S
5	Brotas	SP	Brazil	10	-22.276083 W	-48.11850 S
6	Itapira	SP	Brazil	17	-22.433333 W	-46.821667 S
7	Braúna	SP	Brazil	13	-21.482306 W	-50.212278 S
8	Fusquinha	SP	Brazil	18	-22.11100 W	-52.252472 S
9	Campo Grande	MS	Brazil	10	-20.469056 W	-54.777361 S
10	Dourados	MS	Brazil	20	-22.26276 W	-54.837889 S

ID	Population	Estate*	Country	No. Plants	Geographical coordinates	Geographical coordinates
11	Corumbá	MS	Brazil	22	-19.351261 W	-57.563611 S
12	Formosa	FMS	Argentina	11	-25.232972 W	-58.256639 S

*States: MG= Minas Gerais, SP= São Paulo, MS= Mato Grosso do Sul, FMS= Formosa

Table 2. Characteristics of the five gSSR and eight EST-SSR primers and summary statistics of the genetic diversity at each locus.

Primer Name	Primer sequence (5´-3´)	Motif	TA (°C)	Amplicon size	Na	Pa
Aacu10	F: TGCCACATAGAGTGCTTGCT R: CTACCACATCCCCGTGAGTT	(AG)16	56	168–186	10	3
Aacu12	F: GAATGTGCGTGCTCAAAATG R: AATGCCAAGTGACCAAGTCC	(TC)20	56	190–202	12	3
Aacu26	F: ACTTGCAGCCCCATATTCAG R: CAGGAACAGAGGCAAGTTC	(AC)13(AG)14	56	273–316	14	4
Aacu38	F: TTCTCAGTTTCGTGCGTGAG R: GGGAGGCATGAGGAATACAA	(TC)15	56	316–346	11	2
Aacu45	F: CAGACTACCAGGCTTCCAGC R: TCATCATCGCAGCTTGACTC	(CGAC)5	56	260–284	6	1
gSSR mean					**10.60**	**2.60**
Acro16	F: GTCATATGGCTGGTGAGATT R: GTTCCTTCTCTTGGTGGAAT	(GCC)8	55	270	7	3
Acro20	F: CCACCCTTAAGTTCATCTTCT R: GACTGTTGGTGTTAAGGTTCA	(CCT)8	55	294	5	1
Acro46	F: CAGATTATAGCACAGCTGGAG R: AGTGACTTGAAGCTCATGTTG	(ATC)5	55	398	11	2
Acro64	F: GTATGGATGTCGTCGTTGAT R: GACTATGGTAATGGACCAACA	(CTG)12	55	168	10	2
Acro102	F: GGCTAAGATCATTAATGGGAC R: GGACCATACCAATTTCCTTAG	(CCA)5	55	268	3	0
Acro246	F: GAGAAGGTAAGGTAGACGAGG R: ATGGATCAAGAACCCGAC	(GGC)9	57	275	7	3
Acro280	F: ATCTGAGACTGAAGCTGATGA R: GATCTGCATACATCCATCTGT	(GAA)5	55	204	3	2
Acro278	F: GAAGAGTTTTCCTCTCTGCTC R: AGATGCCCTATTGCTCAAG	(CCG)6	55	311	2	0
EST-SSR mean					**6.0**	**1.63**

Na=number of alleles per locus; Pa= Private alleles; PIC = polymorphic information content; HE = expected heterozygosity; HO = observed heterozygosity

Table 3. Genetic diversity estimates for Acrocomia taxa using 13 microsatellite loci.

Specie	N	Na	Ne	I	Ar	H_O	H_E	f	PA
A. aculeata	74.385	5.308	2.329	0.943	3.880	0.208	0.466	0.650	9
Hybrids	19.923	4.846	2.711	1.083	4.290	0.331	0.538	0.420	2
A. totai	65.077	6.692	3.288	1.211	5.070	0.287	0.551	0.583	16
Total	53.128	5.615	2.776	1.079	4.413	0.275	0.518	0.558	

Mean of different alleles (Na), mean of effective allele (Ne), allelic richness (Ar), Shannon's Index (I), Observed (H_O) and Expected (H_E) Heterozygosity, Fixation index (f) and total number of private alleles (PA).

Table 4. Genetic diversity estimates at the population level using 13 microsatellite loci.

Species	Population	N	Na	Ne	I	Ar	H_O	H_E	f
A. aculeata	**Luz**	15.000	2.692	1.628	0.570	2.110	0.158	0.316	0.659
	Patos Minas	14.154	2.846	1.807	0.682	2.350	0.273	0.401	0.392
	Ibituruna	11.538	3.077	2.144	0.805	2.580	0.244	0.457	0.537
	Rifaina	8.615	2.538	1.954	0.640	2.280	0.203	0.368	0.466
	Brotas	11.154	2.769	1.848	0.636	2.260	0.204	0.361	0.465
	Itapira	13.923	2.154	1.574	0.483	1.840	0.162	0.291	0.561
Hybrids	**Braúna**	11.385	3.923	2.384	0.931	3.090	0.347	0.482	0.306
	Campo Grande	8.538	2.846	2.059	0.719	2.440	0.295	0.406	0.369
A. totai	**Fusquinha**	15.769	4.308	2.328	0.924	2.960	0.274	0.460	0.474
	Dourados	19.000	3.923	2.407	0.891	2.970	0.299	0.464	0.487
	Corumbá	20.923	5.000	2.928	1.088	3.450	0.316	0.545	0.516
	Argentina	9.385	3.615	2.880	0.916	2.880	0.222	0.455	0.576
	All populations	13.282	3.308	2.162	0.774	2.601	0.250	0.417	0.485

Mean of different alleles (Na), mean of effective allele (Ne), allelic richness (Ar), Shannon's Index (I), Observed (H_O) and Expected (H_E) Heterozygosity and Fixation index (f).

Table 5. Matrix of pairwise F_{ST} values (below diagonal) among *A. aculeata A. totai* and Hybrids, based on 13 microsatellites loci.

	A. aculeata	**Hybrids**	A. totai
A. aculeata	0.000		
Hybrids	0.079	0.000	
A. totai	0.099	0.059	0.000

Table 6. Matrix of pairwise F_{ST} values (below diagonal) among populations based on 13 microsatellites.

	Luz	**Patos Minas**	**Ibituruna**	**Rifaina**	**Brotas**	**Itapira**	**Braúna**	**Fusquinha**	**Campo**
Luz	0.000								
Patos Minas	0.148	0.000							
Ibituruna	0.116	0.119	0.000						
Rifaina	0.142	0.120	0.088	0.000					
Brotas	0.117	0.102	0.076	0.067	0.000				
Itapira	0.167	0.150	0.111	0.056	0.077	0.000			
Braúna	0.194	0.146	0.119	0.128	0.125	0.162	0.000		
Fusquinha	0.226	0.168	0.146	0.141	0.161	0.203	0.081	0.000	
Campo Grande	0.235	0.220	0.165	0.203	0.183	0.235	0.154	0.164	0.000
Dourados	0.187	0.180	0.127	0.140	0.132	0.175	0.090	0.056	0.126
Corumbá	0.172	0.176	0.119	0.131	0.148	0.175	0.109	0.064	0.123
Argentina	0.287	0.210	0.213	0.240	0.235	0.283	0.134	0.093	0.212