UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

BRENDA GABRIELA DIAZ HERNANDEZ

GENÔMICA E ESTRUTURA POPULACIONAL DA MACAÚBA [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.], VISANDO SUBSIDIAR À DOMESTICAÇÃO DA ESPÉCIE.

CAMPINAS, SP

2020

BRENDA GABRIELA DIAZ HERNANDEZ

GENÔMICA E ESTRUTURA POPULACIONAL DA MACAÚBA [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.], VISANDO SUBSIDIAR À DOMESTICAÇÃO DA ESPÉCIE.

Tese apresentada ao Instituto de Biologia da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do Título de Doutora em GENÉTICA E BIOLOGIA MOLECULAR, *na Área de* GENÉTICA VEGETAL E MELHORAMENTO

Orientadora: Profa. Dra. Maria Imaculada Zucchi

ESTE ARQUIVO DIGITAL CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELA ALUNA BRENDA GABRIELA DIAZ HERNANDEZ E ORIENTADA PELA PROFA. DRA. MARIA IMACULADA ZUCCHI.

CAMPINAS

2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

D543g

Diaz Hernandez, Brenda Gabriela, 1983-
Genômica e estrutura populacional da macaúba [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.], visando subsidiar à domesticação da espécie / Brenda Gabriela Diaz Hernandez. – Campinas, SP : [s.n.], 2020.

Orientador: Maria Imaculada Zucchi.
Coorientador: Jorge Maurício Costa Mondego.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Óleos vegetais. 2. Genotipagem por sequenciamento. 3. Polimorfismo de nucleotídeo único. 4. Repetições de microssatélites. I. Zucchi, Maria Imaculada. II. Costa Mondego, Jorge Maurício. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Genomics and population structure of Macaúba [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.], aimed the species domestication
Palavras-chave em inglês:
Vegetable oils
Genotyping by sequencing
Polymorphism single nucleotide
Microsatellite repeats
Área de concentração: Genética Vegetal e Melhoramento
Titulação: Doutora em Genética e Biologia Molecular
Banca examinadora:
Maria Imaculada Zucchi [Orientador]
Anete Pereira de Souza
Vera Nisaka Solferini
Sérgio Yoshimitsu Motoike
Erick Mauricio Goes Cordeiro
Data de defesa: 27-11-2020
Programa de Pós-Graduação: Genética e Biologia Molecular

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0001-7787-852
- Currículo Lattes do autor: http://lattes.cnpq.br/6152103895832535

Campinas, 27 de novembro de 2020.

COMISSÃO EXAMINADORA

Profa. Dra. Maria Imaculada Zuchhi

Profa. Dra. Anete Pereira de Souza

Profa. Dra. Vera Nisaka Solferini

Prof. Dr. Sérgio Yoshimitsu Motoike

Prof. Dr. Erick Mauricio Goes Cordeiro

Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa em Genética e Biologia Molecular da Unidade Instituto de Biologia.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão do auxílio à pesquisa ao projeto temático (2014/23591-7) e pelo apoio através da concessão da bolsa de estudos de doutorado (2017/22987-2).

Ao Programa de Pós-Graduação em Genética e Biologia Molecular da UNICAMP, pela oportunidade de realizar o doutorado e pela infraestrutura e recursos humanos.

À minha orientadora, Profa. Dra. Maria Imaculada Zucchi, pela oportunidade de trabalhar ao seu lado, com profissionalismo, confiança e amizade. Obrigada pelas palavras de incentivo, sempre no momento certo, e por todo apoio profissional e pessoal.

Ao meu coorientador, Dr. Jorge Mondego, pelas discussões e por todas as contribuições durante o desenvolvimento do projeto.

Ao Dr. Carlos Colombo, pela disponibilidade da estrutura necessária para o desenvolvimento deste trabalho e por toda contribuição intelectual.

À Profa. Dra. Anete Pereira, por disponibilizar seu laboratório e ao Alessandro Alves e Aline Morais, pela ajuda no desenvolvimento das bibliotecas genômicas.

A todos meus colegas e amigos do Centro de Recursos Genéticos Vegetais do IAC pela convivência dentro e fora do ambiente de trabalho, que tornaram mais ameno este caminho.

Aos colegas e amigos do grupo de Genética e Genômica da conservação, pela convivência e o conhecimento compartilhado.

Ao Diego Wassner, Suelen Vianna, Luiz Berton, Joaquim Azevedo e Anna Neubauer, pelo auxílio nas coletas de material botânico.

A meu filho, pela paciência nos momentos em que não consegui ser a super herói que ele desejava que eu fosse, e ao meu marido, pelo amor, apoio incondicional e incentivo, sempre dizendo “vai dar tudo certo”.

A minha família, pelo apoio e carinho de sempre.

A todos que de forma direta ou indireta, contribuíram para o êxito deste trabalho.
Obrigado!

RESUMO

Acrocomia (Arecaceae) é um gênero com ampla distribuição na América neotropical e ao qual são atribuídas oito espécies. *A. aculeata* é a espécie mais importante devido ao maior potencial em produzir óleo, semelhante ao da palma de óleo ou dendê, porém com a vantagem de poder ser cultivada em áreas com déficit hídrico, característica que tem motivado seu interesse econômico e o crescente aumento das áreas de cultivo comercial. Considerada uma espécie com grau de domesticação incipiente, cultivares comerciais ainda não estão disponíveis e os desafios quanto à consolidação da cadeia produtiva dessa espécie vem merecendo a atenção da comunidade científica nacional e internacional. Neste contexto, tendo em vista a importância da domesticação e melhoramento genético de *A. aculeata* e outras espécies do gênero, os objetivos deste projeto foram: a) analisar e compreender o conteúdo e a estruturação da diversidade genômica por SNPs, visando elucidar as relações genéticas entre as espécies do gênero *Acrocomia* e *A. aculeata* de diferentes países americanos; b) elucidar o status taxonômico entre *A. aculeata* e *A. totai* e identificar a ocorrência de possíveis híbridos interespecíficos, utilizando marcadores de repetição de sequência simples (SSR) e, c) analisar o sistema de cruzamento e a diversidade genética de progênies de polinização aberta (OPP) de *A. aculeata* para prever sua vulnerabilidade genética. Nossos resultados mostraram-se parcialmente congruentes com a atual classificação taxonômica baseada em caracteres morfológicos, recuperando a separação das espécies *A. aculeata*, *A. totai, A. crispa* e *A. intumescens* como grupos taxonômicos distintos. *A. aculeata* apresentou uma subestrutura genética acentuada, revelando dois grandes grupos genéticos, correspondendo a uma divisão norte-sul do equador. Com base na utilização de SNPs e marcadores microssatélites, foi corroborada uma diferenciação genética entre *A. aculeata* e *A. totai* como espécies diferentes, sendo que *A. totai* possui o maior nível de diversidade. Os resultados também mostraram que não houve perda significativa na diversidade genética das OPPs derivadas de seleção artificial em populações naturais. Em conclusão, nossos resultados evidenciaram a aplicabilidade dos marcadores SNPs e SSR como referência para futuros estudos sobre o gênero Acrocomia, além de fornecer uma visão mais global da diversidade genômica de *A. aculeata* e outras espécies de Acrocomia, informações úteis para definir estratégias de conservação e melhoramento genético nos diferentes países onde a espécie ocorre.

Palavras-chave: Óleo vegetal, Genotipagem por sequenciamento, Polimorfismo de nucleotídeo único, Repetições de microssatélites

ABSTRACT

Acrocomia (Arecaceae) is a neotropical genus that is attributed eight species with a wide distribution in America. *A. aculeata* is the most important species because its potential to supply oil with the same production capacity as oil palm, but which can be grown in areas with water deficit, characteristic that has motivated its commercial interest and increased commercial planting. Considered a species with a certain degree of incipient domestication, commercial cultivars are not yet available, and the challenges regarding the consolidation of the productive chain of this species deserve the attention of the scientific community. In this context, in view of the importance of domestication and genetic improvement of *A. aculeata* and other species of the genus, the objectives of this research project were a) to analyze and understand the content and structuring of *A. aculeata* genomic diversity by SNPs, with the objective of elucidating the genetic relationships between species Acrocomia and *A. aculeata* from different American countries; b) to elucidate the taxonomic status between *A. aculeata* and *A. totai* and identify the occurrence of natural interspecific hybrids, using simple sequence repeat markers (SSR) and c) to analyze the mating system and genetic diversity of open-pollinated progenies (OPP) of *A. aculeata* to predict their genetic vulnerability. Our results showed partially congruent with the current taxonomic classification based on morphological characters, recovering the separation of the species *A. aculeata*, *A. totai*, *A. crispa* and *A. intumescens* as distinct taxonomic groups. *A. aculeata* showed an accentuated substructure within, showing two genetic groups, corresponding to a north– south split. Based on both using SNPs and microsatellites markers, was corroborated as a genetic differentiation between *A. aculeata* and *A. totai* as different species, and that *A. totai* has the highest level of diversity. The results also showed that there was no significant loss in the genetic diversity of OPPs derived from an artificial selection in natural populations. In conclusion, our results evidenced the applicability of SNPs and SSR markers as a reference for future studies on the Acrocomia genus, as well as provide a more global view of the genomic diversity of *A. aculeata* and another species of the genus, information that will be useful to define strategies for conservation and breeding in the different countries where the species occur.

Keywords: Vegetable oil**,** Genotyping by Sequencing, Single-nucleotide polymorphism, Simple sequence repeat

SUMÁRIO

INTRODUÇÃO GERAL

As palmeiras são um grupo de plantas econômica e ecologicamente muito importantes, pois fornecem uma grande variedade de produtos. Entre os quais estão produtos alimentícios obtidos dos frutos, sementes, palmito, bebidas obtidas da seiva e frutas, fibras, óleo, ceras, contribuindo assim para a alimentação, moradia, energia, uso medicinal e vestimenta (Rivas et al., 2012). Tal é o caso de espécies do gênero Acrocomia que possuem múltiplos usos.

A palmeira macaúba [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.] é a espécie de maior interesse econômico do gênero. É uma planta perene, nativa da América tropical e subtropical, com ampla distribuição geográfica, ocorrendo desde o norte do México e Antilhas até o Sul do Brasil (Henderson et al., 1995, Lorenzi et al., 2010). No Brasil, é considerada como a palmeira de maior dispersão (Lorenzi 2006).

Na última década a macaúba vem sendo apontada em vários países onde ela ocorre como palmeira de alto potencial bioeconômico em razão da variedade de aplicações e produtos obtidos, tendo praticamente todas as partes aproveitadas, desde a alimentação humana e animal, bioenergia até o uso medicinal (Lorenzi et al. 2006; Ramos et al., 2008). O fruto é a parte mais utilizada da planta, uma vez que pode ser totalmente aproveitado, a polpa é consumida in natura, como farinha ou usada para extração de gordura comestível (Lorenzi et al. 2006;). Tanto o óleo da polpa como o óleo da amêndoa são muito semelhantes ao do dendê em termos de composição, o óleo da polpa é rico em ácido oleico, o que confere o título de “high oleic”, com alto valor, grande demanda na indústria alimentícia e, sobretudo, energética, em função da maior estabilidade à oxidação e tolerância a baixas temperaturas (Silva et al., 1986, Colombo et al., 2018). O teor de óleo extraído da amêndoa é rico em ácidos graxos saturados de cadeia curta, especificamente o ácido láurico (Coimbra e Jorge, 2011; Bora e Rocha, 2004), constituindo-se em uma valiosa matéria prima para fabricação de cosméticos, produtos de saponificação e bioquerosene. Além da qualidade dos óleos, sua produção pode atingir até 6 mil litros de óleo.ha.ano^{-1}, superando em muito a soja que produz apenas 400 litros.ha.ano^{-1} e igualando-se ao dendê, considerado a oleaginosa de maior rendimento de óleo por área com uma produção de até 6 mil litros de óleo.ha.ano-1 (César et al., 2015).

Além das vantagens econômicas, por ser uma espécie perene, a macaúba apresenta importantes benefícios ambientais, como o manejo e a conservação de solo, recuperação de áreas degradadas, além de mobilizar CO_2 atmosférico, minimizando assim os impactos negativos causados pela emissão de gases (Cargnin et al., 2008) e, como espécie

pioneira, abre caminho para a sucessão de outras espécies, assegurando a conservação e recuperação ambiental (Moura, 2007). A macaúba apresenta também adaptação a solos arenosos e com baixo índice hídrico (Costa e Marchi, 2008; Lorenzi et al., 1996; Motta et al., 2002).

Apesar de ser amplamente utilizada em diversos países, a macaúba é uma planta que se encontra em um estado insipiente de domesticação. Atualmente, não existem cultivares disponíveis para cultivo comercial e sua exploração é de forma local e extrativista. No entanto, as vantagens econômicas da espécie nos últimos anos e o interesse comercial aumentam vertiginosamente em muitos locais, havendo necessidade de se disponibilizar urgentemente materiais para plantio com o mínimo de informações agro técnicas.

O conhecimento da diversidade genética e estrutura populacional da espécie é uma premissa crucial para um adequado manejo e aproveitamento i*n situ* e *ex situ* dos recursos genéticos. Em culturas emergentes e de domesticação insipiente, como é o caso da macaúba, análises sobre diversidade genética são cruciais para orientar a seleção de materiais mais promissores e com caracteres de interesse agronômico, maximizar os ganhos genéticos e a criação de cultivares comerciais de forma mais eficaz, bem como, fornecer informações básicas da variação genética intra e inter específica para outras áreas de estudo.

Estudos prévios usando marcadores moleculares têm avaliado a diversidade genética de *A. aculeata*, porém focado em populações brasileiras e, sobretudo, oriundas dos estados de São Paulo e Minas Gerais (Abreu et al., 2012; Lanes, et al., 2015; Mengistu 2015; Nucci, 2007; Silva 2013; Oliveira et al. 2012), revelando um panorama limitado da diversidade da espécie. Portanto, devido à a ampla distribuição geográfica da espécie (Américas Tropical e Subtropical), existe a necessidade de se conhecer e explorar sua diversidade e estrutura genética populacional em toda sua área de ocorrência e sua relação com outras espécies do gênero.

Dentro de *Acrocomia* um aspecto importante a ser elucidado é a situação taxonômica das espécies, haja vista não existir consenso na classificação das espécies para o gênero. Para Henderson et al. (1995) só existem duas espécies dentro gênero: *A. aculeata*, (Jacq.) Lodd. ex Mart. e *A. hassleri* (Barb. Rodr.) W.J. Hahn. Já para Lorenzi (2010) o gênero é representado por sete espécies, *A. aculeata* (Jacq.) Lodd. ex Mart., *A. intumescens* Drude, *A. totai* Mart., *A. crispa* (Kunth), de porte arbóreo, e *A. hassleri*, *A. glaucescens* Lorenzi e *A. emensis*, de porte baixo. Assim mesmo, os sites The Plant List (2020) e The Palmweb (2020) também reconhecem a existência de *A. media* O.F. Cook, endêmica de Poerto Rico.

Das espécies, *A. aculeata* e *A. totai*, são as que apresentam maior dificuldade na sua distinção taxonômica. Esta falta de consenso pode ser devido à plasticidade dos atributos morfológicos, à grande variação das características adotadas para diferenciação das espécies (Crocomo e Melo, 1996), ausência de informação sobre os padrões de variações morfológicas intraespecíficas, bem como possível ocorrência de hibridação interespecífica. Para fins aplicados ao melhoramento genético, a elucidação das relações taxonômicas entre as supostas espécies *A. aculeata* e *A. totai* permitirá estabelecer estratégias de melhoramento genético próprios de cada espécie. Ademais, a comprovação de hibridização interespecífica pode ser uma ferramenta para aumentar a diversidade genética, assim como gerar novos genótipos com características complementares e de interesse agronômico.

Se bem o conhecimento da diversidade da espécie é ponto de partida para atividades de gestão dos recursos genéticos, entender os impactos do melhoramento genético na diversidade e os fatores que o influenciam, sendo de fundamental importância para definir estratégias de seleção através dos ciclos de melhoramento, maximizar os ganhos genéticos e garantir a manutenção da diversidade genética para mitigar a vulnerabilidade da espécie causada por fatores bióticos e abióticos.

Considerando o potencial econômico de *A. aculeata*, no presente projeto de doutorado, analisamos a diversidade genética e genômica de *A. aculeata* em populações naturais, sua relação com outras espécies do gênero e o sistema reprodutivo, utilizando marcadores microssatélites (SSR) e polimorfismo de nucleotídeo único (SNP).,

A tese foi estruturada em três estudos apresentados no formato de manuscrito científico, tendo por objetivo auxiliar os programas de melhoramento e planejar estratégias de manejo *in situ* e *ex situ* para a conservação da espécie nos diferentes países onde ela ocorre, assim como fornecer informações básicas para outras áreas de estudo. Além disso foi possível analisar a diversidade genética de *A. aculeata* e outras espécies do gênero, visando auxiliar sua domesticação.

Sabendo da importância dos estudos de diversidade genética para diferentes áreas e para o desenvolvimento de estratégias de melhoramento genético, de manejo e conservação *in situ* e *ex situ* da espécie, o primeiro capítulo teve como objetivo caracterizar a diversidade e estrutura genética populacional de sete espécies de *Acrocomia*, com ênfase em *A. aculeata* abrangendo uma ampla área de distribuição natural da espécie. Para tal fim foram utilizados marcadores de SNPs.

No segundo capítulo, considerando as dificuldades para diferenciar morfologicamente as espécies *A. aculeata* e *A. totai* e a ausência de informações sobre variabilidade genética em populações naturais, foram utilizados marcadores microssatélites para elucidar a relação taxonômica entre *A. aculeata* e *A. totai* a nível molecular, identificar a ocorrência de híbridos interespecíficos e comparar a variação genética dentro e entre as populações analisadas. As informações obtidas neste estudo permitiram definir critérios de seleção de germoplasma para fins de melhoramento genético específicos para cada espécie.

O terceiro capítulo teve como objetivo avaliar o sistema de cruzamento de *A. aculeata*, caracterizar sua taxa de cruzamento e/ou autofecundação a nível populacional e comparar os efeitos da seleção na diversidade genética de progênies de polinização aberta em relação às populações nativas. Os resultados deste capítulo permitiram melhorar nossa compreensão sobre a taxa de cruzamento em *A. aculeata*. e forneceram informações importantes sobre os impactos da intensidade da seleção na diversidade genética das populações melhoradas.

Gênero Acrocomia

O gênero Acrocomia é representado por espécies que ocorrem unicamente em áreas tropicais e subtropicais do continente americano. Devido à grande variação fenotípica que algumas das espécies do gênero apresentam nas diferentes regiões onde se distribui, têm sido descritas 46 espécies (Mobot, 2018), a maioria atualmente consideradas como sinonímias de *A. aculeata* (The Plant List 2013). Para Henderson et al. (1995), apenas duas espécies são atribuídas ao gênero: *A. aculeata* e *A. hassleri*, sendo a primeira de porte arbóreo e amplamente distribuída ao longo das Américas Central e do Sul, enquanto a segunda é rizomatosa, de pequeno porte e restrita às áreas do Cerrado do Brasil e parte do Paraguai. No entanto, de acordo com Lorenzi (2010), o gênero é representado por sete espécies diferenciadas entre si principalmente pela altura das plantas, das quais seis são encontradas no Brasil: *A. aculeata* (Jacq.) Lodd. ex Mart., *A. intumescens* Drude e *A. totai* Mart. apresentam porte arbóreo,enquanto *A. hassleri*, *A. glaucescens* Lorenzi e *A. emensis* são de porte baixo. Além das setes espécies aceitas por Lorenzi et al., (2010) os sites The Plant List (2020) e The Palmweb (2020), também consideram a existência de *Acrocomia media*, espécie endêmica de Porto Rico e Ilhas Virgens.

***Acrocomia aculeata* (Jacq.) Lodd. ex Mart.**

Atualmente, *A. aculeata* é a espécie de maior interesse econômico do gênero. É uma espécie diploide (2n = 30), com um genoma de 2.8 Gb (Abreu et al., 2011). É uma palmeira

perene, heliófila, de porte alto, podendo atingir de 10 a 15 metros de altura na idade adulta. O estipe é cilíndrico, espinescente, anelado, coberto pelos remanescentes dos pecíolos foliares. As folhas, que se dispõem aglomeradas no ápice do estipe, são compostas, pinadas e pecioladas, com folíolos dispostos alternadamente ao longo da raque. É uma espécie monoica e com um sistema reprodutivo misto (Scariot et al. 1995; Scariot et al. 1991: Abreu et al. 2012; Lanes et al. 2016). Apresenta inflorescência interfoliar e ramificada. As flores pistiladas se localizam na base das ráquilas, sempre formando tríades; no entanto, as flores estaminadas se encontram nos dois terços superiores da ráquila (Lorenzi et al. 2010). Fruto drupáceo, globoso, com 3,0 a 5,0 cm de diâmetro (Lorenzi et al., 2010; Vianna et al., 2017), o epicarpo é lenhoso, verde-amarelado e liso; mesocarpo variando de cor creme a alaranjado, fibro-mucilaginoso; endocarpo duro, fortemente aderido ao mesocarpo. A semente é adnata ao mesocarpo, revestida por um fino tegumento (Lorenzi et al., 2010). O fruto é rico em ácidos graxos, contendo até 75% de óleo no mesocarpo e 65 % na amêndoa (Bora & Rocha 2004; Hiane et al., 2006; Ciconini et al., 2013).

A espécie apresenta ampla distribuição geográfica, ocorrendo na América tropical e subtropical, desde o norte do México (CONABIO 2020) até o sul do Brasil, com exceção do Peru e Equador (Henderson et al. 1995; Lorenzi et al. 2010; Vianna e Campos-Rocha, 2020). No Brasil, é considerada a espécie de maior ocorrência, apresentando ampla distribuição e presente nos Estados do Pará, Maranhão, Ceará, Minas Gerais, Goiás, Mato Grosso e São Paulo, com possível ocorrência nos estados de Mato Grosso do Sul, Paraná, Tocantins, Rondônia, Roraima, Espírito Santo e Pernambuco (Scariot et al., 1991; 1995, Lorenzi et al., 2010, Vianna e Campos-Rocha, 2020). Por ser uma espécie heliófila, sua ocorrência está fortemente associada às áreas abertas e às áreas de pastagem. Se distribui principalmente em áreas de Cerrado, cerradão e bosques semi-caducifólios (Almeida et al. 1998; Lorenzi et al. 2010), porém sua ocorrência natural também tem sido registrada em áreas de Floresta Tropical e Subtropical e na Caatinga, tanto em áreas altas como ao nível do mar (Vianna & Colombo, 2013), embora se desenvolva melhor em baixas altitudes (entre 500 a 1000 m). No entanto, tem sido encontrada em altitudes de até 1600 m e ao nível do mar (Observação pessoal).

***Acrocomia totai* Mart.**

Palmeira arborescente de até 15 m de altura, perene, heliófila, com estipe único, cilíndrico, liso, desprovido dos remanescentes da bainha das folhas, com presença de espinhos unicamente na fase juvenil das plantas. Folhas cobertas por espinhos, apresentando abscisão nítida, pinas irregularmente distribuídas e inseridas em planos diferentes. Inflorescência

interfoliar com flores pistiladas localizadas somente na base das ráquilas, sempre formando tríades e com flores estaminadas localizadas nos dois terços superiores das ráquilas. Frutos globosos, de 2,5 a 3,5 cm de diâmetro, epicarpo castanho-amarelado e mesocarpo fibro-mucilaginoso (Lorenzi *et al*., 2010). Os frutos são ricos em lipídeos, com um teor de óleo na polpa variando de 26 a 33% (Vianna et al. 2015, Hiane et al. 2005).

A espécie também apresenta ampla distribuição geográfica, porém ocorre unicamente na América do Sul, no nordeste da Argentina, no leste da Bolívia, no Paraguai e no Estado brasileiro do Mato Grosso do Sul (Markley, 1956; Rodríguez e Aschero, 2005, Lorenzi et al., 2010), embora com possível ocorrência no Paraná, Rio Grande do Sul, São Paulo e Tocantins (Lorenzi et al., 2010 e Vianna e Campos-Rocha, 2020). A espécie prefere terrenos temporariamente alagados de várzeas úmidas, ocorrendo eventualmente em áreas mais secas. Segundo Markley (1956), *A. totai* é a única das espécies do gênero nativa de zona mais temperada.

***Acrocomia media* O.F. Cook**

A espécie é endêmica de Porto Rico e das Ilhas Virgens, ocorrendo em habitats moderadamente secos a úmidos, principalmente em altitudes inferiores a 200 m (Proctor 2005).

Esta palmeira tem uma altura que geralmente varia entre 8 a 10 m, podendo chegar a 15 m. Apresenta estipe solitário, cilíndrico, anelado, um pouco mais grosso na parte superior ou ligeiramente fusiforme, com espinhos negros que podem ser deiscentes e deixar o tronco liso. As folhas são numerosas, cada uma com até 4 metros de comprimento, com pecíolo curto e densamente espinescente e com folíolos acuminados ou atenuados no ápice. A espata interna possui até 60 cm de comprimento, sendo densamente pubescente, marrom, com presença de alguns espinhos na superfície exterior. A inflorescência possui cerca de 1,5 m de comprimento, com pedúnculo espinescente, curto e grosso. Ráquilas com 10 a 25 cm de comprimento. Na ráquila, as flores estaminadas se localizam na parte apical e as flores pistiladas na parte basal. Os frutos são achatados, globosos, com diâmetro entre 3,5 a 4,5 cm, apresentando epicarpo liso, sem brilho e amarelado quando maduro (Bailey, 1941; Proctor, 2005).

***Acrocomia intumescens* Drude**

Esta palmeira apresenta um intumescimento no terço médio do estipe. As folhas são pinadas, com folíolos dispostos em intervalos regulares ou agrupados em vários planos de inserção, com bainha e pecíolo coberto por espinhos. As inflorescências são ramificadas, com flores pistiladas na base das ráquilas formando tríades e flores estaminadas imersas em alvéolos

profundos nos dois terços superiores das ráquilas. Os frutos são globosos a subglobosos, com 3,7 a 5,5 cm de comprimento e 3,8 a 5,4 cm de largura. O epicarpo é lenhoso, coberto por tomento castanho, brilhante quando maturo (Lorenzi *et al*., 2010).

A espécie é endêmica do nordeste brasileiro, se distribui na Mata Atlântica ao longo da costa nordeste brasileira e nas áreas dispersas de extensões florestais na Caatinga, com registros ocorrendo nos Estados de Alagoas, Pernambuco, Paraíba, Bahia e Ceará, e no centro endêmico do rio São Francisco (Lorenzi et al. 2010; Vianna e Campos-Rocha, 2020).

Acrocomia crispa

A espécie apresenta um estipe único, cilíndrico, anelado, com um intumescimento próximo à copa, coberto por espinhos grandes e de coloração castanho-escuro. A planta pode alcançar entre 5 a 7 m de altura. As folhas são verde-acinzentadas com espinhos brancos. Os frutos são globosos, com 2 a 3 cm de diâmetro, amarelos quando maduros, com mesocarpo seco e fibroso (Bailey, 1941; Henderson *et al*., 1995).

A espécie é endêmica de Cuba, com ampla distribuição na ilha, ocorrendo principalmente em áreas secas de planície, áreas de pastagem e campo aberto (Bailey, 1941; Borhidi, 1996), em áreas com solos calcários (Dransfield et al., 2008).

Acrocomia glaucescens Lorenzi

É uma palmeira solitária que apresenta duas formas de crescimento, uma com estipe subterrâneo e outra caulescente. O estipe é cilíndrico ou irregular, frequentemente engrossado na parte mediana. As folhas são verde-azuladas, raque foliar com pinas inseridas quase regularmente, em diferentes planos. A bráctea é peduncular e a inflorescência é ramificada. Apresenta fruto globoso, verde-amarelado, com diâmetro entre 1.5 a 2.6 cm de diâmetro. Sua ocorrência tem sido reportada nos Estados de Goiás e Mato Grosso do Sul, em vegetação de cerrado e sobre solo arenoso (Lorenzi et al.,2010; Martins, 2012).

Acrocomia emensis (*Toledo*) Lorenzi.

Esta palmeira solitária possui porte arbustivo, com 40 a 60 cm de altura. É rasteira, acaulescente e espinescente. Apresenta rizoma subterrâneo alongado, às vezes coberto por alguns espinhos. As folhas são pinadas e arqueadas, com folíolos irregularmente

distribuídos e dispostos em vários planos. As folhas são fixadas ao rizoma e emergem ao nível do solo. As inflorescências são dispostas em tríades, com flores estaminadas nos dois terços superiores e imersas em alvéolos. Os frutos são globosos, pequenos e com epicarpo coriáceo (Lorenzi *et al*., 2010).

Esta espécie é característica do Cerrado *lato sensu* e ocorre nos Estados de Goiás, Mato Grosso do Sul, Minas Gerais, São Paulo e Paraná (Leitman et al., 2012; Vianna e Campos-Rocha, 2020). A espécie é considerada em risco de extinção (Lorenzi et al., 2010).

Acrocomia hassleri (Barb. Rodr.) W.J. Hahn

Esta espécie é heliófila, acaulescente, terrícola, solitária, com altura entre 30 a 40 cm, com estipe levemente espinescente, apresentando rizoma com cerca de 30 cm abaixo da superfície. Apresenta entre 4 a 7 folhas contemporâneas, verde-azuladas, com pinas lineares, com bainha e pecíolo espinescentes e folíolos distribuídos regularmente em dois planos distintos formando um "V". As inflorescências são ramificadas, localizadas ao nível do solo, apresentando entre três a cinco ráquilas. A bráctea é peduncular, marrom avermelhada e pubescente. O fruto é uma drupa globosa, com aproximadamente 3 cm de diâmetro (Leitman *et al*., 2012; Lorenzi *et al*., 2010; Rolón et al., 2017). A espécie ocorre no Brasil em áreas de cerrado aberto sobre solos arenosos, principalmente no Estado do Mato Grosso do Sul, e também é encontrada no Paraguai (Lorenzi *et al*., 2010; Rolón et al., 2017).

CAPÍTULO I -Whole-genome SNP analysis elucidates the genetic population structure and diversity of Acrocomia species.

Brenda G. Díaz, Maria I. Zucchi, Alessandro. Alves-Pereira, Caléo P. de Almeida, Aline C. L. Moraes, Suelen A. Vianna, Joaquim A. Azevedo-Filho, Carlos A Colombo

Artigo submetido na revista PLOS ONE

Whole-genome SNP analysis elucidates the genetic structure and diversity of Acrocomia species

B. G. Díaz[1], M. I. Zucchi[1,3], A. Alves-Pereira[1], C.P. de Almeida[3], A. C. L. Moraes[1], S. A. Vianna[3], J. A. Azevedo-Filho[2], C. A Colombo[3*]

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

[2] Polo Regional Leste Paulista, Secretaria da Agricultura do estado de São Paulo Monte Alegre do Sul-SP, Brazil.

[3] Centro de Pesquisa e Desenvolvimento de Recursos Genéticos Vegetais, Instituto Agronômico -IAC, Campinas-SP, Brazil,

* Corresponding author

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

Abstract

Acrocomia (Arecaceae) is a genus widely distributed in tropical and subtropical America that has been achieving economic interest due to the great potential of oil production of some of its species. In particular *A. aculeata,* due to its vocation to supply oil with the same productive capacity as the oil palm even in areas with water deficit. Although eight species are recognized in the genus, the taxonomic classification based on morphology and geographic distribution is still controversial. Knowledge about the genetic diversity and population structure of the species is limited, which has limited the understanding of the genetic relationships and the orientation of management, conservation, and genetic improvement activities of species of the genus. In the present study, we analyzed the genomic diversity and population structure of seven species of Acrocomia including 117 samples of *A. aculeata* covering a wide geographical area of occurrence, using single nucleotide Polymorphism (SNP) markers originated from Genotyping By Sequencing (GBS). The genetic structure of the Acrocomia species were partially congruent with the current taxonomic classification based on morphological characters, recovering the separation of the species *A. aculeata*, *A. totai*, *A. crispa* and *A. intumescens* as distinct taxonomic groups. However, the species A. media was attributed to the cluster of *A. aculeata* while *A. hassleri* and *A. glauscescens* were grouped together with *A. totai*. The species that showed the highest and lowest genetic diversity were A. totai and A. media,

respectively. When analyzed separately, the species A. *aculeata* showed a strong genetic structure, forming two genetic groups, the first represented mainly by genotypes from Brazil and the second by accessions from Central and North American countries. Greater genetic diversity was found in Brazil when compared to the other countries. Our results on the genetic diversity of the genus are unprecedented, as is also establishes new insights on the genomic relationships between Acrocomia species. It is also the first study to provide a more global view of the genomic diversity of *A. aculeata*. We also highlight the applicability of genomic data as a reference for future studies on genetic diversity, taxonomy, evolution and phylogeny of the Acrocomia genus, as well as to support strategies for the conservation, exploration and breeding of Acrocomia species and in particular *A. aculeata*.

Keywords *Acrocomia aculeata*, macaúba palm, coyol palm, taxonomy, genetic resources, Genotyping by Sequencing, species conservation, domestication.

Introduction

The genus Acrocomia is endemic to tropical and subtropical America. This genus is one of the most taxonomically complex concerning species in the family Arecaceae [1]. Taxonomic classifications of Acrocomia are mostly limited to the description of species based on morphological and geographical distribution information. However, extensive morphological plasticity, especially for species with wide geographical distribution, has hindered the taxonomic resolution of species. Since the description of the genus Acrocomia by Martius in 1824 [2], many species have been included and removed from the genus. From the most recent classifications, Henderson et al. [3] attributed only two species to the genus. One is *A. aculeata* (Jacq.) Lodd. ex Mart., which is large (arboreal) and widely distributed throughout Central, North, and South America. The other is *A. hassleri* (Barb. Rodr.) WJ Hahn, which is small in size and is restricted to the Cerrado savanna in Brazil and part of Paraguay. Lorenzi et al. [4] recognized seven species for the genus. Six of these are found in Brazil: *A. aculeata*, *A. intumescens,* and *A. totai* have an arboreal size and are mainly differentiated by the stipe characteristics. *A. hassleri, A. glaucescens*, and *A. emensis* are small size and are differentiated by their height. The seventh species, *A. crispa*, has an arboreal size and is endemic to Cuba. The Plant List [5] and The Palmweb [6] recognized *A. media* as the eighth species. It is endemic to Puerto Rico. Therefore, the systematics of the genus Acrocomia remain controversial, with

the number of species not well resolved and very few studies having addressed species delimitation, population genetic diversity and structure, and inter-species relationships.

A. aculeata, A. totai, and *A. intumescens* are the species of greatest economic interest, mainly due to their many applications and products obtained, with practically all parts of the palms used. The fruits are important for the production of vegetable oil as a bioenergy source and flour for human and animal consumption [7] as well as for medicinal uses [7, 8]. Of these three species, *A. aculeata* is distinguished by its high productive capacity and oil quality [9]. The oil production of 4,000 oil L/ha/year estimated in Brazil far surpasses soybeans (400 L/ha) [9] and equals the oil palm, which is considered the oilseed with the highest oil yield per area, with an oil production volume of up to 6,000 L/ha [10, 11].

A. aculeata is an arborescent heliophile and monoecious. This species produces unisexual flowers in the same inflorescence [3, 4]. It has a mixed reproductive system, with a preference for allogamy [12]. It is a diploid species (2n = 30), with a genome size of 2.8 Gbp [13]. *A. aculeata* has a wide geographic distribution, occurring naturally from northern Mexico and the Antilles to southern Brazil [3, 4, 14, 15]. It is commonly found in savanna areas, but also is found in tropical and subtropical forests, and in the dry forests of Caatinga [3, 4, 16] and has adapted to sandy soils and regions with low water availability [17]. Besides being a perennial species, it is beneficial for soil management and conservation since its useful life can exceed 50 years. Colombo et al. [9] identified *A. aculeata* as a promising resource for sustainable large-scale production of vegetable oil.

Although the economic interest in some Acrocomia species is growing, little is known about infrageneric relationships, levels of genetic diversity and structure, and patterns of gene flow at the genus level. The population genetics approach can assist in species delimitation and provide reference information on the genetic diversity and structure within and between species. Such knowledge is essential for more efficient management and economic exploration of the species and can guide strategies for domestication and conservation of these genetic resources. *A. aculeata* is an emerging crop with incipient domestication. The analysis of genetic diversity of *A. aculeata* is crucial to guide the selection of the most promising materials for crop use, to maximize genetic gains, and to more effectively contribute to the creation of commercial cultivars.

In this context, molecular markers have been broadly adopted in plants as an essential tool to investigate genetic diversity in ecological, phylogenetic, and evolutionary studies. In

addition, they have been widely used for direct management, conservation, and genetic breeding of several species [18]. More recently, next-generation sequencing (NGS) has facilitated the identification of single nucleotide polymorphisms (SNPs), which have emerged as the most extensively used genotyping markers due to their abundance and distribution in the genome. The use of SNPs has considerably expanded knowledge of the genetic diversity of genomes of various plant species [19] at low cost and without the need for reference genomes [20-22]. However, SNPs have not been used as markers in genetic studies of Acrocomia species.

In *Acrocomia,* microsatellites or simple sequence repeats (SSR) have been the most used molecular markers, with the main objective of evaluating the genetic diversity and structure of natural populations and germplasm banks [12, 23-26]. Other approaches include the use of internal transcribed ribosomal 18S-26S spacer (ITS region) [27] and random amplification of polymorphic DNA (RAPD) markers [28]. However, most studies have focused on *A. aculeata* [12, 23, 25, 26, 29]. Only one study has analyzed the genetic diversity of *A. totai* (Lima et al., 2020).

Considering the wide distribution of *A. aculeata* in the Americas, all the studies carried out using molecular markers have revealed a limited panorama of species genetic diversity because they considered a very small geographic sampling, with genotypes obtained mainly from the states of São Paulo and Minas Gerais in Brazil (Abreu et al., 2012; Lanes et al., 2015; Mengistu 2016; Oliveira et al., 2012; Coelho et al., 2018). Only a single study has evaluated the genetic diversity of natural populations of *A. aculeata* (termed *A. mexicana)* from another country besides Brazil, that being Mexico [30].

Faced with the difficulties of taxonomic resolution of the genus Acrocomia, our study aimed to apply a population genomic approach to elucidate the genetic diversity and genetic relationships of species through genomic polymorphism data. Recognizing the increasing importance of *A. aculeata* and the lack of genetic reference data in the different countries where it grows, we analyzed the genetic diversity and structure of natural populations of *A. aculeata*, considering its wide occurrence in the American continent.

The present study is unprecedented because it was conducted using seven Acrocomia species and a wide sampling of *A. aculeata* from several countries in the American continent. This is the first study carried out with SNP markers for the genus.

Material and Methods

Plant material and DNA extraction

In the present study, we considered 172 samples to represent seven from eight Acrocomia species: *A. aculeata, A totai, A. intumescens, A. media, A. crispa*, *A. hassleri*, *A. glaucescens*. The samples were obtained from different locations in order to represent the entire geographic distribution described in the literature for the respective species [3, 4]. The species *A. aculeata*, with a greater distribution in America, was represented by samples from five countries (Fig 1 and S1 Table).

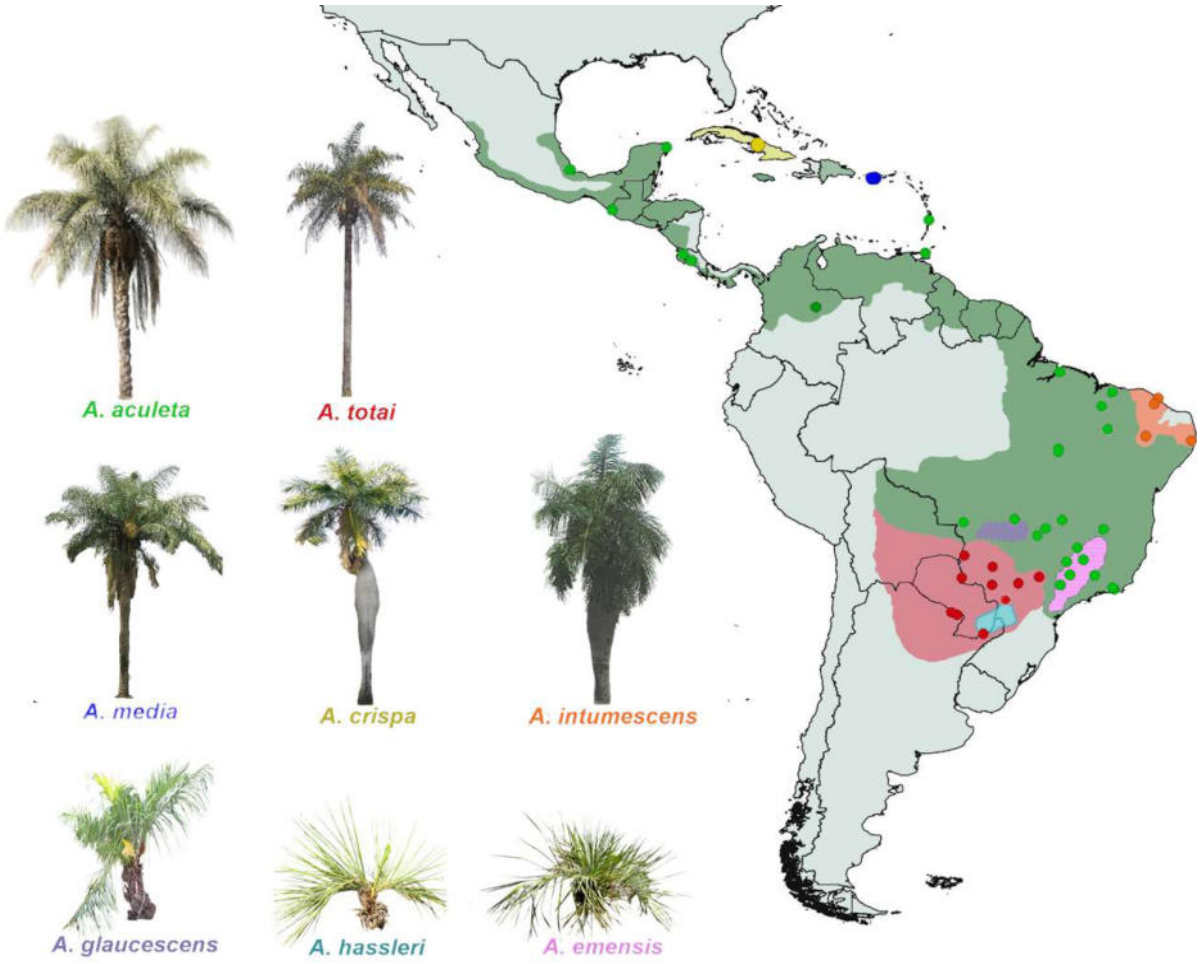


Fig 1. Schematic map of Acrocomia species distribution and geographic location and origin of samples. Data used to generate the species distribution (Colored shading) are based on occurrence record data from GBIF (Global Biodiversity Information Facility www.gbif.org) and Lorenzi et al., [4]. Circles represent geographical location and origin of samples in this study. Image sources: *A. aculeata, A. totai, A. hassleri, A. glaucescens. A. emensis* (B. G. Díaz); *A. intumescens*; *A. media* and *A. crispa* (S. A. Vianna.)

The total genomic DNA was extracted from leaf material using the Doyle & Doyle [31] protocol. We evaluated the quality and quantity of DNA on a 1% agarose gel, on the NanoVue ™ Plus spectrophotometer (GE Healthcare), and through fluorescence using the QubitTM dsDNA BR Assay (Qubit - Life Technologies). Based on the obtained reading, we standardized the DNA to a concentration of 30ng.µl-1.

GBS library preparation and High-Throughput Sequencing

To obtain SNPs, we developed genomic libraries using the GBS (Genotyping by Sequencing) technique according to the protocol described by Poland et al. [32], with modifications. We digested 7 µl of the genomic DNA [30ng.µl-1] from each sample at 37 ° C for 12 hrs with the enzymes *NsiI* and *MspI*. Subsequently, 0.02 µM of specific adapters for the Illumina technology (containing the barcode sequences and complementary to the Illumina ™ primers for sequencing) were connected to the fragments ends generated in the digestion. The ligation reaction was carried out at 22°C for 2 h; 65°C for 20 min; 10°C indefinitely.

After adapters ligation, we purified the samples using QIAquick PCR Purification Kit (Qiagen). The library was enriched by PCR. We performed eight replicates, each one containing 10 µL of purified and amplified ligation, using 12.5 µL of Phusion® High-Fidelity PCR Master Mix NEB (New England Biolabs Inc.), and 2 µl of Illumina forward and reverse [10 µM] primers ™, in a final volume of 25 µL, using the following amplification program: 95°C for 30 s, followed by 16 cycles of 95°C for 10 s, 62°C for 20 s, 72°C for 30 s, ending at 72°C for 5 min. Finally, we purified the library using QIAgen's QIAquick PCR Purification Kit.

The verification of average size of the DNA fragments using the Agilent DNA 12,000 kit and the 2100 Bioanalyzer System (Agilent) equipment. The libraries were quantified by qPCR using the CFX 384 real-time thermocycler (BioRad) with the aid of the KAPA Library Quantification kit (KAPA Biosystems). We prepared two libraries of 96 samples each, which were sequenced using Illumina's NextSeq 500/550 Mid Output Kit v2.5 (150-cycle), on the NextSeq550 platform (Illumina Inc., San Diego, CA).

SNP identification

We performed the identification of SNP markers using the Stacks *v.* 1.42 *pipeline* [33]. We used the *process_radtags* module to demultiplex the samples and to remove the low-quality

readings. As there is no reference genome for Acrocomia, we aligned the sequences and organized the loci using the *ustacks* module with the following parameters: the minimum sequencing depth (-m) ≥ 3, the maximum distance between *stacks* (-M) = 2; and the maximum distance between primary and secondary sequences (-N) = 2. Subsequently, a locus catalog was built using the *cstacks* module*,* allowing a maximum of 2 differences between stacks (-n) from different individuals. We eliminated loci with lower values of probability (lnl_lim -10) by the *rxstacks* correction module. The SNPs were filtered using the *populations* module*,* retaining only one SNP per sequence, with a minimum depth of 3X sequencing, minor allele frequency ≥ 0.01, and minimum occurrence in 75% of individuals in each location/population. After filtering, we identified 3269 SNPs (S1 File).

Identification of putatively neutral and under selection loci

We identified neutral SNPs and loci putatively under selection (outliers). To reduce the possibility of identifying false positives, we applied three approaches to identify outlier loci. For the first approach, we used the method based on Principal Component Analysis (PCA) from the *pcadapt* package [34], on the R platform [35]. The *pcadapt* method assumes that SNPs excessively related to the population structure are candidates to be under adaptive selection. In this approach, no a priori information about the number of populations was introduced. Initially, we carried out the principal component analysis (PCA) to define the structure of the data set, adopting the Mahalanobis distance from the z-scores in the first k-components of each locus to identify the most related loci to the population structure. In the second approach, we used the *fsthet* package [36] based on Wright's F_{ST} fixation index [37] to identify the loci with deviation from the expected relationship between F_{ST} and heterozygosity (H_E), using the island migration model [38].

The third approach we adopted to test the association of environmental variables with the genetic variation of SNP markers was the LFMM (Latent Factor Mixed Models) [39], using the LEA package (Landscape Genomics and Ecological Association Test) [40] on R. platform [35]. We used nineteen bioclimatic variables related to precipitation and temperature, in addition to the minimum, average and maximum values of wind speed, vapor pressure, and solar radiation, obtained from the WorldClim database [41]. We performed the analyzes with the following variables (correlation ≤ 0.8): average annual temperature, average daytime variation, isothermality, average temperature of the wettest four-month period, annual

temperature variation, annual precipitation, precipitation in the driest month, precipitation seasonality, radiation maximum solar radiation, minimum solar radiation, and average wind speed. For the lfmm function, five replicates were performed with 200,000 MCMC interactions after 50,000 burn-ins. For the association tests, the genetic structure presented between the individuals was considered with the SNMF analysis [39], determining the most likely number of genetic groups for the different data sets, using 100,000 MCMC interactions, and 10 repetitions for the number of groups (K) varying between 1 to 15. The LFMM analysis considered K = 8 (species) and 6 (*A. aculeata* Americas). The associations with environmental variables were identified for the loci with corrected p-values, considering FDR = 0.1 of environmental variable association to detect SNPs, from the LEA package [40], using the sparse non-negative matrix factorization function (snmf) [42]. We carried out the identification of environmental variables by the principal component analysis, adopting 19 bioclimatic variables from the WorldClim database [41], and selecting the variables that showed the highest correlation. For the snmf function, the most likely number of populations for the different data sets was determined using 100,000 interactions, and 10 repetitions for K = 1-15.

The identification of SNPs hypothetically under selection (outliers) was performed for the following groups: 1) In the genus Acrocomia, considering the species as groups, and 2) within *A. aculeata*, considering as groups the samples' countries of origin. We considered as loci putatively under selection those shared between the three identification methods (*fsthe*t, *pcadapt* and LFMM) (S2 Table). Consequently, we adopted the remaining SNPs considered neutral for the analysis of population genomic diversity and structure.

Population structure

We used all samples (S1 Table) to perform the analysis of the genomic structure for de Acrocomia genus and to infer the number of the most likely groups using the software Structure v.2.3.4 [43], considering only neutral SNPs (3227). We also used the same software to access the genomic structure of *A. aculeata* separately, considering 3259 neutral SNPs identified for the species. Each analysis in Structure was performed with a burn-in of 100,000 interactions, followed by 500,000 repetitions of the Markov Chain Monte Carlo (MCMC) in 10 independent simulations, and without prior information to define the clusters. The number of clusters (K) was determined using the average likelihood values of the ΔK method [44] implemented in the program Structure Harvester [45]. The participation coefficient for each access was given by

the alignment of five repetitions of the best K through the CLUMPP method [46] by the software CLUMPAK [47].

To visualize the genetic relationships among Acrocomia species and within the *A. acuelata*, we obtained the Nei genetic distance [48] between the individuals of each data set, and the Neighbor-Joining (NJ) hierarchical classification method with 20000 bootstrap repetitions, using the poppr package[49] on R [35].

In addition, the Principal Component Analysis (PCoA) was also carried out through the ADE 4 package [50]to explore the genetic structure of the different groups using only neutral SNPs, and was visualized graphically by the ggplot2 package [51].

Analysis of genomic diversity

We conducted the population diversity analysis only with the SNP data set identified as neutral for two groups or taxonomic levels: 1) The genus Acrocomia (except the species *A. hassleri* and *A. glaucescens* as they contain only one individual for each species), and 2) *A. aculeata.* Population estimates of allelic richness, percentage of total alleles by locus, observed heterozygosity, expected heterozygosity, and inbreeding coefficient were calculated using the diveRsity [52], poppr [49], and the PopGenKit packages [53] on R platform [35]. To minimize the effect of differences in the number of samples of each population, we calculated the allelic richness (Ar) and the richness of private alleles (ap) for populations of each group or taxonomic level, by the rarefaction method implemented in the software HP-Rare v.1.1 [54].

Results

In population genetics, neutral loci are genomic regions that are influenced by mutational dynamics and demographic effects, and not by selection. However, loci under selection (i.e., outliers) generally behave differently and therefore reveal "extreme" patterns of variation [55, 56]. Since most population genetic inferences are based on neutral loci, the loci under selection can greatly influence the estimates of genetic parameters. In this sense, it is important to identify and remove the outlier loci from the analysis, with the aim to infer more reliable parameters of population genetic diversity and structure.

Based on pcadapt, fsthet, and LEA, we identified 42 outlier loci for all samples or taxonomic groups for the genus Acrocomia, and 10 outlier loci for the taxonomic group formed

by samples of *A. aculeata*. The neutral datasets for the different groups were constructed by removing the outliers. After the removal of outlier loci (S2 Table), genus Acrocomia (all species) and *A. aculeata* contained 3227 nd 3259 neutral loci, respectively.

Genomic structure of Acrocomia spp.

Structure 2.3.4 software [43] was initially used to access the genomic structures of 172 samples of Acrocomia species based on 3227 neutral SNPs. ΔK had a maximum value of K = 7 (S1 Fig), indicating the existence of seven genetic groups (Fig 2). Samples with an attribution probability score > 0.75 and < 0.75 were assigned to the "pure group" and "admixture group", respectively. Based on the classification of Lorenzi [4] and the geographic distribution of the species, we observed a substructure of samples considered to be *A. aculeata*. Two well-defined subgroups (clusters 1 and 3) strongly associated with the geographical origin of the samples were evident. Cluster 1 (Fig 2) was composed of 38 samples of *A. aculeata* from Central and North America (Costa Rica, Trinidad and Tobago, Puerto Rico, and Mexico) and Colombia. Cluster 2 (Fig 2) comprised 39 samples of *A. totai* and five samples considered as *A. aculeata*. Of the latter, four were collected in the state of Parana, southeastern Brazil, (XAM, PR) and one in state of Tocantins, northern Brazil (PAL). The samples from Campo Grande (CGR) showed low mixture levels with clusters 1 and 5 of *A. aculeata.* Cluster 3 (Fig 2) consisted of 39 samples from Brazil. The majority (n = 34) of these samples were from the southeast region of the country, with five from the north region (BEL population). Cluster 4 was exclusively *A. crispa* samples, with a 100% probability of assignment to the cluster.

Based on assignment probabilities ≤ 0.75, some samples were assigned to an admixture group. Twenty samples of *A. aculeata* from the central-west, north, and northeast regions of Brazil, and all samples of *A. intumescens* displayed a similar genomic composition, with a median level of assignment (≥ 0.50). A genetic admixture of *A. aculeata* samples in cluster 5 (Fig 2) with samples mainly from clusters 1 and 3 was evident. *A. intumescens* samples presented a mixture of clusters 5 and 6, with cluster 6 being practically exclusive to the species. Individuals from Cáceres, MT (CAC), and Braúna, São Paulo (SP) (BRA), with a greater assignment to cluster 2, also showed a significant degree of admixture with clusters 3 and 5.

The NJ and PCoA analyses (Figs 3a and 3b) performed with all the samples showed strong agreement with the results of the Bayesian analysis performed using Structure software. However, the NJ tree showed higher resolution in group/cluster recovery than the PCoA. In

both analyses, *A. crispa* was clearly separated from the rest of the Acrocomia species. In addition, there is a clear genomic differentiation between *A. aculeata* and *A. totai.* Similar to the results obtained using the Structure software, the NJ analysis also recovered the substructure within *A. aculeata,* separating the Brazilian samples from those from other countries (Fig 3a). This separation did not result from the PCoA (Fig 3b). In agreement with the results obtained using the Structure software, both PCoA and NJ grouped *A. media* and *A. intumescens* samples into the cluster formed mainly by *A. aculeata,* with *A. hassleri* and *A. glaucescens* grouped into the *A. totai* cluster. The results of NJ and PCoA also agreed concerning the allocation of samples from Xambré, PR (XAM) originally considered as *A. aculeata* in the cluster of *A. totai*. Samples from Braúna, SP (BRA) and Cáceres, MT (CAC), which were identified as an admixture by the Structure software, occupied an intermediate position between the clusters formed mainly by *A. aculeata* and *A. totai* in the PCoA.

Based on the Structure software results (Fig 2) and NJ and PCoA data (Fig 3a and 3b), the samples from Xambrê, PR (XAM) previously considered *A. aculeata* were treated as *A. totai* species for further analysis of differentiation and genomic diversity. The F_{ST} values enabled a moderate genetic differentiation between species, with an average value of 0.469. The F_{ST} values between species (Table 1) ranged from 0.083 (*A. aculeata* vs. *A. totai)* to 0.946 (*A. media* vs. *A. crispa*). In agreement with the genomic structure analysis findings, all comparisons between *A. crispa* and the other species showed higher values of F_{ST}, demonstrating a greater degree of genetic differentiation of *A. crispa* with the other species.

Table 1. Pairwise FST estimates among five species of Acrocomia

	A. aculeata	***A. totai***	***A. intumescens***	***A. média***	***A. crispa***
A. aculeata	0.000				
A. totai	0.083	0.000			
A. intumescens	0.128	0.194	0.000		
A. média	0.133	0.235	0.700	0.000	
A. crispa	0.673	0.687	0.912	0.946	0.000

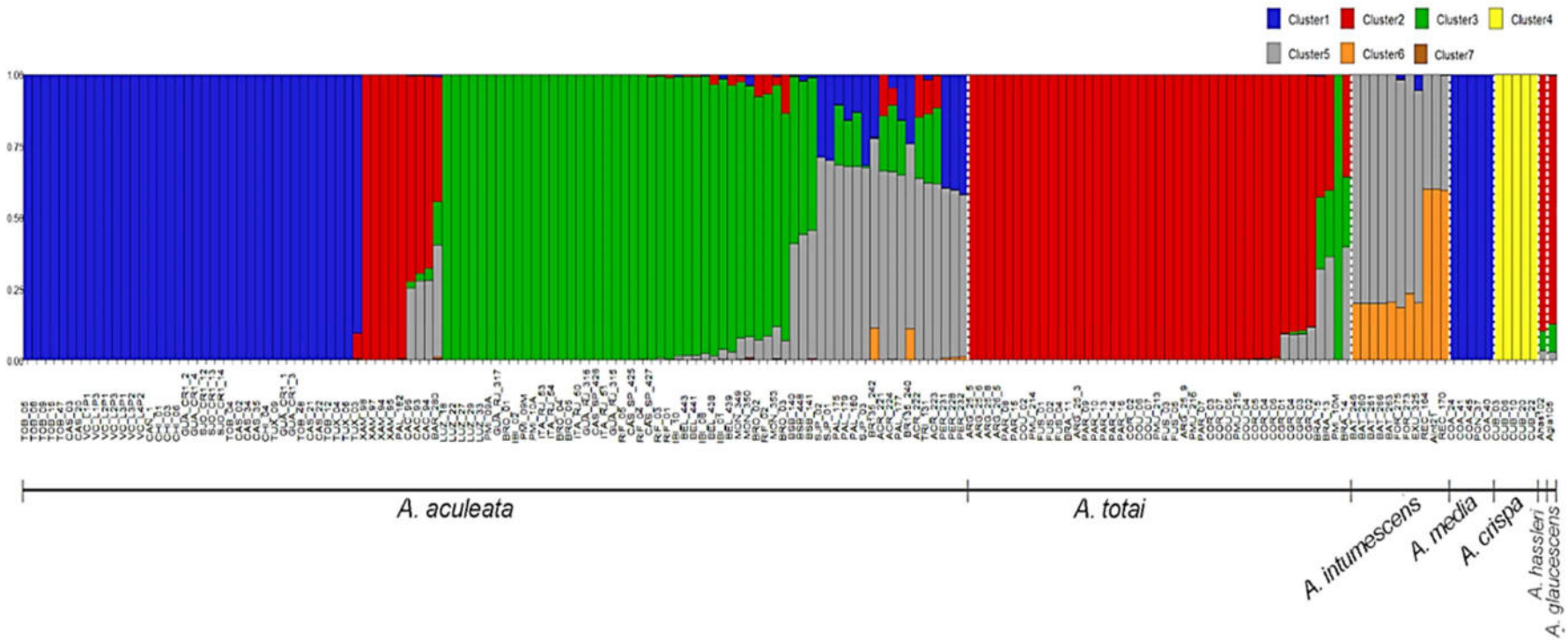


Fig 2. Genomic structure of 172 samples from Acrocomia species based on 3227 neutral SNPs loci. The y-axis is the population membership, and the x-axis is the sample. Each vertical bar represents a sample and color represent separate clusters (k =7)

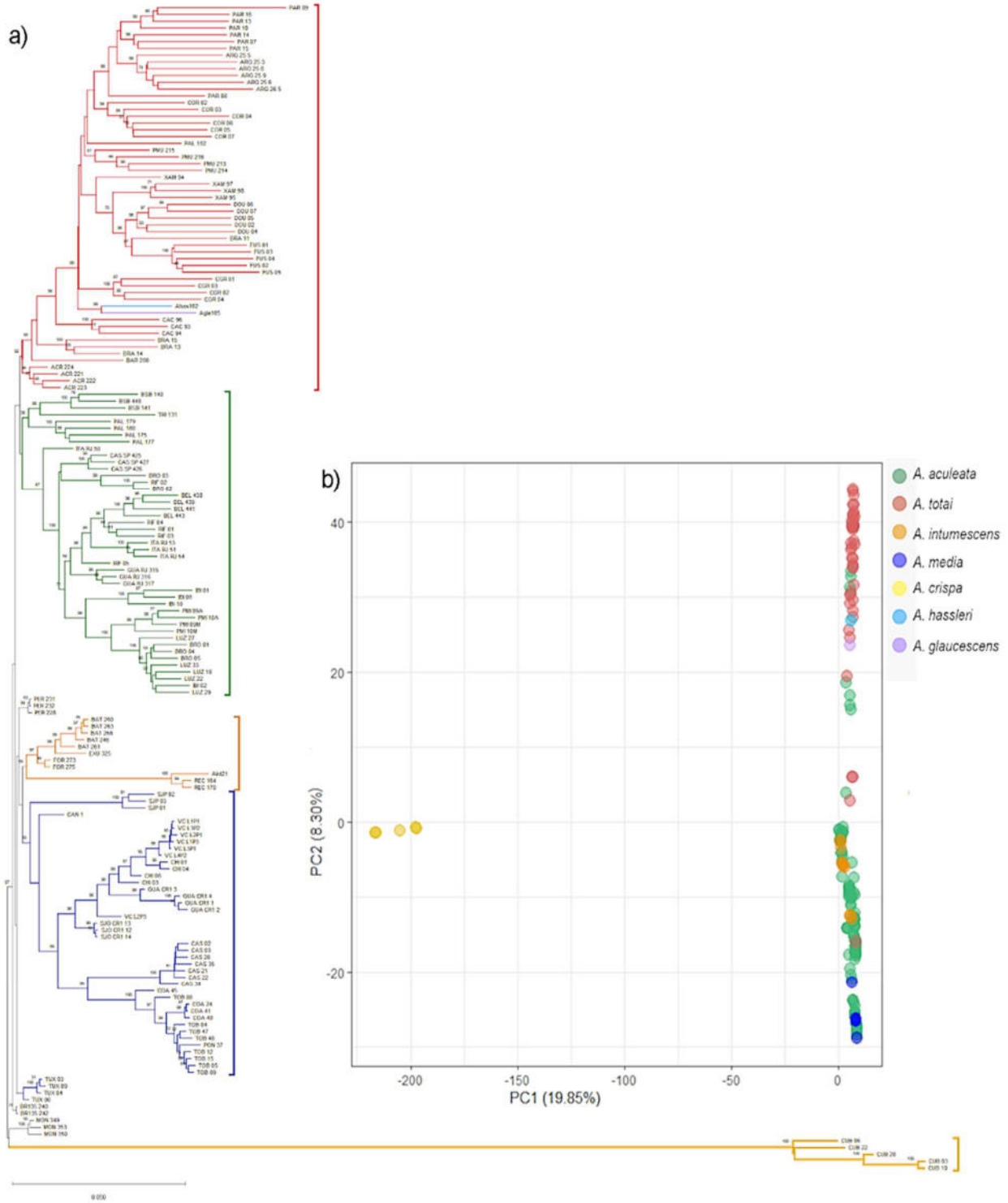


Fig 3. Neighbor-joining (NJ) tree and principal components analysis (PCoA) of Acrocomia species.a) Scatterplot of the principal components analysis (PCoA) showing the dispersion of samples across the first two principal components and b) Neighbor-Joining dendrogram based on Nei´s genetic distance. Bootstrap support of nodes is shown.

Genomic diversity between species

The number of polymorphic loci of the five Acrocomia species ranged from 0.017 to 0.601. *A. aculeata* had the highest mean and *A. media* had the lowest mean (Table 2). The genomic diversity based on the average expected heterozygosity (H_E) in the species ranged from 0.106 in *A. totai* to 0.005 in *A. media*. However, *A. crispa* was the species with the highest allelic richness (2.29) and the highest allelic richness of private alleles (0.17), while *A. media* presented the lowest values of allelic richness and allelic richness of private alleles (1.08 and 0.01, respectively). The inbreeding coefficient (F) values were high for all species, indicating relatively high levels of inbreeding in Acrocomia species, with the exception of *A. media*, which presented negative values (Table 2).

Table 2. Genetic diversity parameter estimates for Acrocomia species calculated from 3227 neutral loci SNPs.

Species	Na	Ne	I	H_O	H_E	Ar	PAr	f
A. aculeata	1.601	1.142	0.157	0.031	0.093	1.20	0.03	0.479
A. totai	1.534	1.160	0.176	0.074	0.106	1.23	0.07	0.262
A. intumescens	1.053	1.027	0.037	0.011	0.025	1.10	0.02	0.483
A. média	0.993	0.984	0.007	0.006	0.005	1.08	0.01	-0.145
A. crispa	0.630	0.619	0.028	0.006	0.020	2.29	0.21	0.591

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

Genomic structure of *A. aculeata*

The population structure of all the *A. aculeata* samples was evaluated using 3259 hypothetically neutral SNPs. Using the method of Evanno [44] the most probable Δk was K = 2 (S2 Fig). This finding supported the presence of two genetically distinct subpopulations previously identified in the structure analysis at the genus level (Fig 2). The two groups were mainly associated with geographical origin, given that samples from Central and North America (Colombia, Costa Rica, Trinidad and Tobago, and Mexico) were grouped in cluster 1, and most of the collected in Brazil were grouped in cluster 2 (Fig 4).

The same two groups identified using the Structure software were also visualized by using the first two PCoA axes as well as the NJ dendrogram. These analyses clearly revealed the formation of

two distinct genetic groups within *A. aculeata*, which are suggested to be geographically separated by the Amazon Rainforest (Fig 4b and 4c). Thus, pronounced differentiation was observed between the individuals of *A. aculeata*. Based on the NJ dendrogram, two large groups were assigned based on geographical origin, separating all individuals from the North and Central America in one node (red) from the Brazilian samples (blue). Two main subgroups were evident in the Northern group. One subgroup contained samples from Peritoró (PER) and São Jose dos Patos (SJP) from Maranhão, Brazil. The other subgroup contained the remaining samples. Interestingly, individuals from Tuxtla Chico, Chiapas (TUX) in Mexico formed a separate cluster from the other samples from Mexico and Colombia, Costa Rica, Trinidad and Tobago, and Puerto Rico (Fig 4b).

The second PCoA axis comprised three samples from Cáceres, MT (CAC). These samples formed a subgroup that was very distant from the other samples of *A. aculeata.* However, the Structure and NJ dendrogram data were not able to discriminate these samples, and grouped with individuals from Brazil (Fig 4a and 4c).

The 'South' group (Cluster 2 in Fig 4a) contained most of the samples from Brazil. The samples collected in Maranda formed a different cluster from the other samples. However, most clusters reflected a strong relationship with the samples geographic origins, with the exception of samples collected in Belém, PA (BEL), northern Brazil, which were more closely related to samples from Rio de Janeiro and São Paulo located in southeastern Brazil. It is also noteworthy that five samples collected in the Brazilian State of Maranhão (PER and SJP) were more closely related with the 'North' group, as evident by the cluster 1 considering the assignment probability of 0.75 in the Structure software analysis (Fig 4a). This result was also corroborated by the NJ and PCoA hierarchical classification (Fig 4b and 4c).

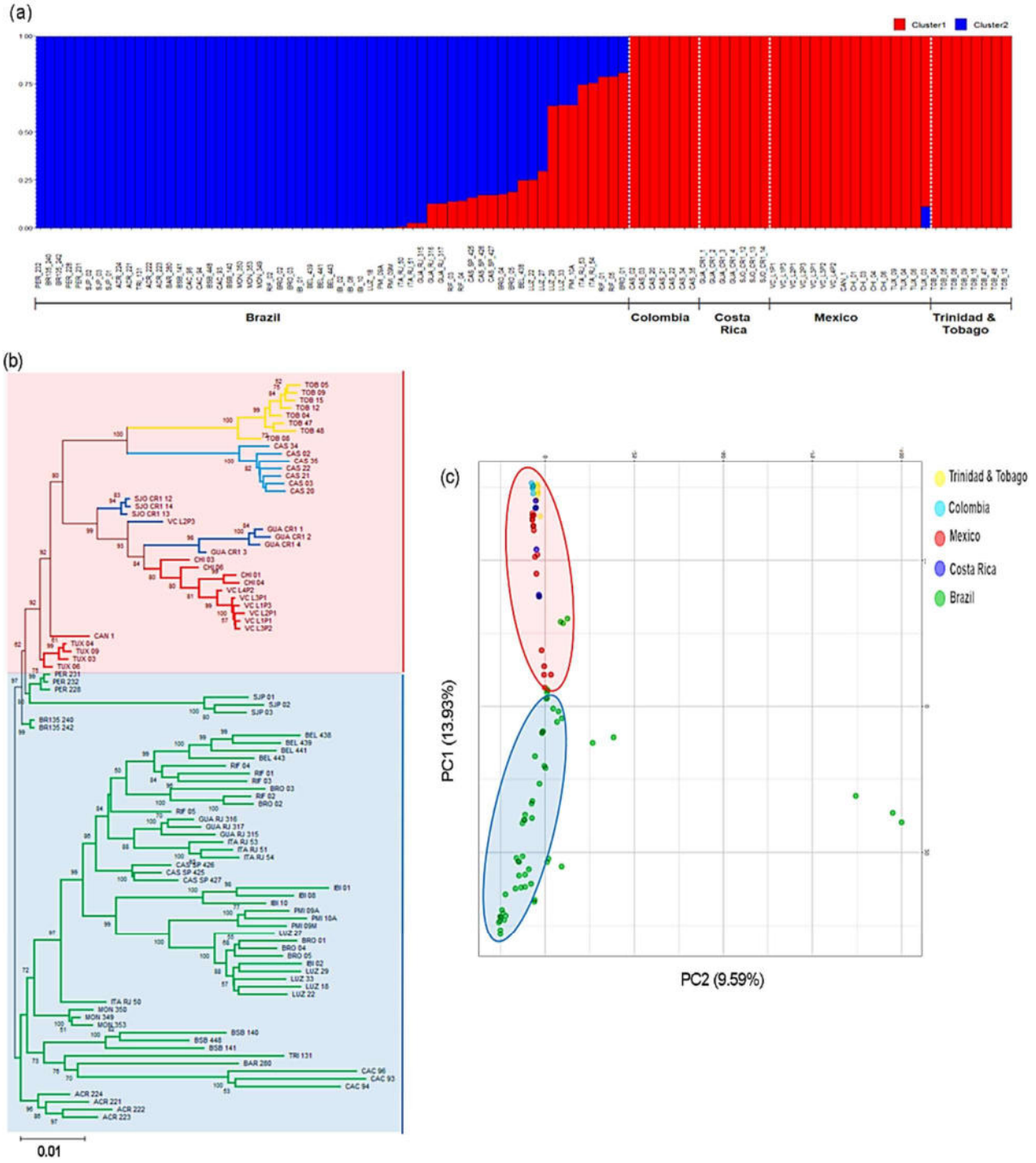


Fig 4. Population genomic structure within *A. aculeata*, based on 3256 neutral loci SNPs. a) Genomic structure from Bayesian analyses (k =2). The y-axis is the population membership, and the x-axis is the sample. Each bar represents an individual and each color is inferred membership in each of the cluster; b) Neighbor-Joining dendrogram based on Nei´s genetic distance. Bootstrap support of nodes is shown. Groups: northern genetic group (Blue); southern genetic group (Red) and c) Scatterplot of the principal components analysis (PCoA) showing the dispersion of samples across the first two principal components.

Genomic diversity of *A. aculeata*

Concerning the genomic diversity within *A. aculeata* species, the greatest diversity was found in Brazil (H_E = 0.081) and the lowest diversity in Mexico (H_E = 0.005). Likewise, the allelic richness values were similar for all populations in the 'North' samples, varying from 1.09 to 1.11. However, the greatest allelic richness for the species was registered in Brazil (Ar = 1.44) (Table 3).

Table 3. Genetic diversity parameter estimates for *A. aculeata* calculated from 3259 neutral loci of SNPs

Country	Na	Ne	I	H_O	H_E	Ar	PAr	f
Trinidad &Tobago	1.001	0.983	0.013	0.008	0.008	1.09	0.01	0.038
Colombia	1.009	0.996	0.018	0.012	0.012	1.09	0.01	-0.014
Mexico	0.994	0.975	0.009	0.004	0.005	1.09	0.01	0.147
Costa Rica	0.981	0.971	0.012	0.009	0.008	1.11	0.01	-0.139
Brazil	1.441	1.124	0.135	0.043	0.081	1.44	0.33	0.377
Brazilian State								
Pará	1.090	1.044	0.059	0.046	0.039	1.11	0.01	-0.168
Mato Grosso	1.090	1.031	0.091	0.062	0.061	1.23	0.04	-0.023
Góias	0.957	0.935	0.041	0.031	0.028	1.26	0.01	-0.068
Distrito Federal	0.870	0.836	0.053	0.033	0.035	1.49	0.01	0.015
Minas Gerais	1.190	1.094	0.089	0.043	0.058	1.09	0.01	0.198
Rio de Janeiro	1.024	0.987	0.037	0.024	0.024	1.14	0	-0.006
São Paulo	1.200	1.089	0.091	0.044	0.059	1.10	0.01	0.179
Maranhão	1.000	0.979	0.042	0.039	0.029	1.17	0.02	-0.350

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

Due to the vast territory and the greater number of *A. aculeata* samples from Brazil, genetic diversity analyses were conducted considering the Brazilian States as populations. Greater diversity (H_E) was found in the states of Mato Grosso (MT), São Paulo (SP), and Minas Gerais (MG), with values of 0.061, 0.059, and 0.058, respectively. In terms of allelic richness (Ar), the most accentuated

values were located in the central-west region of the country, in Distrito Federal (DF), Goiás (GO), and Mato Grosso (MT), with values of 1.49, 1.26, and 1.23, respectively.

Discussion

To our knowledge, this is the first study using GBS for identifying genome-wide SNPs and their application for inferring the genetic diversity and population structure in Acrocomia species and within *A. aculeata*. Sampling was broad in terms of the occurrence of Acrocomia species and comprehensively captured the genomic diversity and structure of the species.

A. aculeata

At the genus level, the distinction of *A. aculeata* as an independent genetic group or taxon was supported through the results obtained with the Bayesian analyses (Fig 2), and by the PCoA and the NJ tree (Figs 3a and 3b). A notable finding was the identification of an accentuated substructure within *A. aculeata,* showing two genetic groups, corresponding to a north– south split in which the samples from Brazil (Northern group, blue cluster in Fig 3) were separated from those of Central and North America (Southern group, red Cluster in Fig 3). This result was evident in the Bayesian analysis performed at the genus level (Fig 2) as well as with only samples of *A. aculeata* (Fig 4a).The substructure identified in *A. aculeata* has not been previously reported and can be attributed to the greater number of samples included in this study, which covered a wide geographic occurrence of the species in the American continent. The presence of two genetic groups may be the result of reproductive isolation due to the Amazon Rainforest acting as a geographical barrier that prevented gene flow between them and with an independent evolution. Another hypothesis is that these two gene pools support the existence of more than one species, as reported in a previous taxonomic classification in Central and North America Countries [57].

Another interesting result observed was that individuals from the population of Maranhão presented as an admixture between the Northern and Southern groups of *A. aculeata* (Fig 4). The origin of the genus Acrocomia is uncertain. However, in the case of *A. aculeata*, based on the dates of archeological records of human use, the most accepted hypothesis suggests that the species originated in northern Brazil (in the region of Santarém, State of Pará) approximately 11,200 MY, and was later dispersed by humans to Central America [58]. According to our results, the admixture observed in the populations of Maranhão (neighboring to Pará State) (Fig 4a) may support this hypothesis, suggesting a common geographical origin of the to two genetic groups in the northeast

region of Brazil. In agreement with the *A. aculeata* dispersion routes from South to Central and North America [58], the low values of genetic diversity for the species found in the Northern group may have resulted from a founder effect, since all population of this cluster presented lower values of genetic diversity than those observed in the populations of the southern cluster (Brazil) (Table 3).

Bayesian analysis identified individuals of *A. aculeata* with a degree of genetic admixture with *A. totai* (Cluster 2, in Fig 2) and *A. intumescens* (Cluster 6, in Fig 2), suggesting gene flow between species. As *A. aculeata* is dispersed mainly by cattle [59, 60], the agricultural expansion and livestock may have favored the dispersion of the species to areas where *A. totai* and *A. intumescens* occur, creating opportunities for hybridization due to secondary contact. There have been no reports of interspecific hybridization in the Acrocomia genus. However, a recent study using microsatellite markers also detected connectivity between populations of *A. aculeata* and *A. totai* in Brazil [61].

A. aculeata displays the greatest geographical distribution of the genus [3, 4, 14, 15]. As expected for a species with a wide distribution that has adapted to diverse environmental conditions, the genetic diversity of *A. aculeata* was high when compared to other species (Table 3). At the intraspecific level, the highest genetic diversity for the species was found in Brazil, especially in the States of Minas Gerais and São Paulo (Table 3). Although it is not possible to make direct comparisons due to the different types of molecular markers used, previous studies also identified a high genetic diversity for *A. aculeata* in the States of Minas Gerais and São Paulo [23, 28, 29].

An unexpected result was the low genetic diversity of *A. aculeata* in Mexico, where the species is also distributed in an extensive geographical area, from the north to the south of the country (Table 4). These results could reflect the use and exploration of the species in that country and other Central American countries, where adult plants are harvested as the raw material for a fermented drink called “taverna” [62, 63]. This kind of exploration is one of the main factors driving the reduction size or elimination of the natural populations, which affects the reproductive capacity of the species and its natural regeneration [63] and might also been reducing the genetic diversity.

A. aculeata is strongly associated with humans [58, 59]. Even though it is considered an incipiently domesticated species, it has a wide range of uses in different countries of the Americas [7, 8, 64]. Therefore, patterns of genetic diversity and structure can also be the result of different states of domestication, with different intensities of selection in each region, as also reported for other species, such as beans [65], tomato [66], and cacao [67].

A. totai

A. totai was the second most geographically dispersed species in the genus. It has been documented in eastern Bolivia, Paraguay, Central-west Brazil to northern Argentina [4, 15]. The taxonomic distinction of the species has been demonstrated based on morphological data and geographic distribution [4], leaf anatomy [1], and fruit biometry [68]. However, *A. totai* is commonly regarded as *A. aculeata* due to the pronounced morphological similarity of both species, and because both have fruits with similar biometric and color characteristics [68]. Our results were congruent with the current taxonomic classification of the species. Almost all samples initially considered as *A. totai* (94%) belonged to cluster 2 with a high assignment probability (> 0.75), according to Structure analysis (Fig 2), and corroborated with PCoA and NJ analysis (Figs 3a and 3b). Our results agreed with those of Lima et al. [61], that documented the clear genetic differentiation between *A. aculeata* and *A. totai* (treated as ecotypes) using microsatellite markers. Although not treated as distinct species, but considering the geographical distribution of both, several studies using molecular and morphological markers also reinforced the classification of *A. totai* as a distinct taxon. Lanes et al. (23) used microsatellite markers to demonstrate the marked genetic differences of A. aculeata between individuals from the Pantanal region, State of Mato Grosso do Sul, Brazil, and other regions of the country. Similarly, Silva et al. (27) analyzed the variation in the internal transcribed spacer (ITS) region and identified four haplotypes. Two were shared by genotypes from São Paulo and Minas Gerais, and one was exclusive to genotypes collected in Mato Grosso do Sul. The morphological characteristics of A. aculeata include larger fruits (3.5 and 5.0 cm) and a pulp oil content that can reach approximately 78% (27, 68-70) while the fruits of A. totai are smaller (2.5 and 3.5 cm) with a pulp oil content between 26% and 33% (68, 71, 72.

In Brazil, *A. totai* is considered to be restricted to the State of Mato Grosso do Sul [4, 69, 70]. An interesting finding of our study was that samples from Xambrê, Paraná (XAM) and a sample from Palmas, Tocantins (PAL_182), considered as *A. aculeata* based on Lorenzi et al., [4] taxonomic classification, were attributed to cluster 2 of *A. totai* by the Bayesian analysis (Fig 2), by PCoA, and by NJ (Figs 3a and 3b). Although the occurrence of *A. totai* in these states has not been proven, our results are consistent with the information reported on the Flora do Brazil 2020 website [15], indicating the possible occurrence of *A. totai* in these states.

Although the genetic structure and separation of *A. aculeata* from *A. totai* was evident based on the cluster analyses, the genetic differentiation (F_{ST}) between species was 0.083, which was the

lowest value (Table 1). This result was consistent with the value obtained using microsatellite markers (F_{CT} = 0.07) by Lima et al., [61]. The findings may reflect the retention of ancestral polymorphisms, the hybridization or gene flow between species in convergent areas [61] or could be evidence of an ongoing speciation process [23].

Based on the H_E and Ar values, *A. totai* was the species with the highest level of genetic diversity (Table 2). Our results are comparable to those found in a recent study using microsatellite markers [61], in which the genetic diversity of *A. totai* was greater than that of *A. aculeata*. Similar, previous studies also identified greater genetic diversity in populations from Mato Grosso do Sul than population from other location of Brazil, although the authors did not consider the populations to be *A. totai* [23, 25]. The high diversity observed in *A. totai* could reflect its geographically widespread occurrence and expansion of genetic diversity promoted by the interspecific hybridization with *A. aculeata*.

The results of cluster analysis and genetic differentiation corroborated the classification of *A. totai* as an independent taxon based on morphological [4], anatomical [1], and molecular markers [61]. This taxonomic separation seems to be more appropriate than that proposed for Henderson et al. [3], which considered all tree-sized Acrocomias as a single taxonomic group called *A. aculeata*.

A. intumescens

Contrary to the actual taxonomic classification [4-6], our analyses did not show a clear genetic separation of *A. intumescens* (Figs 2, 3a, and 3b). All the samples of *A. intumescens* were assigned to cluster 6, however presented high levels of admixture with *A. aculeata* (cluster 5, Fig 2). *A. intumescens* also showed a moderate genetic differentiation with *A. aculeata* (F_{ST} = 0.128, Table 1), reinforcing the close genetic relationship among both species as described by Vianna et al. [1] based on leaf anatomy. Morphologically, *A. intumescens* is distinguished mainly by the swelling of the stipe [4]. However, botanical characters suggested to delimit Acrocomia species have revealed an overlapping in size of fruits [68] and for oil content in the mesocarp, ranging from 37 to 78% in *A. aculeata* [71, 72] and from 34 to 41% in *A. intumescens* [71, 73].

A phylogenetic study by Meerow et al. [74], estimate the divergence of *A. intumescens* and *A. aculeata* 5 MA ago. The genetic structure we observed may reflect the maintenance of ancestral polymorphism, possibly as a result of the recent divergence of these species with insufficient time for the appearance of reproductive isolation mechanisms, allowing the interspecific hybridization. *A.*

intumescens is endemic to northeast Brazil and has a restricted distribution [4, 15]. Species with a restricted geographical distribution tend to have reduced genetic diversity than species with a wide geographical distribution [75, 76]. Consistent with this trend, *A. intumescens* showed lower values of heterozygosity and allelic richness than the wide geographical distribution species *(A. aculeata* and *A. totai*) (Table 2). However, the genetic diversity found in *A. intumesces* was comparable to that observed in other plant species associated with restricted geographic distribution [77-79].

A. crispa

A. crispa is an insular species with a distribution restricted to Cuba. A clear separation and a strong genetic divergence compared to the other species, as evidenced in the cluster analysis (Fig 2, 3a, and 3b) and by the high values of F_{ST} (Table 1). These expectations were understood if considered that the gene flow through pollen or seed dispersal between island populations and continental populations is limited such that a strong genetic structure and a high degree of differentiation between them is expected, as reported for several species [80, 81]. Our results are congruent with those reported for other tree species, which also showed high levels of genetic differentiation between island populations compared to continental populations and lower levels of genetic diversity on the islands than on the continent [82-85]. *A. crispa* displayed low values of genetic diversity ($H_E = 0.020$) compared with other Acrocomia species, although these values are expected for endemic island species. However, interestingly, *A. crispa* presented the greatest allele richness (2.29) and allele richness of private alleles (0.17) (Table 2). Based on chlorosplastic and nuclear genes, the time of divergence estimated for *A. crispa* as 16 Mya, while *A. aculeata* and *A. intumescens* diverged 5 Mya [74]. This more ancient divergence associated with geographic isolation may support the allelic richness and the greater number of private alleles found in *A. crispa*, as well as the strong genetic differentiation of from other Acrocomia species. This hypothesis has also been posited for other endemic species of islands that have congeners on the continent [86, 87].

There is no detailed information about the morphological characteristics of *A. crispa*. However, some morphological differences have been described, such as the presence of swelling in the median region of the stipe as the most discriminating botanical characteristic Bailey [57], the smaller fruits, varies from 1 to 3 cm [3], than that described in *A. totai* (2.5 to 3.5 cm), *A. intumescens* (3.0 to 4.0 cm) and *A. aculeata* (3.5 to 5.0 cm) [68] and also differences in pollen morphology with

trichotomocolpated pollen in *A. aculeata* and monocolpous pollen in *A. crispa* (named *Gastrococos crispa* by the authors) [88].

A. crispa, previously designated to the genus Gastrococos by Moore [89], was recently allocated to the genus Acrocomia, mainly due to the sequencing of the nuclear *prk* gene [90]. Although most phylogenetic studies that analyzed support the relationship between *A. aculeata* and *A. crispa* as sisters in a single monophyletic group [90-95], other phylogenetic [74] and cladistic studies [96] shown that they are sister species in paraphyletic groups. However, these phylogenetic studies were conducted at higher taxonomic levels (families, subfamilies, and tribes), with the inclusion of few species of Acrocomia. Therefore, they have limited ability to accurately reveal phylogenetic relationships of Acrocomia species.

The morphological characteristics of the species, the divergence time and our results of genetic differentiation, diversity, and structure may collectively support an independent taxonomic status of *A. crispa* within the genus Acrocomia. Therefore, we suggest a revision of the taxonomy for the species.

A. media

In contrast to the evidence of genetic divergence for *A. crispa*, the recognition of *A. media* as an independent taxonomic unit was not supported by our study. As *A. media* is also an island species, it would be expected to have a strong genetic structure when compared to other Acrocomia species with a continental distribution. Contrary to this assumption, all samples considered as *A. media* were assigned to the northern group of *A. aculeata*, as evidenced by three cluster analyses (Figs 2, 3a, and 3b). In addition, the F_{ST} values (Table 1) also indicated low genetic differentiation of *A. media* compared to *A. aculeata.*

The patterns of genetic diversity observed in *A. media* were the lowest compared to other species (H_E= 0.005 and Ar = 1.08), but were consistent with several studies of population genetics in plants, which predicted that island populations have reduced levels of genetic diversity compared to continental populations [80, 97]. The low genetic diversity observed in *A. media* can be attributed to the founder effect associated with the establishment of populations with only a few individuals [97, 98] or to genetic drift due to stochastic events inherent in the islands and/or fragmentation during its formation [99]

A. media was first described in Puerto Rico by Cook [100]. The author adopted the shortest trunk and the smallest diameter of the stipe as the differentiating characteristics of *A. media* from *A. aculeata*. However, *A. media* was considered synonymous with *A. aculeata* for a long time due to the absence of consistent botanical characteristics for differentiation. In 2013, The Plant List recognized *A. media* as a distinct species based on the floristic palm inventory of Proctor [101]. However, the same author mentioned that the existing information about *A. media* was very old and based on few individuals, suggesting an increase in the number of evaluated individuals to guarantee a more consistent morphological description of the species. The only phylogenetic study performed with *A. media* included an individual from Puerto Rico, and a sample of *A. aculeata* from Brazil revealed that both species were closely related [90].

Based on the lack of genetic differentiation of *A. media*, low genetic diversity in the species, and low pairwise F_{ST} value between *A. media* and *A. aculeata*, we hypothesize that *A. media* is synonymous with *A. aculeata*. Thus, a recent introduction in Puerto Rico was not sufficient to characterize the reproductive isolation needed for the differentiation of *A. aculeata.*

A. hassleri and *A. glauscescens*

The genomic data of our study did not allow the assignment of distinct taxonomic units to the species *A. hassleri* and *A. glauscescens*. Based on morphological characters, the species are clearly differentiated from the others by their small size. However, based on the results obtained from the cluster analysis, they were assigned to cluster 2, being closely related to *A. totai* (Fig. 2, 3a, and 3b). However, this result should be considered with caution, as we only used one sample of each species in the analyses, which could limit the comparison of genetic estimates and decrease the probability of detecting genetic structure, as evidenced in similar studies with a low number of samples [102, 103]. Further studies with a greater number of accessions are needed to increase the species representation, and to establish reliable genetic relationships between *A. hassleri* and *A. glauscescens* and other Acrocomia species.

Conclusions

Our study is the first to offer evidence of the efficiency of NGS through the application of the GBS protocol in Acrocomia. The data may constitute a reference for the application of this protocol in the genus. Even without a reference genome, we successfully identified a large number of SNPs for several species, revealing potentially valuable markers for future studies in the genus Acrocomia.

The SNPs yielded unprecedented results of the genetic relationships between Acrocomia species as well as at the population level for *A. aculeata.* In general, our results were partially congruent with the taxonomy of the genus, supporting the current separation of some species. The genomic structure revealed the formation of well-defined genetic groups and confirmed the distinction of *A. aculeata, A. totai*, *A. intumescens*, and *A. crispa*, with the latter showing a strong genetic differentiation as well as the absence of genetic distinction of *A. media*. We recommend a review of the current taxonomic classification of *A. crispa* and *A. media*. In addition, SNPs also allowed the identification of gene flow patterns and/or hybridization between species.

In the case of *A. aculeata*, the data provide an overview of the genomic diversity and structure from sampling over a wide area of occurrence. The genomic data showed the existence of two large gene pools in the species at the continental level (north and south), with greater genomic diversity in the latter populations. The results from this study will serve as a reference for current and future studies on genetic diversity, taxonomy, evolution, ecology, and phylogeny of the genus Acrocomia, and will support genetic breeding, conservation, and management activities for *A. aculeata*.

References

1. Vianna SA, Carmelo-Guerreiro SM, Noblick LR, Colombo CA. Leaf anatomy of Acrocomia (Arecaceae): an additional contribution to the taxonomic resolution of a genus with great economic potential. Plant Syst Evol. 2017;303:233-248.
2. Martius CFPV. Historia Naturalis Palmarum. Leipzig, TC Weigel. 1824:285-286.
3. Henderson A, Galeano G, Bernal R. Field guide to the palms of the Americas. New Jersey, USA: Princeton University Press; 1995.
4. Lorenzi H, Noblick L, Kahn F. Flora Brasileira-Arecaceae (Palmeiras) Nova Odessa, SP: Instituto Plantarum. 2010.
5. list TP. Version 1.1 2013 [Available from: http://www.theplantlist.org/
6. Palmweb. Acrocomia Mart. Hist. Nat. Palm. n. d [Available from: http://www.palmweb.org/.
7. Lorenzi GMAC. *Acrocomia aculeata* (Jacq.) Lodd. ex Mart. Arecaceae: bases para o extrativismo sustentável. Curitiba, PR: Universidade Federal do Paraná; 2006.

8. Ramos MIL, Ramos Filho MM, Hiane PA, Braga Neto JA, Siqueira EMA. Qualidade nutricional da polpa de bocaiúva *Acrocomia aculeata* (Jacq.) Lodd. Food Sci Technol 2008;28:90-94.

9. Colombo CA, Berton LHC, Diaz BG, Ferrari RA. Macauba: a promising tropical palm for the production of vegetable oil. OCL. 2018;25(1):D108.

10. Coimbra MC, Jorge N. Characterization of the pulp and kernel oils from Syagrus oleracea, Syagrus romanzoffiana, and *Acrocomia aculeata*. J Food Sci. 2011;76(8):C1156-C1161.

11. da Silva César A, de Azedias Almeida F, de Souza RP, Silva GC, Atabani A. The prospects of using *Acrocomia aculeata* (macaúba) a non-edible biodiesel feedstock in Brazil. Renew Sust Energ Rev. 2015;49:1213-1220.

12. Abreu AG, Priolli RHG, Azevedo-Filho JA, Nucci SM, Zucchi MI, Coelho RM, et al. The genetic structure and mating system of *Acrocomia aculeata* (Arecaceae). Genet Mol Biol. 2012;35(1):116-121.

13. Abreu IS, Carvalho CR, Carvalho GMA, Motoike SY. First karyotype, DNA C-value and AT/GC base composition of macaw palm (*Acrocomia aculeata*, Arecaceae)–a promising plant for biodiesel production. Aust J Bot. 2011;59(2):149-155.

14. Dransfield J, Uhl NW, Lange CBA, Baker WJ, Harley MM, Lewis CE. Genera Palmarum: the evolution and classification of palms: Kew Publishing; 2008.

15. Vianna SA, Campos_Rocha A. *Acrocomia in* Flora do Brasil 2020 em construção: Jardim Botânico do Rio de Janeiro; 2020 [Available from: <http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB15662.

16. Vianna S, Colombo C, editors. Distribuição geográfica de acrocomia aculeata (jacq,) lodd ex mart.(arecaceae) em sua região de ocorrência. I Congresso brasileiro de macaúba; 2013.

17. Costa CJ, Marchi ECS. Germinação de sementes de palmeiras com potencial para produção de agroenergia. Embrapa Cerrados-Documentos (INFOTECA-E). 2008.

18. Vinson C, Mangaravite E, Sebbenn A, Lander T. Using molecular markers to investigate genetic diversity, mating system and gene flow of Neotropical trees. Braz J Bot. 2018;41(2):481-496.

19. Poland JA, Rife TW. Genotyping-by-sequencing for plant breeding and genetics. The Plant Genome. 2012;5(3):92-102.

20. Boutet G, Carvalho SA, Falque M, Peterlongo P, Lhuillier E, Bouchez O, et al. SNP discovery and genetic mapping using genotyping by sequencing of whole genome genomic DNA from a pea RIL population. BMC genom. 2016;17(1):1-14.

21. Davey JW, Blaxter ML. RADSeq: next-generation population genetics. Brief Funct Genomics. 2010;9(5-6):416-423.

22. Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS genet. 2010;6(2):e1000862.

23. Lanes ÉC, Motoike SY, Kuki KN, Nick C, Freitas RD. Molecular characterization and population structure of the macaw palm, *Acrocomia aculeata* (Arecaceae), ex situ germplasm collection using microsatellites markers. J Hered. 2015;106(1):102-112.

24. Lanes ÉC, Motoike SY, Kuki KN, Resende MD, Caixeta ET. Mating system and genetic composition of the macaw palm (*Acrocomia aculeata*): implications for breeding and genetic conservation programs. J Hered. 2016;107(6):527-536.

25. Mengistu FG, Motoike SY, Cruz CD. Molecular characterization and genetic diversity of the macaw palm ex situ germplasm collection revealed by microsatellite markers. Diversity. 2016;8(4):20.

26. Nucci SM. Desenvolvimento, caracterização e análise da utilidade de marcadores microssatélites em genética de população de macaúba. : Instituto Agronômico de Campinas; 2007.

27. Silva LCC, Lemos RdC, Carvalho CGI, Good-God PIV, Oliveira LOd, Costa MD-BL, et al. Genetic diversity and structure of macaw palm: implications for genetic variability sampling. Rev Arvore. 2017;41(5).

28. Oliveira D, Melo Júnior A, Brandão M, Rodrigues L, Menezes E, Ferreira P. Genetic diversity in populations of Acrocomia aculeata (Arecaceae) in the northern region of Minas Gerais, Brazil. Genet Mol Res. 2012;11(1):531-538.

29. Coelho NHP, Tambarussi EV, Aguiar BI, Roque RH, Portela RM, Braga RC, et al. 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. Conserv Genet. 2018;19(4):879-891.

30. Velazquez MAM. Caracterización fitoquímica, molecular y evaluación de respuestas del cultivo in vitro de coyol (Acrocomia mexicana Kraw ex Marti)/por Antonio Magdiel Velazquez Mendez: Universidad Autónoma Chapingo; 2013.

31. Doyle JJ, Doyle JL. Isolation ofplant DNA from fresh tissue. Focus. 1990;12(13):39-40.

32. Poland JA, Brown PJ, Sorrells ME, Jannink J-L. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS one. 2012;7(2):e32253.

33. Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes genom genet. 2011;1(3):171-182.

34. Luu K, Bazin E, Blum MG. pcadapt: an R package to perform genome scans for selection based on principal component analysis. Mol Ecol Resour. 2017;17(1):67-77.

35. Core TR. R: A language and environment for statistical computing. Vienna, Austria; 2018.

36. Flanagan SP, Jones AG. Constraints on the F ST–heterozygosity outlier approach. J Hered. 2017;108(5):561-573.

37. Wright S. Isolation by distance. Genetics. 1943;28(2):114.

38. Wright S. Evolution in Mendelian populations. Genetics. 1931;16(2):97.

39. Frichot E, Schoville SD, Bouchard G, François O. Testing for associations between loci and environmental gradients using latent factor mixed models. Mol Biol Evol 2013;30(7):1687-1699.

40. Frichot E, François O. LEA: An R package for landscape and ecological association studies. Methods Ecol Evol. 2015;6(8):925-929.

41. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 2005;25:965–1978.

42. Martins H, Caye K, Luu K, Blum MG, François O. Identifying outlier loci in admixed and in continuous populations using ancestral population differentiation statistics. Molecular ecology. 2016;25(20):5029-5042.

43. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945-959.

44. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 2005;14(8):2611-2620.

45. Earl DA, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2012;4(2):359-361.

46. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007;23(14):1801-1806.

47. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Molecular ecology resources. 2015;15(5):1179-1191.

48. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 1978;89(3):583-590.

49. Kamvar ZN, Brooks JC, Grünwald NJ. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Frontiers in genetics. 2015;6:208.

50. Dray S, Dufour AB, Chessel D. The ade4 package-II: Two-table and K-table methods. R news. 2007;7(2):47-52.

51. Villanueva RAM, Chen ZJ. ggplot2: Elegant graphics for data analysis. Taylor & Francis; 2019.

52. Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods in ecology and evolution. 2013;4(8):782-788.

53. Paquette SR, Paquette MSR. Package 'PopGenKit'. 2011.

54. Kalinowski ST. hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Molecular ecology notes. 2005;5(1):187-189.

55. Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet. 2003;4(12):981-994.

56. Allendorf FW, Hohenlohe PA, Luikart G. Genomics and the future of conservation genetics. Nature Rev Genet. 2010;11(10):697-709.

57. Bailey LH. *Acrocomia* preliminary paper. Gentes Herbarum – Occasional Papers on the Kinds Plants; 1941.

58. Morcote-Rios G, Bernal R. Remains of palms (Palmae) at archaeological sites in the New World: a review. Bot Rev. 2001;67(3):309-350.

59. Lentz DL. *Acrocomia mexicana*: Palm of the ancient Mesoamericans. J Ethnobiol. 1990;10(2):183-194.

60. Scariot A. Seed dispersal and predation of the palm *Acrocomia aculeata*. Principes. 1998;42:5-8.

61. de Lima NE, Meerow AW, Manfrin MH. Genetic structure of two Acrocomia ecotypes (Arecaceae) across Brazilian savannas and seasonally dry forests. Tree Genet Genomes. 2020;16(4):1-12.

62. Balick MJ. Production of coyol wine from Acrocomia mexicana (Arecaceae) in Honduras. Econ Bot. 1990;44(1):84-93.

63. H. D-FV, Amín R-CP, Beau M-VB, Moises A-B. Caracterización morfológica del coyol Acrocomia aculeata (Jacq.), y su potencial productivo de aceite vegetal. In: Instituto Nacional de Investigaciones Forestales AyPCdIRPSCERI, editor. Folleto Técnico Tuxtla Chico, Chiapas, México2020. p. 54.

64. Ramírez Hernández BC, Zañudo Hernández J, García de Alba Verduzco JE, Délano Frier JP, Pimienta Barrios E, García Martínez MÁ. Importancia agroecológica del coyul (*Acrocomia mexicana* Karw. ex Mart.). Estudios sociales (Hermosillo, Son). 2013;21(41):95-113.

65. Rossi M, Bitocchi E, Bellucci E, Nanni L, Rau D, Attene G, et al. Linkage disequilibrium and population structure in wild and domesticated populations of Phaseolus vulgaris L. Evol Appl. 2009;2(4):504-522.

66. Blanca J, Cañizares J, Cordero L, Pascual L, Diez MJ, Nuez F. Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato. PloS one. 2012;7(10):e48198.

67. Cornejo OE, Yee M-C, Dominguez V, Andrews M, Sockell A, Strandberg E, et al. Population genomic analyses of the chocolate tree, Theobroma cacao L., provide insights into its domestication process. Communications biology. 2018;1(1):1-12.

68. Vianna SA, Berton LHC, Pott A, Guerreiro SMC, Colombo CA. Biometric characterization of fruits and morphoanatomy of the mesocarp of Acrocomia species (Arecaceae). Int J Biol. 2017;9(3):78-92.

69. Markley KS. Mbocayá or Paraguay Cocopalm—an important source of oil. Econ Bot. 1956;10(1):3-32.

70. Rodríguez MF, Aschero CA. Acrocomia chunta (Arecaceae) raw material for cord making in the Argentinean Puna. J Archaeol Sci. 2005;32(10):1534-1542.

71. Conceição L, JUNQUEIRA NTV, Licurgo FdS, ANTONIASSI R, WILHELM AE, BRAGA MF, editors. Teor de óleo em frutos de diferentes espécies de macaubeira (Acrocomia spp.). Embrapa Cerrados-Artigo em anais de congresso (ALICE); 2012: In: CONGRESSO BRASILEIRO DE FRUTICULTURA, 22., 2012, Bento Gonçalves. Anais ….

72. Berton L. Avaliação de populações naturais, estimativas de parâmetros genéticos e seleção de genótipos elite de macaúba (Acrocomia aculeata). Campinas, SP. Brazil: Instituto Agronômico de Campinas - IAC; 2013.

73. Bora PS, Rocha R. Macaiba palm: fatty and amino acids composition of fruits macaíba: composición de aminoácidos y ácidos grasos de frutos macaíba: composición de aminoácidos e ácidos graxos de froitos. CYTA-J Food. 2004;4(3):158-162.

74. Meerow AW, Noblick L, Salas-Leiva DE, Sanchez V, Francisco-Ortega J, Jestrow B, et al. Phylogeny and historical biogeography of the cocosoid palms (A recaceae, A recoideae, C ocoseae) inferred from sequences of six WRKY gene family loci. Cladistics. 2015;31(5):509-534.

75. Cole CT. Genetic variation in rare and common plants. Annu Rev Ecol Evol Syst. 2003;34(1):213-237.

76. Hamrick J, Godt M. Allozyme diversity in plant species. In 'Plant population genetics, breeding and genetic resources'.(Eds AHD Brown, MT Clegg, AL Kahler, BS Weir) pp. 43–63. Sinauer Associates: Sunderland, MA; 1989.

77. Ellis J, Pashley C, BURKE* J, McCauley D. High genetic diversity in a rare and endangered sunflower as compared to a common congener. Mol Ecol 2006;15(9):2345-2355.

78. Furches MS, Small RL, Furches A. Genetic diversity in three endangered pitcher plant species (Sarracenia; Sarraceniaceae) is lower than widespread congeners. Am J Bot. 2013;100(10):2092-2101.

79. Levy E, Byrne M, Coates D, Macdonald B, McArthur S, Van Leeuwen S. Contrasting influences of geographic range and distribution of populations on patterns of genetic diversity in two sympatric Pilbara Acacias. PLoS One. 2016;11(10):e0163995.

80. Franks SJ. Genetics, evolution, and conservation of island plants. J Plant Biol. 2010;53(1):1-9.

81. Holsinger KE, Weir BS. Genetics in geographically structured populations: defining, estimating and interpreting F ST. Nat Rev Genet. 2009;10(9):639-650.

82. Chen D, Zhang X, Kang H, Sun X, Yin S, Du H, et al. Phylogeography of Quercus variabilis based on chloroplast DNA sequence in East Asia: multiple glacial refugia and mainland-migrated island populations. PLoS One. 2012;7(10):e47268.

83. Qi X-S, Yuan N, Comes HP, Sakaguchi S, Qiu Y-X. A strong 'filter'effect of the East China Sea land bridge for East Asia's temperate plant species: inferences from molecular phylogeography and ecological niche modelling of Platycrater arguta (Hydrangeaceae). BMC Evol Biol 2014;14(1):41.

84. Qiu YX, Sun Y, Zhang XP, Lee J, Fu CX, Comes HP. Molecular phylogeography of East Asian Kirengeshoma (Hydrangeaceae) in relation to Quaternary climate change and landbridge configurations. New Phytol. 2009;183(2):480-495.

85. Sakaguchi S, QIU YX, LIU YH, QI XS, KIM SH, Han J, et al. Climate oscillation during the Quaternary associated with landscape heterogeneity promoted allopatric lineage divergence of a temperate tree Kalopanax septemlobus (Araliaceae) in East Asia. Mol Ecol. 2012;21(15):3823-3838.

86. Mitsui Y, Chen S-T, Zhou Z-K, Peng C-I, Deng Y-F, Setoguchi H. Phylogeny and biogeography of the genus Ainsliaea (Asteraceae) in the Sino-Japanese region based on nuclear rDNA and plastid DNA sequence data. Ann Bot. 2008;101(1):111-124.

87. Nakamura K, Denda T, Kokubugata G, Huang C-J, Peng C-I, Yokota M. Phylogeny and biogeography of the Viola iwagawae-tashiroi species complex (Violaceae, section Plagiostigma) endemic to the Ryukyu Archipelago, Japan. Plant Syst Evol. 2015;301(1):337-351.

88. Machado SR. Variaciones en la morfología polínica de Arecaceae en Cuba: abertura tricotomosulcada y estratificación de la exina. Revista Jard Bot Nac Univ Habana. 2003:71-79.

89. Moore HE. *Gastrococos crispa* (Kunth). Principes. 11.1968.

90. Gunn BF. The phylogeny of the Cocoeae (Arecaceae) with emphasis on Cocos nucifera. Ann Missouri Bot Gard. 2004:505-522.

91. Hahn WJ. A phylogenetic analysis of the Arecoid line of palms based on plastid DNA sequence data. Mol Phylogenet Evol. 2002;23(2):189-204.

92. Asmussen CB, Dransfield J, Deickmann V, Barfod AS, Pintaud J-C, Baker WJ. A new subfamily classification of the palm family (Arecaceae): evidence from plastid DNA phylogeny. Bot J Linn Soc. 2006;151(1):15-38.

93. Couvreur TL, Hahn W, Granville J-Jd, Pham J-L, Ludena B, Pintaud J-C. Phylogenetic relationships of the cultivated Neotropical palm Bactris gasipaes (Arecaceae) with its wild relatives inferred from chloroplast and nuclear DNA polymorphisms. Syst Bot. 2007;32(3):519-530.

94. Eiserhardt WL, Pintaud J-C, Asmussen-Lange C, Hahn WJ, Bernal R, Balslev H, et al. Phylogeny and divergence times of Bactridinae (Arecaceae, Palmae) based on plastid and nuclear DNA sequences. Taxon. 2011;60(2):485-498.

95. Roncal J, Kahn F, Millan B, Couvreur TL, Pintaud J-C. Cenozoic colonization and diversification patterns of tropical American palms: evidence from Astrocaryum (Arecaceae). Bot J Linn Soc. 2013;171(1):120-139.

96. Noblick LR, Hahn WJ, Griffith MP. Structural cladistic study of Cocoseae, subtribe Attaleinae (Arecaceae): Evaluating taxonomic limits in Attaleinae and the neotropical genus Syagrus. Brittonia. 2013;65(2):232-261.

97. Whittaker RJ, Fernández-Palacios JM. Island biogeography: ecology, evolution, and conservation: Oxford University Press; 2007.

98. Templeton AR. The reality and importance of founder speciation in evolution. BioEssays. 2008;30(5):470-479.

99. Stuessy TF, Takayama K, López-Sepúlveda P, Crawford DJ. Interpretation of patterns of genetic variation in endemic plant species of oceanic islands. Bot J Linn Soc. 2014;174(3):276-288.

100. Cook OF. A synopsis of the palms of Puerto Rico. . The Botanical Club. 281901. p. 565-567.

101. Proctor GR. Arecaceae (Palmae). P Acevedo-Rodríguez, M T Strong. Monocots and Gymnosperms of Puerto Rico and the Virgin Islands.: Smithsonian Institution: Contributions of the United States National Herbarium; 2005.

102. Cavers S, Degen B, Caron H, Lemes MR, Margis R, Salgueiro F, et al. Optimal sampling strategy for estimation of spatial genetic structure in tree populations. Heredity. 2005;95(4):281-289.

103. Morin PA, Martien KK, Taylor BL. Assessing statistical power of SNPs for population structure and conservation studies. Mol Ecol Resour. 2009;9(1):66-73.

Supporting information

S1 Table. Geographical location and origin of the Acrocomia species samples.

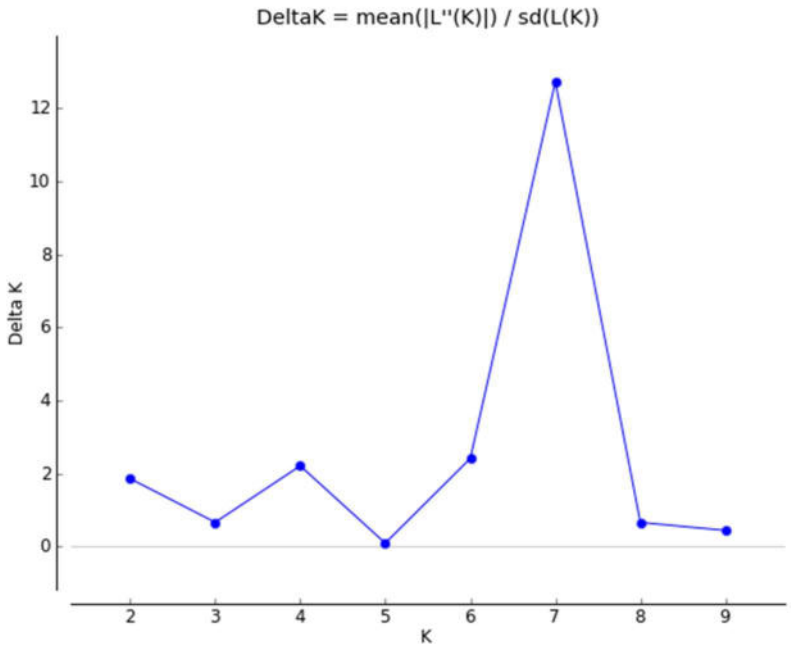
Species	Country	State	Locality	ID	Lat.	Lon.	No. of samples	Sample origin*
A. aculeata	Brazil	São Paulo	Cássia dos Coqueiros	CAS_SP	21°16'58" S	47°10'11" W	3	NP
A. aculeata	Brazil	São Paulo	Rifaina	RIF	19°59'10.6" S	47°30'30.5" W	5	NP
A. aculeata	Brazil	São Paulo	Brotas	BRO	22°16'33.9" S	48°7'6.6" W	5	NP
A. aculeata	Brazil	São Paulo	Fusquinha	FUS	22°6'39.6" S	52°15'8.9" W	5	NP
A. aculeata	Brazil	São Paulo	Brauna	BRA	21°28'56.3" S	50°12'44.2" W	4	NP
A. aculeata	Brazil	Rio de Janeiro	Guapimirim	GUA_RJ	22°32'14"S	42°58'55" W	3	NP
A. aculeata	Brazil	Rio de Janeiro	Itaboraí	ITA_RJ	22°42'69.5"S	42°48'40.6" W	4	NP
A. aculeata	Brazil	Minas Gerais	Montes Claros	MON	16°44'49.99"S	43°53'10.68" W	3	NP
A. aculeata	Brazil	Minas Gerais	Patos de Minas	PMI	18°35'29"S	46°27'18.4" W	4	NP
A. aculeata	Brazil	Minas Gerais	Ibituruna	IBI	21°20'34.96"S	44°44'23.05" W	4	NP
A. aculeata	Brazil	Minas Gerais	Luz	LUZ	19°46'26.30"S	45°51'52.70" W	5	NP
A. aculeata	Brazil	Paraná	Xambrê	XAM	23°44'10"S	53°29'24" W	4	NP
A. aculeata	Brazil	Maranhão	Santa Rita	BR135	3°11'10.99"S	43°1'53.22" W	2	NP
A. aculeata	Brazil	Maranhão	Peritoró	PER	4°31'13.51"S	44°3'13.86" W	3	NP
A. aculeata	Brazil	Maranhão	São José dos Patos	SJP	6° 81' 04.6" S	43° 46' 63" W	3	NP
A. aculeata	Brazil	Tocantins	Palmas	PAL	9°2'38"S	48°19'27.3" W	5	NP
A. aculeata	Brazil	Pará	Belém	BEL	1°08'40.84"S	48°08'45.79" W	4	NP
A. aculeata	Brazil	Goiás	Acreúna	ACR	17°22'40.80"S	50°23'16.80" W	4	NP
A. aculeata	Brazil	Goiás	Trindade	TRI	16°39'46.16"S	49°33'07.24"W	1	NP
A. aculeata	Brazil	Goiás	Brasília, DF	BSB	15°47'38"S	47°52'58" W	3	NP
A. aculeata	Brazil	Mato Grosso	Barra do Garças	BAR	15°44'43.8"S	52°36'57.6"W	1	NP
A. aculeata	Brazil	Mato Grosso	Cáceres	CAC	16°2'21.3"S	57°38'43" W	3	NP
A. aculeata	Costa Rica	San Jose	San Jose	SJO_CR	9°53'50.93"N	84°24'48.09" W	4	NP
A. aculeata	Costa Rica	Guanacaste	Bagaces	GUA_CR	10°31'30.40"N	85°12'25.40" W	4	NP
A. aculeata	Colombia	Casanares	Aguazul	CAS	5°09' 92" N	72°47' 3.0" W	7	NP

Continuation of S1Table.

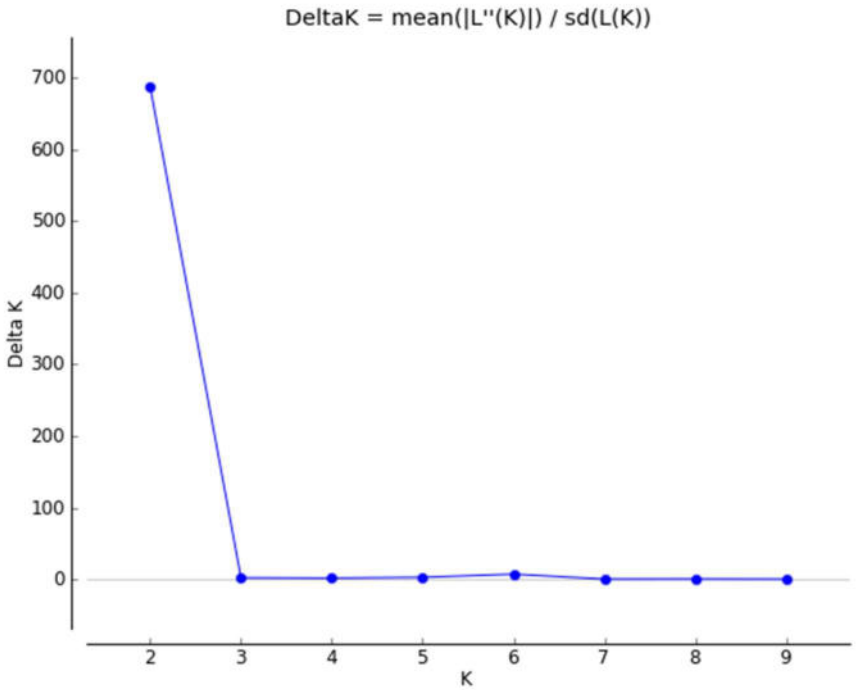
Species	Country	State	Locality	ID	Lat.	Lon.	No. of samples	Sample origin*
A. aculeata	Trinidad and Tobago	Tunapuna piarco	Saint George	TOB	10°39'51.9"N	61°23'56.5" W	6	MBC
A. aculeata	Trinidad and Tobago	Mayaro	Mayaro	TOB	10°19'52.1"N	60°58'48.1"W	2	MBC
A. aculeata	Mexico	Veracruz	Medellin	VC_L1	19°03'11.1"N	96°09'48.9"W	2	NP
A. aculeata	Mexico	Veracruz	La capilla	VC_L2	18°54'24.9" N	96°12'49" W	2	NP
A. aculeata	Mexico	Veracruz	Mangal	VC_L3	19°00'08.8"N	96°09'35.7"W	2	NP
A. aculeata	Mexico	Veracruz	Tinaja	VC_L4	18°54'58.4"N	96°19'29.0"W	1	NP
A. aculeata	Mexico	Quintana Roo	Cancun	CAN	21°9'13.3" N	86°50'31.2" W	1	NP
A. aculeata	Mexico	Chiapas	Tuxtla chico	TUX	14°56´19.5” N	92°11´04.0” W	4	NP
A. aculeata	Mexico	Chiapas	Cacahuatan	CHI	14°58'34.0" N	87 °50'37.2" W	4	NP
A. totai	Brazil	Mato Grosso do Sul	Corumbá	COR	19°21'4.54"S	57°33'49.07" W	6	NP
A. totai	Brazil	Mato Grosso do Sul	Porto Murtinho	PMU	21°33'40.66"S	57°48'40.33" W	4	NP
A. totai	Brazil	Mato Grosso do Sul	Campo Grande	CGR	20°28'8.6"S	54°46'38.5" W	4	NP
A. totai	Brazil	Mato Grosso do Sul	Dourados	DOU	22°15'45.9"S	54°50'16.4" W	5	NP
A. totai	Argentina	Formosa	Col. Pastoril	ARG_25	25°13'58.70"S	58°15'23.90" W	4	NP
A. totai	Argentina	Formosa	Misión Tacaaglé	ARG_26	24°58'53.60"S	58°50'36.50" W	1	NP
A. totai	Paraguai	Itapúa	Bella Vista	PAR	27 14' 14.0" S	55 °35' 49.3" W	8	NP
A. intumescens	Brazil	Pernambuco	Recife	Aint	08°49'5.29"S	48º19'24.1" W	1	MBC
A. intumescens	Brazil	Pernambuco	Exú	EXU	7°50' 63" S	39°72'84" W	1	NP
A. intumescens	Brazil	Pernambuco	Recife	REC	08°49'5.29"S	48°19'24.1" W	2	NP
A. intumescens	Brazil	Ceará	Baturité	BAT	4°19'1.76"S	38°53'33.64" W	5	NP
A. intumescens	Brazil	Ceará	Fortaleza	FOR	3°45'39.23"S	38°31'1.49" W	2	NP
A. media	Puerto Rico	Salinas	Coamo	COA	18°04'50.7" N	66°21'10.8" W	4	MBC
A. media	Puerto Rico	Salinas	Ponce	PON	18°03'31.5" N	66°37'31.8" W	1	MBC
A. crispa	Cuba	Camagüey	Camagüey	CUB	21°23'37.0"N	77°52'29.0"W	5	MBC
A. hassleri	Brazil	São Paulo	Nova Odessa	Ahas	22° 46' 49"S	47° 18' 48" W	1	IP
A. glaucescens	Brazil	São Paulo	Nova Odessa	Agla	23° 46' 49"S	48° 18' 48" W	1	IP

*NP= Native population, MBC = Living Collection Montgomery Botanical Center, IP= Living Collection Instituto Plantarum

S1 Fig. Delta (Δ) K values for different numbers of populations assumed (K) in the STRUCTURE analysis, estimated based on Evanno method for all Acrocomia species.



S2 Fig. Delta (Δ) K values for different numbers of populations assumed (K) in the STRUCTURE analysis, estimated based on Evanno method for *A. aculeata.*



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

Brenda G. Díaz, Maria I. Zucchi, Alessandro Alves-Pereira, Joaquin A. Azevedo-Filho, Mariana Sanitá, Carlos A. Colombo

Artigo submetido para publicação na revista Ecology and Evolution

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

Brenda G. Díaz[a], Maria I. Zucchi[a,c], Alessandro Alves-Pereira[a], Joaquin A. Azevedo-Filho[b], Mariana Sanitá[c] Carlos A. Colombo[c*]

[a] University of Campinas UNICAMP, Institute of Biology, Cidade Universitária Zeferino Vaz, R. Monteiro Lobato 255, Campinas-SP, Brazil, 13083-862.

Brenda G Díaz 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

[b] Polo Regional Leste Paulista, Secretaria da Agricultura do estado de São Paulo, Estrada Vicinal Nelson Taufic Nacif, Km 03, Monte Alegre do Sul-SP, Brazil, 13910-000.

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

[c] Instituto Agronômico do estado de São Paulo-IAC, Centro de Pesquisa de Recursos Genéticos Vegetais, Av. Theodureto de Almeida Camargo, 1500, Campinas-SP, Brazil, 13075-630, Brazil.

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

Mariana Sanitá 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 (HE=0.551) than in A. aculeata (HE=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: Macaúba 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 Pará, Maranhão, Ceará, Minas Gerais, Goiás, Mato Grosso, Mato Grosso do Sul, Sao Paulo, Paraná, 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; Rodríguez 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 São 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; César 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 São 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.

ID	Population	Estate*	Country	No. Plants	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	Brauna	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
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

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).

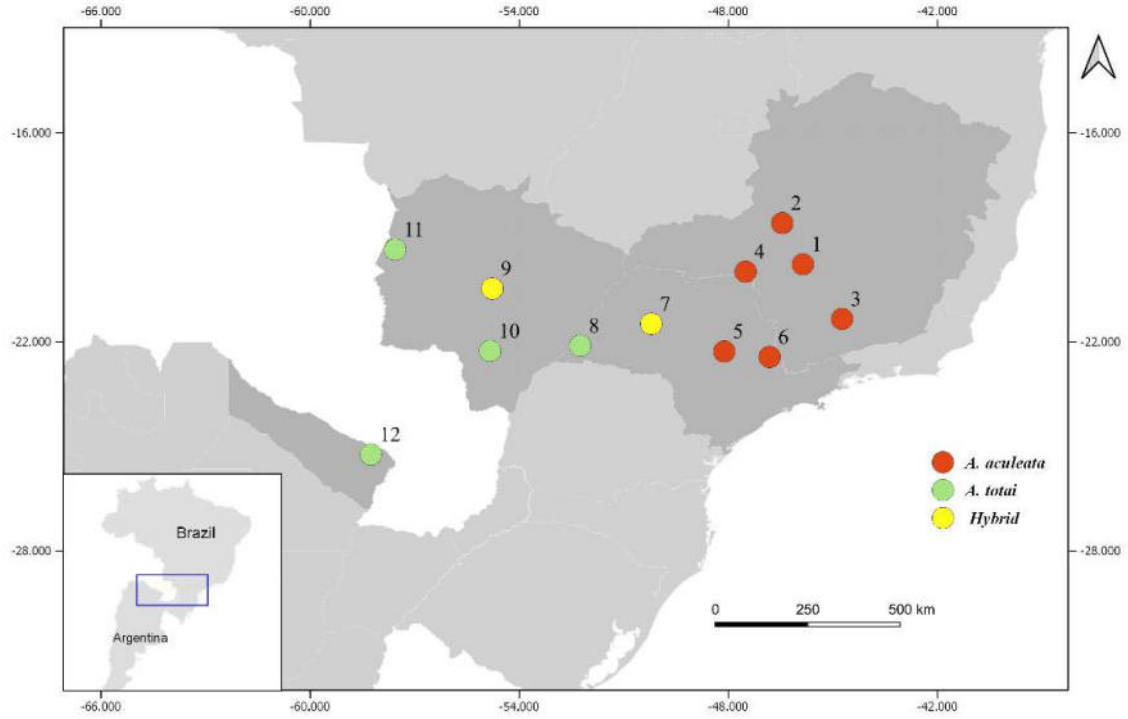


Figure 1. Map of geographic distribution of *Acrocomia spp.* sampled populations

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 geneti

Primer Name	Primer sequence (5´-3´)	Motif	TA (°C)	Amplicon size	Na	Pa	PI
Aacu10	F: TGCCACATAGAGTGCTTGCT R: CTACCACATCCCCGTGAGTT	(AG)16	56	168–186	10	3	0.7
Aacu12	F: GAATGTGCGTGCTCAAAATG R: AATGCCAAGTGACCAAGTCC	(TC)20	56	190–202	12	3	0.7
Aacu26	F: ACTTGCAGCCCCATATTCAG R: CAGGAACAGAGGCAAGTTC	(AC)13(AG)14	56	273–316	14	4	0.8
Aacu38	F: TTCTCAGTTTCGTGCGTGAG R: GGGAGGCATGAGGAATACAA	(TC)15	56	316–346	11	2	0.8
Aacu45	F: CAGACTACCAGGCTTCCAGC R: TCATCATCGCAGCTTGACTC	(CGAC)5	56	260–284	6	1	0.5
gSSR mean					**10.60**	**2.60**	**0.7**
Acro16	F: GTCATATGGCTGGTGAGATT R: GTTCCTTCTCTTGGTGGAAT	(GCC)8	55	270	7	3	0.4
Acro20	F: CCACCCTTAAGTTCATCTTCT R: GACTGTTGGTGTTAAGGTTCA	(CCT)8	55	294	5	1	0.5
Acro46	F: CAGATTATAGCACAGCTGGAG R: AGTGACTTGAAGCTCATGTTG	(ATC)5	55	398	11	2	0.7
Acro64	F: GTATGGATGTCGTCGTTGAT R: GACTATGGTAATGGACCAACA	(CTG)12	55	168	10	2	0.7
Acro102	F: GGCTAAGATCATTAATGGGAC R: GGACCATACCAATTTCCTTAG	(CCA)5	55	268	3	0	0.2
Acro246	F: GAGAAGGTAAGGTAGACGAGG R: ATGGATCAAGAACCCGAC	(GGC)9	57	275	7	3	0.2
Acro280	F: ATCTGAGACTGAAGCTGATGA R: GATCTGCATACATCCATCTGT	(GAA)5	55	204	3	2	0.2
Acro278	F: GAAGAGTTTTCCTCTCTGCTC R: AGATGCCCTATTGCTCAAG	(CCG)6	55	311	2	0	0.1
EST-SSR mean					**6.0**	**1.63**	**0.4**

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

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).

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 Corumbá (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). Braúna 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.

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	**Brauna**	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).

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 (S1 Table).

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

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 I grouped most individuals from populations ascribed as *A. aculeata,* and cluster II 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 Cluster I conformed by the *A. aculeata* populations and Cluster II by population of *A. totai*, 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.)

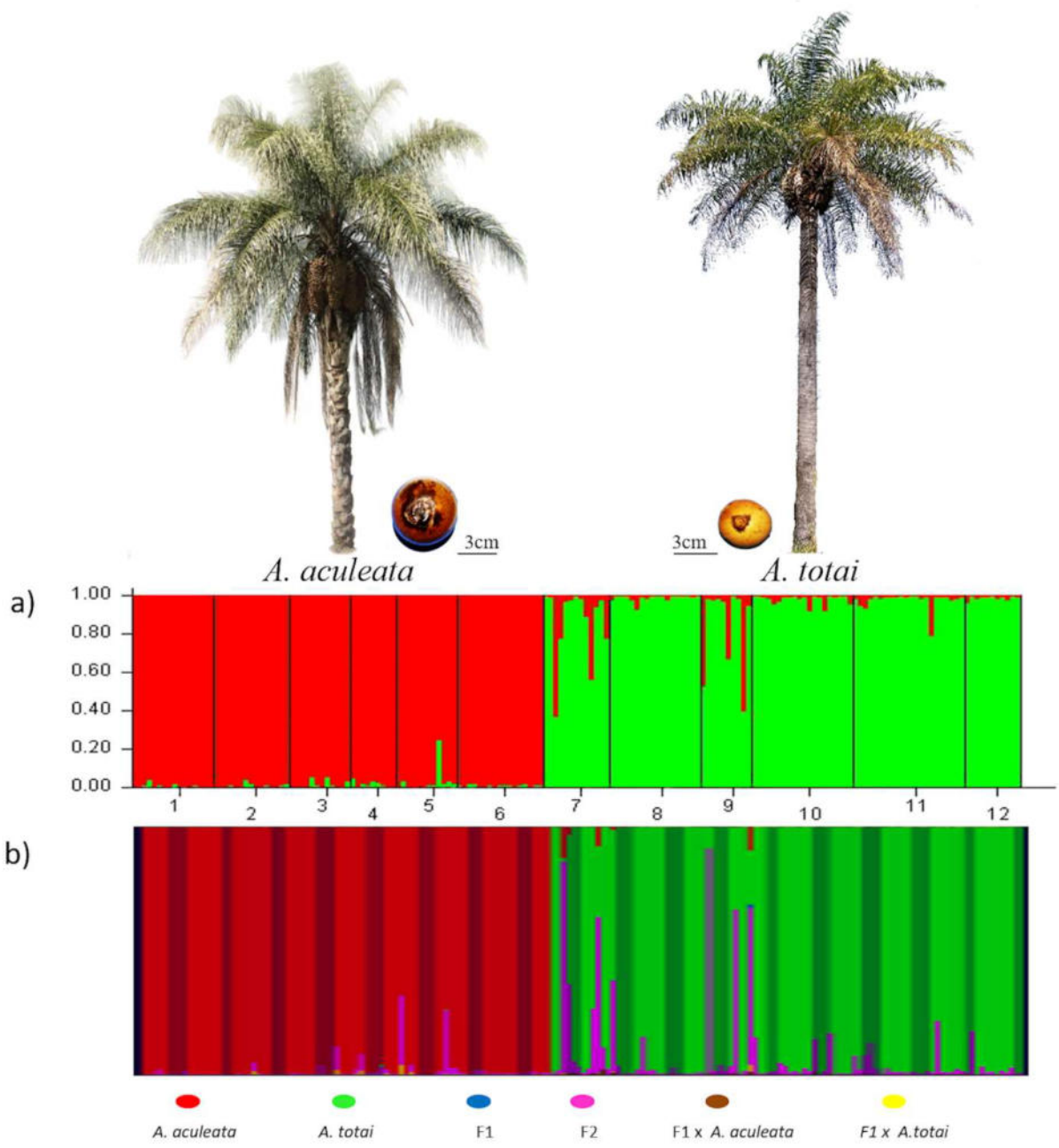


Figure 2. Results of Bayesian clustering assignment with the software's a) STRUCTURE (K=2) and b) NewHybrids. Individuals are represented by vertical bars where the color represents the posterior probability of assignment to each group.

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 II). 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).

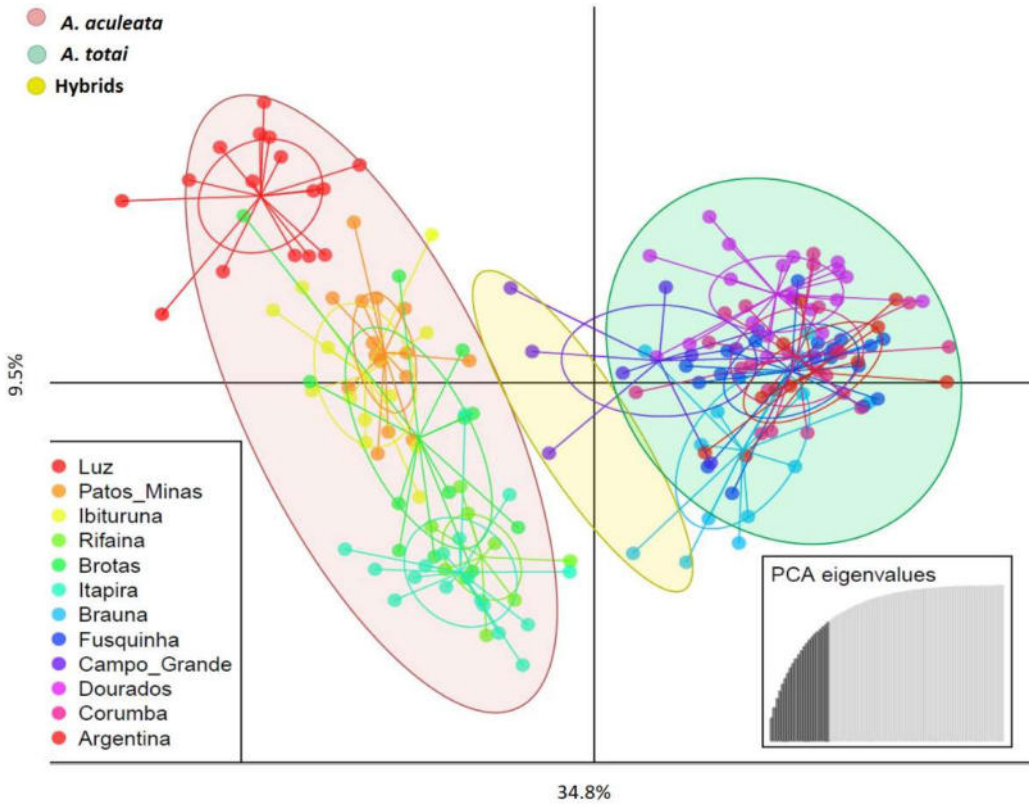


Figure 3. Discriminant analysis of principal components (DAPC) among populations of *A. aculeata* (Cluster I) and *A. totai* (Cluster II) based on the microsatellite data set. Only the two-first axes showing the two higher discriminant eigenvalues are presented.

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 divergent 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.

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 etnobotânica de palmeiras no Território Indígena Krahô. Tocantins. Brasil. Floresta 40:209-220. http://dx.doi.org/10.5380/rf.v40i1.17112

Neiva DS, Melo Júnior AF, Oliveira DA, Royo VA, Brandão 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, Brandão 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.

Patiño RVM (2002) Historia y dispersión 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, Bordács S, Burg K, Coart E, Cottrell J, van Dam B, Deans JD, Dumolin-Lapègue S, Fineschi S, Finkelday R, Gillies A, Gla I, Goicoechea PG, Jensen JS, König A, Lowe AJ, Madsen SF, Mátyás 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, Mejía K, Millán B, Moraes M, Noblick L, Stauffer FW, Kahn F (2008) Las palmeras de América del Sur: diversidad, distribución e história 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/

Ramírez-Rodríguez R, Tovar-Sánchez E, Jiménez RJ, Vega FK, Rodríguez 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

Rodríguez 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) Floração e Frutificação da Bocaiúva (*Acrocomia aculeata)* e do Carandá (*Copernicia alba*) no Pantanal. Corumbá: Embrapa Pantanal. Comunicado Técnico 78.

Sanjinez-Argandoña EJ, Machado CCA (2011) Caracterização biométrica, física e química 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

Schönswetter P, Solstad H, García 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 Àrvore 41:1-9. http://dx.doi.org/10.1590/1806-90882017000500008

Slovák M, Kucera J, Marhold K, Zozomová-Lihová 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-Sánchez E, Mussali-Galante P, Esteban-Jiménez R, Piñero D, Arias DM, Dorado O, Oyama K (2008) Chloroplast DNA polymorphism reveals geographic structure and introgression in the *Quercus crassipes* × *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, Jordão 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: território e ecologia do povo Xavante de Wedezé. Museu do Índio-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

Supplementary material

S1 Table. Matrix of pairwise F_{ST} values (below diagonal) among populations based on 13 microsatellites.

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

	Luz	Patos Minas	Ibituruna	Rifaina	Brotas	Itapira	Brauna	Fusquinha	Campo Grande	Dourados	Corumbá	Argentina
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						
Brauna	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	0.000		
Corumbá	0.172	0.176	0.119	0.131	0.148	0.175	0.109	0.064	0.123	0.056	0.000	
Argentina	0.287	0.210	0.213	0.240	0.235	0.283	0.134	0.093	0.212	0.113	0.136	0.000

CAPÍTULO III - Mating system and genetic diversity of open-pollinated commercial progenies of *Acrocomia aculeata* to predict their genetic vulnerability

Brenda G. Díaz, Maria I. Zucchi, Mariana Sanitá, Joaquim A. Azevedo-Filho, Carlos A. Colombo

Artigo submetido para publicação na revista Forest Ecology and Mangement

Mating system and genetic diversity of open-pollinated commercial progenies of *Acrocomia aculeata* to predict their genetic vulnerability

Brenda G. Díaz[a], Maria I. Zucchi[a,b*], Mariana Sanitá[b], Joaquim A. Azevedo-Filho[b], Carlos A. Colombo[b]

[a] University of Campinas UNICAMP, Institute of Biology, Cidade Universitária Zeferino Vaz, R. Monteiro Lobato 255, Campinas-SP, Brazil, 13083-862.

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

[b] Instituto Agronômico do estado de São Paulo-IAC, Centro de Pesquisa de Recursos Genéticos Vegetais, Av. Theodureto de Almeida Camargo, 1500, Campinas-SP, Brazil, 13075-630.

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

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

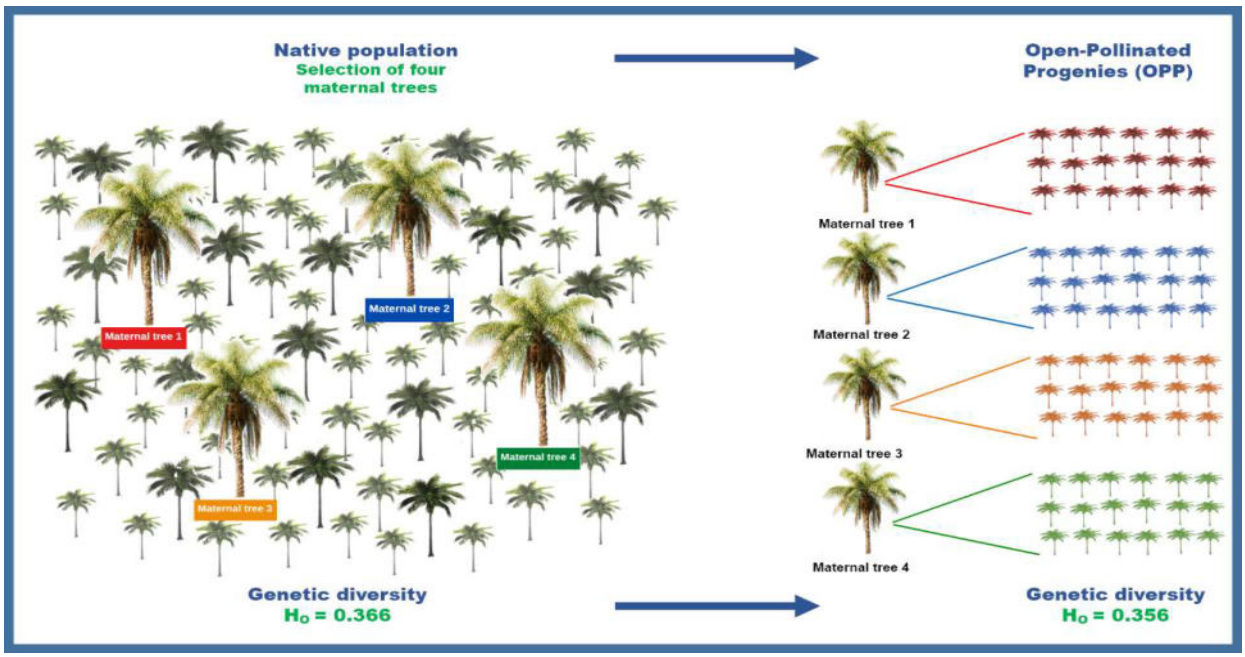
Carlos A. Colombo e-mail carlos.colombo@sp.gov.br ORCID: 0000-0003-3269-6068

Maira I Zucchi e-mail: mizucchi@unicamp.br ORCID: 0000-0002-4863-1843

*Corresponding author:

E-mail: mizucchi@unicamp.br (MIZ)

Graphical abstract



Abstract

At present, intelligent and sustainable agriculture presupposes a compromise between maximizing crop yield and minimizing its risk of genetic vulnerability. Breeding programs should implement strategies that allow the development of superior genotypes and preserve genetic variability, especially in the case of perennial plants. *Acrocomia aculeata* is a species with the vocation to supply oil with the same production capacity as the oil palm even in areas with water deficit, has been motivating its commercial interest and increasing the demand for seedlings for commercial plantations from natural populations. Therefore, to evaluate the impact of this selection procedure on the reduction of genetic variability and, consequently, increase in the genetic vulnerability of commercial planting, in the present study we analyzed the population mating system and the genetic diversity of open-pollinated progenies (OPP) from selected matrices in relation to the genetic diversity of respective native populations. The genetic diversity of twenty-five plants and four open-pollinated progenies consisting of 18 individuals from four Brazilian populations was performed using the allelic variation provided by 15 microsatellite markers (SSR). The mating system indices estimated for *A. aculeata* palm indicated that the species present a mixed mating system (*tm* = 0.877) with a predominance of outcrossing. At the population level, the outcrossing estimated was t=1.00. The biparental inbreeding rate was consider moderate (*tm-ts* = 0.303), with some variation between population (from 0.203 to 0.501). The multilocus correlated paternity (*rp*= 0), indicating that all the progeny was composed of half-sibs. The results show that there was no significant loss of the genetic diversity of OPP derived from an artificial selection within natural populations. The OPP (Na = 2.5) and polymorphism (P = 75%) was similar to that of native populations (Na = 2.5 and P = 78.3%). Likewise, the mean values of observed heterozygosity (Ho) between native populations and OPP were similar, with a mean of 0.366 and 0.356, respectively. In conclusion, the procedure adopted to produce commercial seedlings should not offer a high risk of genetic vulnerability. In addition, it will guide breeders on selection intensity in both natural and improved populations for the advancement of generations with a view to maximizing future genetic gains and maintaining genetic diversity.

Keywords: Macauba palm, Vegetable oil, Domestication, Conservation, Plant selection

1. Introduction

The world production of vegetable oils has been growing annually, having jumped from 171 to 204 million tons from 2014 to 2020 (USDA, 2020). In the last decades, there has been an increasing application of vegetable oils in the food industry, cosmetics, chemical oil, and biofuels, a demand

that has been supplied by commodities such as palm oil, soybean, canola, sunflower, cotton, olive, corn and peanuts (Byerlee et al., 2017), with strong participation of palm oil and soybeans, which together represent 68% of all vegetable oil produced worldwide (USDA, 2020). According to the OECD/FAO (2019), demand for vegetable oils will continue to grow. Given the context of high current demand and growing future demand and the situation of dependence on only two raw materials, the diversification of the sources of production of vegetable oils proves to be an important opportunity for other plant species.

However, the choice of raw material to produce vegetable oils must aim not only at economic viability but also at environmental sustainability. In this sense, the cultivation of perennial species can provide an important reduction in environmental impacts compared to the agricultural systems of annual species, through various environmental services such as soil conservation, improving water quality, and improving wildlife habitat (Kantar et al., 2014). In addition to being more efficient in carbon sequestration and significantly reducing nitrogen losses (Pugesgaard et al., 2015)

In this sense, the macaúba palm [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.] has been identified as a promising species to diversify the production of vegetable oil, due to the quantity and quality of the oil produced in its fruits and the generation of high added value by-products, which can meet the demand of the food, cosmetic and, above all, energetic (Bora and Rocha, 2004; Coimbra and Jorge, 2011; Berton et al., 2013).

The potential for oil production of *A. aculeata* is surpassed only by the palm, currently the species with the highest oil yield per unit area (Amaral, 2007). Another advantage of the macaúba palm is its wide occurrence in the American continent, with wide geographic distribution throughout the American tropics from Mexico to the south of Brazil (Henderson et al., 1995; Lorenzi et al., 2010), including areas with levels of precipitation below 1000 mm year $^{-1}$ (Colombo et al., 2018).

With all these advantages, the volume of scientific studies on the *A. aculeata* is growing, with emphasis on research in the agronomic area, whose challenges are associated with the incipient degree of domestication of the species (Colombo et al., 2018). In addition to studies on agronomic assessment of natural populations for selection purposes (Oliveira et al., 2012; Domiciano et al., 2015; Díaz-Fuentes et al., 2019; dos Reis et al., 2019; Alfaro-Solís et al., 2020), studies have also been conducted to understand the genetic diversity and the preferential reproduction system of the species (Abreu et al., 2012; Lanes et al., 2016; Coelho et al., 2018)

The understanding of the organization of the genetic diversity of the species is a starting point for activities of management of genetic resources or to use this diversity for purposes of genetic

improvement. Likewise, understanding the impacts of genetic improvement on plant genetic diversity is also an important study approach. The reproductive system of a species is one of the main factors that influence the levels and dynamics of genetic diversity and structure (Wright, 1940; Charlesworth and Wright, 2001; Sebbenn, 2006) and determines how genes are transferred from one generation to the next and how genes are recombined (Ritland, 2002; Vinson et al., 2018). Thus, it is essential to understand the mating system to define improvement and conservation strategies, as well as to address evolutionary and ecological issues of species.

Microsatellite markers or Simple Sequences Repeats (SSR) have been widely used for several genetic studies in plants due to their abundance, codominant inheritance, and high levels of polymorphism (Kalia et al., 2011; Mason, 2015). In *A. aculeata* these markers have been used to evaluate the diversity and genetic structure of natural populations (Nucci, 2007; Abreu et al., 2012; Lanes et al., 2015; Coelho et al., 2018)and germplasm banks (Mengistu et al., 2016), the reproductive system (Abreu et al., 2012; Lanes et al., 2016; Coelho et al., 2018) and inbreeding depression (Simiqueli et al., 2018).

The degree of knowledge of the reproductive system of the macaúba palm is limited and controversial. According to Nucci et al. (2008) and Abreu et al. (2012), the species has a mixed reproduction system. However, Lanes et al. (2016) and Coelho et al. (2018) indicate that the species is preferably allogamous. In any case, previous studies on the mating system have focused on the discussion at the species level (Abreu et al., 2012; Lanes et al., 2016) or population (Coelho et al., 2018). However, the pattern of the breeding system may vary depending on the population (Sebbenn, 2006; Kubota et al., 2008; Whitehead et al., 2018) and, therefore, have central implications in the definition of breeding strategies and conservation of genetic diversity.

In this sense, the present study aims to: 1) evaluate the crossing system of *A. aculeata* and characterize the crossing rate and/or self-fertilization at the population level, and 2) Compare the effects of selection on the genetic diversity of pollination progenies in relation to native populations and their matrices that gave rise to these progenies.

2. Material and methods

2.1. Plant material

For the present study, four natural populations were selected, two from the state of São Paulo and two from the state of Minas Gerais according to Table 1.

Table 1. Geographic location of *A. aculeata* analyzed populations.

ID	Population	State*	Latitude	Longitude	No. Plants of native population	No. of Maternal trees	No. of Progeny plants
LUZ	Luz	MG	19°46'48.7"S	45°51'00.1"W	25	4	72
PAT	Patos de Minas	MG	18°50'46.0"S	46°27'28.3"W	25	4	72
ITA	Itapira	SP	22°26'00" S	46°49'18" W	25	4	72
DOU	Dourado	SP	22°09'18.7"S	48°19'43.0"W	25	3	54

MG = Minas gerais, SP = São Paulo

To analyze the genetic diversity of each population, 24 plants were considered. Of these, 4 were chosen according to criteria of agronomic superiority for collecting fruits and obtaining their progenies. Thus, the analysis of genetic diversity in the progenies was carried out from 15 families of open pollination (four families from each population, except Dourado, with three families), each family represented by 18 open-pollinated progenies (OPP).

The seeds were obtained and germinated under controlled laboratory conditions and germination chambers, according to the protocol established by Berton et al. (2013). One month after germination, they were transplanted into paperpot bags (850 mL) and kept in a greenhouse for acclimatization. After a month and a half in a greenhouse, leaf samples were collected from the 18 individuals of each matrix for genetic studies. For the analysis of the mating system, only the 15 families of open pollination were used.

2.2. DNA extraction and genotyping of microsatellite loci

The total genomic DNA was isolated from leaf material using the protocol of Doyle and Doyle (1990). The quality and quantity of DNA were evaluated on a 1% agarose gel and on the NanoVue ™ Plus Spectrophotometer (GE Healthcare).

To amplify the SSR loci, 15 primer pairs developed for *A. aculeata*, five from genomic regions (gSSR) and ten from genetic regions (EST-SSR) developed by Nucci et al. (2008) and Bazzo et al. (2018), respectively (Table 2). The PCR reactions were performed in a volume of 15 µL containing 20 ng of DNA, 2.0 µL of each forward and reverse primer (5 µM / µL), 3 µL of Hot Start PCR Master Mix (2X), and 8.2 µL of ultrapure water. The conditions of the PCR reaction were: initial denaturation at 94 ° C for 2 min, followed by 30 cycles at 94 ° C for 1 min, from 55 ° C to 58 ° C, depending on the primer requirement (Table 2) for 1 min, 72 ° C for 1 min and a final extension at 72 ° C for 10 min.

The amplification products were separated by capillary electrophoresis using a 96-capillary Fragment Analyzer ™ Automated CE System (Advanced Analytical Technologies, Ames, IA, USA) using the DNF-905 double-stranded DNA reagent kit (Advanced Analytical Technologies, Ames, IA, USA).

2.3. Analysis of mating system

The parameters for the characterization of the reproductive system were estimated using the MLTR version 3.4 software (Ritland, 2002), assuming the mixed reproduction model (Ritland and Jain, 1981) and based on the multilocus genotype observed in progenies of known matrices. The estimated parameters were: multi-locus crossing rate (*tm*) , single-locus crossing rate (*ts*), biparental inbreeding or crossing between related individuals (*tm-ts*), inbreeding coefficient of the matrices (or maternal genotype) (f_m), multi-locus correlation paternity (*rp*) and self-fertilization correlation (*rs*). The 95% confidence interval was calculated based on the standard error estimated from 1000 bootstraps.

2.4. Analysis of genetic diversity

For each microsatellite locus, the number of alleles per locus (Na), the observed and expected heterozygosity (H_O and H_E), the polymorphism index and inbreeding (F) were estimated. To compare the genetic diversity of the population in relation to the genetic diversity in the OPPs, the following were considered: number of alleles (Na), effective number of alleles (Ne), Shanon index (I), observed heterozygosity (H_O), expected heterozygosity (H_E). Estimates of genetic parameters were obtained using the GenAlEx 6.5 program (Peakall and Smouse, 2006). The content of polymorphic information (Picanço-Rodrigues et al.) was calculated using the MSTools software (Park, 2001).

3. Results and discussion

3.1. Genetic variability of microsatellites

The 15 microsatellite loci analyzed were polymorphic in all populations and progenies evaluated. A total of 68 alleles were detected and an average of 4.5 alleles per locus, ranging from two (Acro 39) to eight (Aacu26) alleles (Table 2). This result confirms the polymorphism found with this marker in *A. aculeata* (Nucci et al., 2008; Lanes et al., 2015; Lanes et al., 2016; Mengistu et al., 2016; Bazzo et al., 2018).

Table 2. Characteristics of 15 microsatellite loci and estimates of genetic diversity parameters for each locus.

Primer	SSR Type	Primer sequence (5´-3´)	Motif	TA (°C)	Amplicon size	A	H_E	H_O	PIC	F
Acro201	EST-SSR	F:GGAGCTAAAGATGAGGAGAAG R:CGGATGGGTGTAGAAGTTATT	(CAG)5	55	312 - 441	3	0.03	0.00	0.31698	1.000
Acro203	EST-SSR	F:AGGGTGCCTATTTTGAGATAC R:GATGACGTCTCCCTCCTC	(AGG)5	58	164 - 227	4	0.37	0.46	0.4539	-0.169
Acro225	EST-SSR	F:AGTCTTGAACTCATTCAGCAG R:TACTTCAGAGTCTTAGCCACG	(GTG)5	58	216 - 285	2	0.01	0.00	0.00595	1.000
Acro227	EST-SSR	F:CAGAGACGACATAGGAATCTG R:GCTGCTCTATTTCAACTCAAG	(GGT)6	55	348 - 417	4	0.13	0.11	0.41801	0.380
Acro111	EST-SSR	F:CAGAAAAGAGACCAGTTGATG R:GTCACAGATTCTAAGCATTGG	(ATG)5	55	154 - 187	3	0.47	0.78	0.37421	-0.622
Acro126	EST-SSR	F:CCTATGCTGCTAGGTAATTCA R:GTTGAGTCTGAAAGGGAGAAG	(TTC)9	55	302 - 350	5	0.49	0.43	0.66618	0.110
Acro39	EST-SSR	F:GATGTTATGCTCAACTCCATC R:GTACCATAGCTTCCTCCACTT	(GCG)7	55	350 - 401	1	0.00	0.00	0	–
Acro93	EST-SSR	F:ATCTAGTCTCAGCCTGACCA R:GAAGGGGAGAGAGAGGAGT	(GCA)5	58	150 - 213	5	0.43	0.42	0.48998	0.062
Acro64	EST-SSR	F: GTATGGATGTCGTCGTTGAT R: GACTATGGTAATGGACCAACA	(CTG)12	55	159 - 180	6	0.51	0.74	0.42924	-0.456
Acro16	EST-SSR	F: GTCATATGGCTGGTGAGATT R: GTTCCTTCTCTTGGTGGAAT	(GCC)8	55	270	6	0.59	0.95	0.73569	-0.659
Aacu07	gSSR	F:ATCGAAGGCCCTCCAATACT R:AAATAAGGGGACCCTCCAA	(GA)13	56	153-177	7	0.39	0.34	0.39636	0.063
Aacu26	gSSR	F: ACTTGCAGCCCCATATTCAG R: CAGGAACAGAGGCAAGTTC	(AC)13 (AG)14	56	273–316	3	0.19	0.18	0.33932	0.024
Aacu38	gSSR	F:TTCTCAGTTTCGTGCGTGAG R: GGGAGGCATGAGGAATACAA	(TC)15	56	316–346	8	0.47	0.21	0.47893	0.563
Aacu45	gSSR	F: CAGACTACCAGGCTTCCAGC R: TCATCATCGCAGCTTGACTC	(CGAC)5	56	260–284	7	0.46	0.45	0.65745	0.038
Aacu74	gSSR	F:TACTGTTGTGCCAAGTCCCA R:GAGCACAAGGGGGATATCAA	(CA)15	56	278–313	4	0.34	0.29	0.49073	0.136
Mean							0.32	0.36	0.42	0.11

A: number of alleles per locus; H_O: observed heterozygosity; H_E: expected heterozygosity; F: inbreeding coefficient; PIC: polymorphic information content.

The average observed (H_O) and expected (H_E) heterozygosity were 0.35 and 0.32, respectively. The highest values were recorded at the Acro16 locus with Ho = 0.95 and H_E = 0.59, although the Acro39 locus did not reveal allelic polymorphism and, therefore, null values for both estimates (Table 2). At the average level, the results of the diversity estimates (H_O, H_E, F and PIC) were lower than those reported by Lanes et al. (2016) and (Mengistu et al., 2016) using microsatellite markers to analyze genetic diversity in natural populations of *A. aculeata*. However, similar to those reported by Lanes et al. (2016) and (Mengistu, 2015) in diversity analyzes in progenies of *A. aculeata* from open pollination.

3.2 Mating system

The rate of biparental inbreeding, both at the species and population levels, suggests that some of these crossings are occurring between related individuals. This type of mating occurs when populations exhibit a spatial genetic structure (Ritland and El-Kassaby, 1985; Ritland, 1988). The spatial grouping of related genotypes can be expected in species such as *A. aculeata* with barochoric dispersion, that is, the seeds are dispersed by gravity with a high probability of remaining close to the mother trees (Abreu et al., 2012; Lanes et al., 2015; Lanes et al., 2016). This phenomenon has been widely characterized in other plant species (Murawski and Hamrick, 1991; Souza et al., 2003; Vinson et al., 2018)

The diversity of pollen detected in progeny matrices can be described by the paternity correlation (rp), which measures the probability that two individuals taken at random from the same matrix are complete siblings (Ritland, 2002). In the present study, the rp found for *A. aculeata* was 0.406. Similar values were reported by Abreu et al. (2012) and Lanes et al. (2016). At the population level, the paternity correlation (rp) showed low and negative values, suggesting that, at the population level, the progenies originated from crosses with several pollen donors and that all their individuals are half-siblings. A similar result was reported by Coelho et al. (2018), who identified low values for the rp. The differences reported in the studies cited above may be the result of several factors, such as flowering synchrony, pollen movement patterns, plant density in populations and pollinator behavior (Murawski and Hamrick, 1992; Barrett, 2003; Breed et al., 2014; Vinson et al., 2018).

Even having low paternity correlation values (rp), it is probable that part of the progeny individuals are related at a higher level than the half-siblings, even if they originated by crossing. This claim can also be supported based on the levels of biparental inbreeding detected**.**

3.3. Genetic diversity in native populations and open pollination progenies

In the present study, we used data produced by 15 microsatellite loci to estimate diversity in native populations and their respective open pollination progenies to assess and compare changes in genetic diversity in the first selection cycle. The evaluation of the variation in the genetic diversity of *A. aculeata* from one generation to another is important to understand its effects not only for the purpose of genetic improvement of the species but also to make predictions of the degree of genetic variability that may be present in progenies obtained from few matrices, given that the current commercial plantations of the species have been carried out with seeds collected from these selected matrices.

Table 3. Estimates of the mating system parameters for *A. aculeata* at population and species level. Values in parentheses represent standard errors

Parameter	Population level				Species level
	LUZ	**PAT**	**ITA**	**DOU**	
Number of maternal plants (m)	4	4	4	3	15
Number of Progeny plants (n)	72	72	72	54	270
Multilocus outcrossing rate (*tm*)	1.200 (0.065)	1.200 (0.076)	1.050 (0.181)	1.200 (0.000)	0.877 (0.153)
Single-locus outcrossing rate *(ts)*	0.821 (0.124)	0.699 (0.111)	0.717 (0.125)	0.997 (0.047)	0.564 (0.047)
Biparental inbreeding (*tm* - *ts*)	0.379 (0.118)	0.501 (0.109)	0.333 (0.197)	0.203 (0.047)	0.313 (0.135)
Correlation of selfing (*rs*)	-0.200 (0.046)	-0.200 (0.049)	-0.200 (0.328)	-0.200 (0.039)	0.406 (0.261)
Multilocus correlation of paternity r_p	0.098 (0.121)	-0.200 (0.049)	-0.200 (0.087)	-0.046 (0.067)	0.406 (0.146)
Fixation index for maternal plants F	-0.190 (0.027)	-0.200 (0.001)	-0.111 (0.043)	-0.200 (0.001)	0.099 (0.053)

The estimates of the genetic diversity reveal that the mean number of alleles per locus was the same in both native populations and in OPP, that is, Na = 2,567 and Na = 2,550, respectively (Table 4). Among native populations, this same index was higher in the population of Patos de Minas (Na = 2,800) and lower in the populations of Luz and Itapira (Na = 2,400). In the case of open pollination progenies, the population of Dourado revealed the highest Na (2,933), while the lowest Na value was detected for the population of Itapira (2,267).

The average percentage of polymorphism (P%) ranged from 78%, in the natural population to 75% in the OPP, a value that was reduced mainly due to the population of Itapira, which was from 86.6% in the native population to 73.3% in the OPP. In the case of the populations of Luz, Patos de Minas and Dourado the value of P% of the population and their respective progenies remained the same.

The observed average heterozygosity (H_O) was higher than expected (H_E) both in native populations and in their respective OPP, indicating an excess of heterozygotes, as well as negative values for the fixation index. On average, native populations and their OPPs exhibited similar values of average observed heterozygosity, $H_O = 0.366$ and $H_O = 0.356$, respectively. In native populations, the observed heterozygosity (H_O) was higher for Itapira ($H_O = 0.379$), while the highest H_O value for OPP was found for Patos de Minas (($H_O = 0.410$) (Table 4).

The mean value of the mean index (F) found in the populations was negative, except for the population of Itapira. In the case of OPPs, all of them also presented negative values, showing a significant deviation from the HW balance as result of the excess of heterozygotes, as previously mentioned. The excess of heterozygotes has been reported for other palm species with a high crossing rate, as reported by Eguiarte et al. (1992) in *Astrocaryum mexicanum* and González-Pérez et al. (2004) in *Phoenix canariensis.*

The excessive number of heterozygotes identified in all populations of the present study and in their respective open pollination progenies, together with the high crossing rate observed in the species *Acrocomia aculeata*, may be the result of self-incompatibility, as suggested by (Abreu et al., 2012). Another possible explanation is that inbreeding depression is also contributing to the reduction of homozygotes. Selection against homozygotes has already been reported in *A. aculeata* by Lanes et al. (2016) and Simiqueli et al. (2018), as was suggested in other palm species, as reported by Picanço-Rodrigues et al. (2015) in a study conducted with the species *Bactris gasipaes*. Selection against homozygotes can eliminate part of the inbreeding generated by self-fertilization, as well as by biparental inbreeding, as identified by

the values of *tm-ts* and, thus, favoring the occurrence of high heterozygosity and the presence of low fixation index as observed in the present study.

The estimated genetic diversity parameters (H_O and H_E) were similar to those found in previous studies in natural populations (Lanes et al., 2015; Mengistu, 2015) and in open pollination progenies (Lanes et al., 2016; Coelho et al., 2018). High levels of genetic diversity are expected in species with high breeding rates. Likewise, a high number of pollen donors, as identified in this study, can contribute to the maintenance of genetic diversity, as also suggested for *Bactris gasipaes* (Picanço-Rodrigues et al., 2015).

The degree of genetic diversity in *A. aculeata* revealed by the genetic parameters adopted for the present study proved to be accentuated both for the populations analyzed and for their respective open pollination progenies. In temperate tree species, a similar result was also observed, showing a low reduction in the heterozygosity of progenies in relation to the native populations from which were derived (Lefèvre, 2004). Likewise, high levels of diversity have been reported in other palm species such as *Cocos nucifera* L. (Liu et al., 2011), *Elaies guineensis* (Billotte et al., 2001), and *Phoenix dactylifera* (Zehdi et al., 2004), species that have undergone several processes of domestication and or improvement, with low reduction of their respective genetic diversity.

Table 4. Genetic parameter estimates in native populations and open-pollinated progenies (OPP) microsatellites markers.

Parameter	Native population					Progen		
	LUZ	PAT	ITA	DOU	Total mean	LUZ	PAT	ITA
N	25	25	25	25	25	72	72	72
Na	2.400	2.800	2.400	2.667	2.567	2.400	2.600	2.267
Ne	1.634	1.866	1.676	1.759	1.734	1.614	1.780	1.434
P%	73.3	80.0	86.6	73.3	78.3	73.3	80.0	73.3
I	0.515	0.638	0.570	0.599	0.580	0.505	0.611	0.402
Ho	0.346	0.376	0.379	0.362	0.366	0.347	0.410	0.304
H_E	0.307	0.372	0.358	0.343	0.345	0.307	0.364	0.248
F	-0.034	-0.039	0.028	-0.027	-0.016	-0.098	-0.103	-0.110

N: average of individuals analyzed; Na: average number of alleles per polymorphic loci and per provenance; Ne: effective number of polymorphism; H_O: observed heterozygosity; H_E: expected heterozygosity; F: inbreeding coefficient

4. Conclusions

The high crossing rate and the high number of pollen donors identified in the present study could have a relevant role in maintaining the genetic diversity of native populations and in the progenies evaluated here. However, mating patterns are not static and may be influenced by several factors, biotic and abiotic. Especially in the case of macaúba palm, occurring in different biomes on the American continent, the rate of crossing or maintenance of genetic variability may vary between locations or populations, and should not present a definitive and or conclusive value.

In the present study, no significant differences were detected in the magnitude of genetic diversity between native populations and their respective open-pollinated progenies (OPP). The high levels of genetic diversity found evidence the initial stage of domestication of *A. aculeata* and suggest that a broad genetic basis would be available in the species for purposes of genetic improvement.

In addition, the maintenance of significant genetic diversity in progenies derived from few mothers selected from natural populations for the formation of commercial seedlings ensures the minimization of future risks of genetic vulnerability that could occur from the genetic bottleneck caused by the selection. However, in order to have a more conclusive idea of the dynamics of diversity during the breeding process of the species, it is necessary to analyze the diversity in the subsequent breeding selection cycles.

The results presented here allowed us to improve our understanding of the crossing rate and the dynamics of genetic diversity in *A. aculeata*. Likewise, they provided important information on the intensity of selection in natural and improved populations, for the advancement of generations, with the objective of maximizing future genetic gains without prejudice to the maintenance of genetic diversity.

References

Abreu, A.G., Priolli, R.H.G., Azevedo-Filho, J.A., Nucci, S.M., Zucchi, M.I., Coelho, R.M., Colombo, C.A., 2012. The genetic structure and mating system of *Acrocomia aculeata* (Arecaceae). Genet Mol Biol. 35, 116-121. https://doi.org/10.1590/S1415-47572012005000002

Alfaro-Solís, J.D., Montoya-Arroyo, A., Jiménez, V.M., Arnáez-Serrano, E., Pérez, J., Vetter, W., Frank, J., Lewandowski, I., 2020. *Acrocomia aculeata* fruits from three regions in Costa Rica: an

assessment of biometric parameters, oil content and oil fatty acid composition to evaluate industrial potential. Agrofor Syst., 1-15. https://doi.org/10.1007/s10457-020-00511-8

Amaral, F.P., 2007. Estudo das características físico-químicas dos óleos da amêndoa e polpa da macaúba [*Acrocomia aculeata* (jacq.) Lodd. Ex mart]. Universidade Estadual Paulista "Júlio de Mequita Filho".

Barrett, S.C., 2003. Mating strategies in flowering plants: the outcrossing–selfing paradigm and beyond. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 358, 991-1004. https://doi.org/10.1098/rstb.2003.1301

Bazzo, B.R., de Carvalho, L.M., Carazzolle, M.F., Pereira, G.A.G., Colombo, C.A., 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, 276. https://doi.org/10.1186/s12870-018-1509-9

Berton, L.H.C., de Azevedo Filho, J.A., Siqueira, W.J., Colombo, C.A., 2013. Seed germination and estimates of genetic parameters of promising macaw palm (*Acrocomia aculeata*) progenies for biofuel production. Ind Crops Prod. 51, 258-266. https://doi.org/10.1016/j.indcrop.2013.09.012

Billotte, N., Risterucci, A.-M., Barcelos, E., Noyer, J.-L., Amblard, P., Baurens, F.-C., 2001. Development, characterisation, and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. Genome 44, 413-425. https://doi.org/10.1139/g01-017

Bora, P.S., Rocha, R., 2004. Macaiba palm: fatty and amino acids composition of fruits macaíba: composición de aminoácidos y ácidos grasos de frutos macaíba: composición de aminoácidos e ácidos graxos de froitos. CYTA-J Food 4, 158-162. https://doi.org/10.1080/11358120409487755

Breed, M.F., Christmas, M.J., Lowe, A.J., 2014. Higher levels of multiple paternities increase seedling survival in the long-lived tree Eucalyptus gracilis. Plos one 9, e90478. https://doi.org/10.1371/journal.pone.0090478

Byerlee, D., Falcon, W.P., Naylor, R., 2017. The tropical oil crop revolution: food, feed, fuel, and forests. Oxford University Press. https://doi.org/10.1111/joac.12272

Charlesworth, D., Wright, S.I., 2001. Breeding systems and genome evolution. Curr Opin Genet Dev. 11, 685-690. https://doi.org/10.1016/S0959-437X(00)00254-9

Coelho, N.H.P., Tambarussi, E.V., Aguiar, B.I., Roque, R.H., Portela, R.M., Braga, R.C., Sanson, D., Silva, R.A., Ferraz, E.M., Moreno, M.A., 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. Conserv Genet. 19, 879-891. https://doi.org/10.1007/s10592-018-1061-z

Coimbra, M.C., Jorge, N., 2011. Proximate composition of guariroba (*Syagrus oleracea*), jerivá (*Syagrus romanzoffiana*) and macaúba (*Acrocomia aculeata*) palm fruits. Food Res Int. 44, 2139-2142. https://doi.org/10.1016/j.foodres.2011.03.032

Colombo, C.A., Berton, L.H.C., Diaz, B.G., Ferrari, R.A., 2018. Macauba: a promising tropical palm for the production of vegetable oil. OCL 25, D108. https://doi.org/10.1051/ocl/2017038

Díaz-Fuentes, V.H., Avendaño-Arrazate, C.H., Gálvez-Marroquín, L.A., Ruíz-Cruz, P.A., Iracheta-Donjuán, L., Nájera-Domínguez, W., Martínez-Valencia, B.B., 2019. Caracterizacion morfoagronomica de coyol (*Acrocomia aculeat*a Jacq.) para determinar su potencial productivo de aceite para biodiesel. AGROProductividad 12, 97-105.

Domiciano, G.P., Alves, A.A., Laviola, B.G., Conceição, L.D.H.C.S.d., 2015. Parâmetros genéticos e diversidade em progênies de Macaúba com base em características morfológicas e fisiológicas. Cienc Rural 45, 1599-1605. https://doi.org/10.1590/0103-8478cr20140909

dos Reis, E.F., Pinto, J.F.N., da Assunção, H.F., da Costa Netto, A.P., da Silva, D.F.P., 2019. Characteristics of 137 macaw palm (*Acrocomia aculeata*) fruit accessions from Goias, Brazil. Comun Sci. 10, 117-124. https://doi.org/10.14295/cs.v10i1.2389

Doyle, J.J., Doyle, J.L., 1990. Isolation of plant DNA from fresh tissue. Focus 12, 39-40.

Eguiarte, L.E., Perez-Nasser, N., Piñero, D., 1992. Genetic structure, outcrossing rate and heterosis in *Astrocaryum mexicanum* (tropical palm): implications for evolution and conservation. Heredity 69, 217. https://doi.org/10.1038/hdy.1992.119

Gaiotto, F.A., Grattapaglia, D., Vencovsky, 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

González-Pérez, M., Caujapé-Castells, J., Sosa, P., 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

Henderson, A., Galeano, G., Bernal, R., 1995. Field guide to the palms of the Americas. Princeton University Press.

Kalia, R.K., Rai, M.K., Kalia, S., Singh, R., Dhawan, A., 2011. Microsatellite markers: an overview of the recent progress in plants. Euphytica 177, 309-334. https://doi.org/10.1007/s10681-010-0286-9

Kantar, M.B., Betts, K., Michno, J.-M., Luby, J.J., Morrell, P.L., Hulke, B.S., Stupar, R.M., Wyse, D.L., 2014. Evaluating an interspecific *Helianthus annuus* × *Helianthus tuberosus* population for use in a perennial sunflower breeding program. Field Crops Res. 155, 254-264. https://doi.org/10.1016/j.fcr.2013.04.018

Kubota, S., Kameyama, Y., Hirao, A.S., Ohara, M., 2008. Adaptive significance of self-fertilization in a hermaphroditic perennial, *Trillium camschatcense* (Melanthiaceae). Am J Bot. 95, 482-489. https://doi.org/10.3732/ajb.95.4.482

Lanes, É.C., Motoike, S.Y., Kuki, K.N., Nick, C., Freitas, R.D., 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, É.C., Motoike, S.Y., Kuki, K.N., Resende, M.D., 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

Lefèvre, F., 2004. Human impacts on forest genetic resources in the temperate zone: an updated review. For Ecol Manag. 197, 257-271. https://doi.org/10.1016/j.foreco.2004.05.017

Liu, X., Tang, H., Li, D., Hou, L., 2011. Genetic diversity of coconut cultivars in China by microsatellite (SSR) markers. Mol Plant Breed. 2. https://doi.org/10.5376/mpb.2011.02.0012

Lorenzi, H.J., Noblick, L., Kahn, F., 2010. Flora brasileira:'Arecaceae'(palmeiras). Nova Odessa, SP: Instituto Plantarum.

Mason, A.S., 2015. SSR genotyping. In: Batley J. (eds) Plant Genotyping. Methods in Molecular Biology (Methods and Protocols), vol 1245. Humana Press, New York, NY. Pp. 77-79 https://doi.org/10.1007/978-1-4939-1966-6_6

Mengistu, F.G., 2015. Cross-species amplification of microsatellite markers and genetic diversity in the macaw palm (*Acrocomia aculeata)*. Universidade Federal de Viçosa

Mengistu, F.G., Motoike, S.Y., Caixeta, E.T., Cruz, C.D., Kuki, K.N., 2016. Cross-species amplification and characterization of new microsatellite markers for the macaw palm, *Acrocomia aculeata* (Arecaceae). Plant Genet Resour. 14, 163-172. https://doi.org/10.1017/S1479262115000179

Murawski, D., Hamrick, J., 1991. The effect of the density of flowering individuals on the mating systems of nine tropical tree species. Heredity 67, 167-174. https://doi.org/10.1038/hdy.1991.76

Murawski, D., Hamrick, J., 1992. Mating System and Phenology *Ceiba Pentadra* (Bombacaeae) in Central Panama. J Hered. 83, 401-404. https://doi.org/10.1093/oxfordjournals.jhered.a111241

Nazareno, A.G., Reis, M.S.d., 2012. Linking phenology to mating system: exploring the reproductive biology of the threatened palm species Butia eriospatha. J Hered. 103, 842-852. https://doi.org/10.1093/jhered/ess070

Nucci, S., Azevedo-Filho, J., Colombo, C., Priolli, R., Coelho, R., Mata, T., Zucchi, M., 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

Nucci, S.M., 2007. Desenvolvimento, caracterização e análise da utilidade de marcadores microssatélites em genética de população de macaúba. Master dissertation. Intituto Agrônomico, Campinas, SP.

OCDE/FAO, 2019. OCDE-FAO Perspectivas Agrícolas 2019-2028, OECD Publishing, París/Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), Roma https://doi.org/10.1787/7b2e8ba3-es

Oliveira, D., Melo Júnior, A., Brandão, M., Rodrigues, L., Menezes, E., Ferreira, P., 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

Park, S., 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. University of Dublin. Ph. D. Thesis. [thesis], pp 254.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population software for teaching and research. Mol Ecol Notes 6, 288-295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Picanço-Rodrigues, D., Astolfi-Filho, S., Lemes, M.R., Gribel, R., Sebbenn, A.M., Clement, C.R., 2015. Conservation implications of the mating system of the Pampa Hermosa landrace of peach palm analyzed with microsatellite markers. Genet Mol Biol. 38, 59-66. https://doi.org/10.1590/S1415-475738120140022

Pugesgaard, S., Schelde, K., Larsen, S.U., Lærke, P.E., Jørgensen, U., 2015. Comparing annual and perennial crops for bioenergy production–influence on nitrate leaching and energy balance. Gcb Bioenergy 7, 1136-1149. https://doi.org/10.1111/gcbb.12215

Ritland, K., 1988. The genetic-mating structure of subdivided populations II. Correlated mating models. Theor Popul Biol. 34, 320-346. https://doi.org/10.1016/0040-5809(88)90027-5

Ritland, K., 2002. Extensions of models for the estimation of mating systems using n independent loci. Heredity 88, 221-228. https://doi.org/10.1038/sj.hdy.6800029

Ritland, K., El-Kassaby, Y., 1985. The nature of inbreeding in a seed orchard of Douglas fir as shown by an efficient multilocus model. Theor Appl Genet 71, 375-384. https://doi.org/10.1007/BF00251176

Ritland, K., Jain, S., 1981. A model for the estimation of outcrossing rate and gene frequencies using n independent loci. Heredity 47, 35-52. https://doi.org/10.1038/hdy.1981.57

Scariot, A., Lleras, E., Hay, J.D., 1995. Flowering and fruiting phenologies of the palm *Acrocomia aculeata*: patterns and consequences. Biotropica, 168-173. https://doi.org/10.2307/2388992

Sebbenn, A., 2006. Sistema de reprodução em espécies arbóreas tropicais e suas implicações para a seleção de árvores matrizes para reflorestamentos ambientais. In: Higa A., R., Sulva, L., D. (Eds.) Pomares de sementes de espécies florestais nativas. Curitiba: FUPEF, 193-198.

Simiqueli, G.F., de Resende, M.D.V., Motoike, S.Y., Henriques, E., 2018. Inbreeding depression as a cause of fruit abortion in structured populations of macaw palm (*Acrocomia aculeata*): Implications for breeding programs. Ind Crops Prod. 112, 652-659. https://doi.org/10.1016/j.indcrop.2017.12.068

Souza, L., Kageyama, P.Y., Sebbenn, A.M., 2003. Sistema de reprodução em população natural de *Chorisia speciosa* A. St.-Hil.(Bombacaceae). Rev Bras Bot 26, 113-121. https://doi.org/10.1590/S0100-84042003000100012

USDA, 2020. United States Department of Agriculture, www.fas.usda.gov

Vinson, C. C., Mangaravite, E., Sebbenn, A., Lander, T., 2018. Using molecular markers to investigate genetic diversity, mating system and gene flow of Neotropical trees. Braz J Bot. 41, 481-496. https://doi.org/10.1007/s40415-018-0472-x

Whitehead, M.R., Lanfear, R., Mitchell, R.J., Karron, J.D., 2018. Plant mating systems often vary widely among populations. Front Ecol Evol. 6, 38. https://doi.org/10.3389/fevo.2018.00038

Wright, S., 1940. Breeding structure of populations in relation to speciation. The American Naturalist 74, 232-248. https://www.jstor.org/stable/2457575

Zehdi, S., Trifi, M., Billotte, N., Marrakchi, M., Christophe Pintaud, J., 2004. Genetic diversity of Tunisian date palms (Phoenixdactylifera L.) revealed by nuclear microsatellite polymprohism. Hereditas 141, 278-287. https://doi.org/10.1111/j.1601-5223.2004.01855.x

CAPÍTULO IV - Macauba: the tropical palm that promises to impact the world oil production

Carlos Augusto Colombo, Luiz Henrique Chorfi Berton, Brenda Gabriela Diaz, Roseli Aparecida Ferrari

Revisão publicada na revista Oilseeds & fats Crops and Lipids

Macauba: the tropical palm that promises to impact the world oil production.

Carlos Augusto Colombo[1*], Luiz Henrique Chorfi Berton[1], Brenda Gabriela Diaz[1], Roseli Aparecida Ferrari [2,]

[1] Genetic Resources Center, Agronomic Institute of São Paulo [IAC], Campinas, SP Brazil

[2] Food Chemistry and Applied Nutrition Center, Institute of Food Technology [ITAL], Campinas, SP Brazil

*Corresponding Author: ccolombo@iac.sp.gov.br

Abstract – The growing global demand for vegetable oils for food and for replacing fossil fuels leads to increased oilseeds production. Almost 122 of the current 187 million tons of vegetable oils produced in the world correspond to palm and soybean oils. The oil palm is cultivated in the tropical zone, in areas formerly occupied by forests, and soybean oil is a by-product of protein meal production. The diversification of raw materials for the vegetable oil market is thus strategic for both food and non-food sectors. Sources for vegetable oil should be economically competitive and provide sustainability indexes higher than that provided by oil palm and soybean. In this context, we describe the potential of Acrocomia aculeata, popularly known as macauba. Macauba is an American palm from the tropical zones which presents oil productivity and quality similar to that of the oil palm. It grows spontaneously in a wide range of environments and it is not very water demanding. Macauba palm has a high potential for oil production and for diversification of co-products with some potential of value aggregation. Such a perennial and sustainable species will probably fulfill the requirements to become an important new commercial oilseed crop

Keywords: Acrocomia aculeata / macauba palm / palm tree / biodiesel / sustainability

Résumé – Macauba : un palmier tropical prometteur pour la production d'huile végétale. La demande mondiale croissante d'huiles végétales pour l'alimentation et pour remplacer les combustibles fossils entraîne une production accrue d'oléagineux. Près de 122 des 187 millions de tonnes d'huiles végétales actuellement produites dans le monde sont issues de palme ou de soja. Le palmier à huile est cultivé en zone tropicale, sur des surfaces anciennement occupées par des forêts; l'huile de soja est un sous-produit de la production de protéines. La diversification des matières premières pour ce marché est donc stratégique pour les secteurs alimentaire et non alimentaire. Les

sources d'huiles végétales devraient être économiquement compétitives et afficher des indices de durabilité plus élevés que le palmier à huile et le soja. Dans ce contexte, nous décrivons le potentiel d'Acrocomia aculeata, connu sous le nom de macauba. Ce palmier américain provenant des zones tropicales présente une productivité et une qualité d'huile similaires à celles du palmier à huile. Il se développe spontanément dans un large éventail d'environnements et n'est pas très exigeant en eau. Le palmier macauba possède un potentiel de rendement élevé en huile et de diversification des coproduits à forte ajoutée. Une telle espèce pérenne et durable répondra probablement aux exigences pour devenir une nouvelle variété commerciale importante.

Mots clés: Acrocomia aculeata / palmier macauba / palmier / biodiesel / durabilité

1 Introduction

Global production of vegetable oil is growing and estimated in 187 million tons for the 2016/2017 crop, according to the Foreign Agricultural Service (USDA, 2017), with a forecast of 195 million due to rising food consumption in emerging countries and in biofuels application. Of these, 37.6% (70.3 million tons) are represented by palm and palm kernel oil and 30% (55 million tons) by soybean oil. The remaining 32.5% are supplied by canola, sunflower, peanut and cottonseed, mainly (Fig. 1).

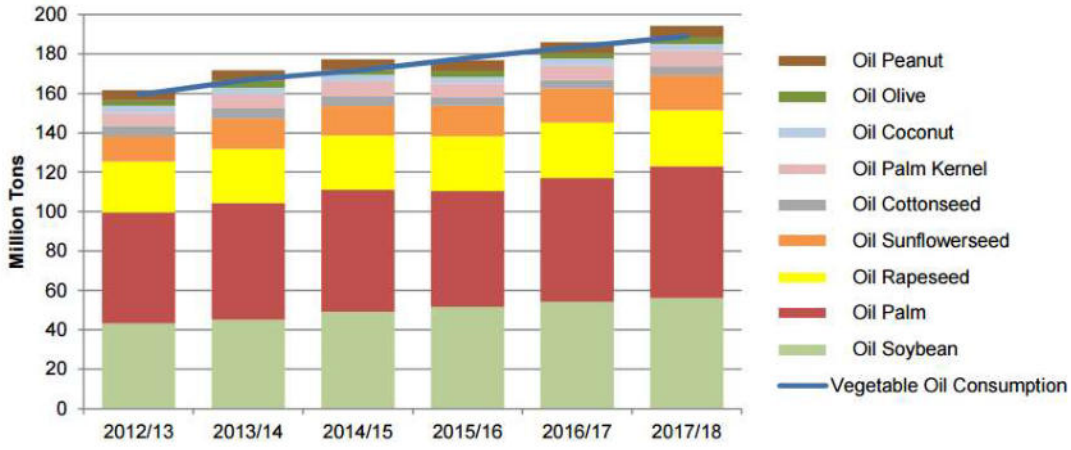


Fig. 1. Global vegetable oil production and consumption (Source: USDA, 2017)

Palm oil represents almost 40% of the world's vegetable oil production, with 70.3 million tons of oil produced in Indonesia and Malaysia accounting for 86.5% of that total (USDA, 2014). The share of the American continent in the world vegetable oil production is around 25%, soybean is the main raw material, with 58% of this value, followed by canola (8%) and palm (5%), according to Carrasco (2013).

The species *Elaeis guineensis* is of tropical origin and cultivated on a large scale only in tropical countries with an adequate water balance, with an annual rainfall above 1,500 mm, with monthly distribution between 120 mm and 150 mm and absence of dry season (Bastos, 2001). Brazil contributes with only 0.54% of palm oil produced in the world, the northern region of the country concentrating 80% of the national production, which is 340 thousand tons. The average Brazilian productivity is 2,000 kilos of oil per hectare against the 4,000 produced in Indonesia and Malaysia, considering the production and occupied area (USDA, 2014).

The USDA forecast (2017) is that world vegetable oil consumption is expected to grow, reflecting population growth and GDP growth in emerging economies. It is also expected to increase the consumption of palm oil in the world by the abandonment of saturated fats in the developed world (Carrasco, 2013) or for the biodiesel production which, according to Byerlee *et al.* (2016), represents almost half of the increase in vegetable oil consumption. Considering the total of palm oil produced in the world, 80% is used by the food sector, 10% in oil chemistry and 10% in bioenergy and biofuels (Andrade, 2015). In the case of biodiesel, the United States, Argentina, China, Brazil, Germany and France are important examples of countries that have adopted the inclusion of biofuel in their economies (Mattei, 2017; Souza *et al.*, 2016). In the specific case of Brazil, biodiesel policy was implemented in 2004 with law 11,097, and the last resolution establishes the proportion of 8% biodiesel to diesel oil sold to the consumer. This measure has increased interest in the diversification of alternative sources of raw materials for oil production, especially of plant species, and species such as jatropha, castor bean and crambe have been the subject of scientific studies aimed at the viability or dynamization of their respective production chains to meet the growing demand for renewable biofuel.

Soybean oil accounted for 78% of all biodiesel manufactured in Brazil, followed by animal fats (18%), cotton oil (1%) and used frying oil (1%), on average between January and May of this year. Approximately 1.5 million tons of soybean oil was destined for the production of biodiesel, according to Abiove (2017). According to data provided by the National Agency of Petroleum,

Natural Gas and Biofuels (ANP), in 2016, the nominal capacity for biodiesel production (B100) in Brazil was around 7.3 million m³, approximately 609 thousand m³/ month. However, domestic production was close to 3.4 million m³, which corresponded to 44.89% of total capacity, as reported by the Ministry of Mines and Energy of the Federal Government (MME, 2017).

Whether for food or industrial use, many factors should motivate the choice of raw material, such as local availability, economic viability, storage conditions, physico-chemical properties of the oil and performance as biofuel. The production of raw material also varies according to the region due to soil and climate characteristics, among others (Haas and Foglia, 2006). In relation to soybeans, homogeneity and great production is due to the degree of agricultural technology involved in its production. However, soybean has low oil yield (400 kg per hectare) and does not favor regional development, since it is concentrated in the south and center-west regions and presents low social insertion and, consequently, few additional occupations with biodiesel (Sauer, 2006; IPEA, 2012).

The economic feasibility of production of jatropha, sunflower, crambe, canola, babassu, among other oilseeds, also depends on research and technological advances. Palm oil has been one of the largest bets, but its planting should be restricted to the north region of Brazil, due to its dependence on the water regime. Given this scenario, ~~we~~ it is presented a new tropical species for the exploration and production of vegetable oil ~~to~~ for diverse use and which has been arousing the interest of the Brazilian scientific community, proving to be spectacularly competitive. It is the palm tree *Acrocomia aculeata*, popularly called as macauba palm that will be presented below.

2 Botanic aspects

The macauba palm (*Acrocomia aculeata*, Arecacea) is a perennial, heliophite with a height varying from 4 to 15m in height, glabrous, fusiform cylindrical stipe, densely acular and ringed. It has 20 to 40 leaves agglomerated at the apex of the stipe, pinnate composites, 4 to 5 m in length, petiolate; alternate leaflets unevenly distributed along the rachis, which may contain numerous spines. Interfoliar and branched inflorescences, long rachis with several branches of equal size, multiflorous. Pistilated flowers, always forming triads and flowers staminated in great number in the apex. The specie is monoic, protogenic and with annual seasonal flowering. In Brazil flowering from september to february, with peak flowering in november and december (Scariot *et al.*, 1995; Lorenzi, 2006, Berton, 2013). Fruiting occurs throughout the year and fruits ripen 12 to 13 months after fertilization (Scariot *et al.*, 1995; Montoya *et al.*, 2015). In the region of Veracruz, Mexico, flowering occurs from

March to September (Quero, 1994) and in Cuba fruits mature between march and may (Pérez *et al.,* (2015).

The fruit is an edible drupe with 3.0 to 5.0 cm diameter, globose, epitaph cart. with varied coloration; mucilaginous fibrous mesocarp of varied staining and sweet taste; endocarp strongly adhered to the mesocarp; seed with a large amount of endosperm, and up to four seeds per fruit (Henderson *et al*., 1995; Lorenzi *et al*., 2010; Berton, 2013) (Fig. 2). When mature, the fruit emits a characteristic aroma and the hull easily detaches from the pulp (Almeida *et al*., 1998; Lorenzi, 2006). The fruits consist of approximately 20% of hull, 40% of pulp, 33% of endocarp and 7% of kernel (Berton, 2013, Ciconini *et al.*, 2013).

Fig. 2. A- Adult plant of *Acrocomia aculeata*; B- Bunch of fruits; C- Detail of the fruit; D- Open inflorescence; E- Detail of the inflorescence, with female flowers in the basal region and male flowers in the apical region.

3 Domestication, occurrence and popular use

Acrocomia aculeata grows in the dry areas of the New World, from Mexico and the Caribbean Islands to northern Argentina (Morcote-Rios and Bernal, 2001). The species occurs in higher densities in open areas, associated to pasture areas, also occurring in semideciduous forests and in locations with rocky outcrops, and can be found throughout tropical and subtropical America between latitudes 22ºN (Mexico) to 28ºS (Argentina). According to Kahn and Moussa (1997), *Acrocomia aculeata* is a peri-Amazonian species, introduced in the Amazon. This palm tree is present from the Atlantic coast of Colombia, Venezuela, Guyana and the Brazilian states of Amapá and Pará, and on the southern outskirts of the central basin of Brazil, Bolivia and Paraguay. It is not found in primary forests but extends in deforested areas of the tropical forest (Henderson *et al.*, 1995). Hernández *et al.* (2013) report that the species tolerates high and low temperatures and droughts, not suffering from diseases, as compared to other crops, according to observations made in the municipality of San Blas, Nayarit, on the Pacific coast of Mexico.

Acrocomia aculeata is the most notable palm tree. Archaeological data suggest that the distribution and abundance of the most productive and more generalist palm trees may have resulted from human use (Morcote-Rios and Bernal, 2001). These authors cite the presence of remains of this in 29 archaeological sites, from Mexico to Brazil. The oldest sites for the species are found in Santarém, northern Brazil (11,200 B.P.), Colombia (9530 B.P.), Panama (8040 B.P.) and Mexico (6750 B.P.). Lentz (1990), evaluating Mesoamerican archaeological sites, reported remains of macauba fruits in Belize: Cerros (200 BC), Colha (900 AD); Honduras: Cerro Palenque (600 AD), Copán Valley (400 AD), Cajón (200 BC); Mexico: Tehuacan Valley (4800 BC); and Panama: Aguadulce (5000 BC) and Chiriqui (4600 BC). The dispersion of macauba from South America to Central America seems more likely. This is not only because of the decline in the oldest dates of archaeological data going north, but also because the genus itself is probably of South American origin, as suggested by the fact that there is another known species, *Acrocomia hassleri*, ~~to be~~ restricted to cerrado areas in Brazil (Henderson *et al.*, 1995).

Macauba palm has been, incipiently domesticated by Mesoamerican peoples for millennia and the beginning of its domestication must be associated with the use of the plant by the pre-Columbian peoples of tropical America, and there are strong reasons for the use of this palm, according to Lévi-Strauss (1952). First, the fruits of *A. aculeata*, unlike other oil palm trees, have an abundant mesocarp that can be consumed directly, and the abundant oil can be easily extracted.

Secondly, the thin thickness of the brittle epicarp provides good protection for the mesocarp and can be easily removed when necessary. Thirdly, the fruits do not ferment rapidly, thus allowing their consumption over a period of several weeks, suggesting a good choice for feeding during the migration of people. They would have fresh fruits to ~~first~~ eat first the fleshy and oily mesocarp during the trip and then the seed or discarding it intact, since the breaking of the endocarp was not an easy option.

There are many examples of the popular use of macauba palm, such as food use by the Mayans (McNeil *et al.*, 2010) and for feeding and use of oil for illumination by the Mbayá-guaicurus people, in the Paraguay river basin, on the second half of the seventeenth century (Carvalho, 2006). As examples of contemporary use, kernel fat is used in French Antilles as a softener in joint diseases, and seed drink applied against internal inflammation (Roig and Mesa, 1945). Pulp oil is used to treat headaches and nevralgias (Corrêa, 1984) and as laxative (Berg, 1984) and fruit pulp used to treat catarrhal conditions, being purgative if taken fasting (Almeida *et al.*, 1998). It is also used in the feeding of different animals in their regions of occurrence (Cavalcante, 1991; Cruz *et al.*, 1984) and the oil extracted from the pulp is applied in the industry of savory, just as the natives of Amazonas anoint the body ~~so as~~ in order to defend against the attack of mosquitos (Pio Corrêa, 1931). It is also used in ~~In~~ cooking, for soft drinks, sweets and jellies (Cavalcante, 1991; Silva *et al.*, 2001) and for the manufacture of ice cream (Conceição and Paula, 1986).

4 Fruit properties

The main products of macauba palm are the oils derived from fruit pulp and kernel. The pulp presents up to 75% of total lipid content and the kernel can contain up to 65%, both on dry basis (Berton, 2013, Ciconini et al., 2013). The oil of the pulp presents minor compounds in its composition, such as tocopherols, phytosterols, β-carotene, flavonoids and vitamin C (Trentini *et al.,* 2016 Rocha *et al.*, 2013, Coimbra and Jorge, 2012). Among the most important fatty acids are oleic acid, with 70% content (Navarro-Diaz *et al*., 2014), followed by palmitic and linoleic acid (Hiane *et al*., 2005; Berton, 2013, Lescano *et al.*, 2015) .

Mesocarp oil is rich in oleic acid, which gives the title "high oleic", with high value and of great demand in the food industry and, especially energy, due to the greater oxidative stability and operability at low temperatures (Berton, 2013, Silva *et al.,* 1986). Some studies have pointed out that the development of acidity in macaúba pulp oil is much slower than for palm oil. Macauba fruits

harvested directly from the bunch or naturally fallen, but without contact with the soil, can remain stored in environmental conditions for longer periods, up to 16 days without exceeding 5% acidity (Souza, 2013; Evaristo *et al.,* 2016a, b) or up to 180 days if in hygienic or thermal treatment (Cavalcanti-Oliveira *et al.,* 2015).

The oil extracted from the endosperm is rich in saturated fatty acids of short chain, constituting a valuable raw material for pharmaceutical and cosmetic use (Beltrão and Oliveira, 2007). It is possible to obtain 50% of the fatty acids present in the lipid extract of the seed (Belén-Camacho et al., 2005).

Pulp oil is light yellow to dark orange due to the presence of carotenoids. From the nutritional point of view, the ß-carotene content found by Ramos *et al.* (2008) in humid pulp was 49.0 mg.g-1, corresponding to about 80% of the total carotenoids found in the pulp. The chromatogram of the carotenoid extract of the pulp shows peaks of 1-zeaxanthin, 4-transcyclopen, 5-α-cryptoxanthin or zeinoxanthin, 6-cis-lycopene, 7 and 8-γ-carotene, 9-trans b-carotene and 10- 13-cis-b-carotene.

The fruit pulp has 4 to 8% of protein and 10 to 57% of fiber. In the kernel can be founded 22 to 50% of proteins and 34 to 64% of fibers (Fig. 3). In addition, considerable amounts of minerals are found in both parts of the fruit, such as calcium and magnesium macronutrients, and micronutrients of copper, manganese, iron and zinc (Bora and Rocha, 2004; Hernández *et al.* Trentini *et al.*, 2016). Ramos *et al.* (2008) carried out comparative analyzes of the mineral content of macauba fruits in relation to tropical fruits commercialized and consumed by the population (avocado, pineapple, banana, papaya, passion fruit, melon and tangerine). According to their report, calcium and potassium contents were higher, the potassium content was twice that found by Franco (2004) in banana (333.4 mg.100 g^{-1}) and in passion fruit (380.0 mg.100 g-1). The mineral composition of the fruit pulp presents high concentrations (mg 100 g -1) of K (1,725 ± 80), Ca (680 ± 69), Mn (20.0 ± 2.0), Fe (101.0 ± 14.0), Cu 3.0), Zn (15.0 ± 2.0) and Br (8.0 ± 1.0), according to a study conducted by Oliveira *et al.* (2006). Bora and Rocha (2004) observed 07 essential amino acids and 10 non-essential amino acids in the mesocarp and endosperm of macauba fruits. According to Cruz *et al.*, (1984), the macauba endosperm residue presents high levels of aspartic and glutamic acid, as well as excellent threonine/serine balance.

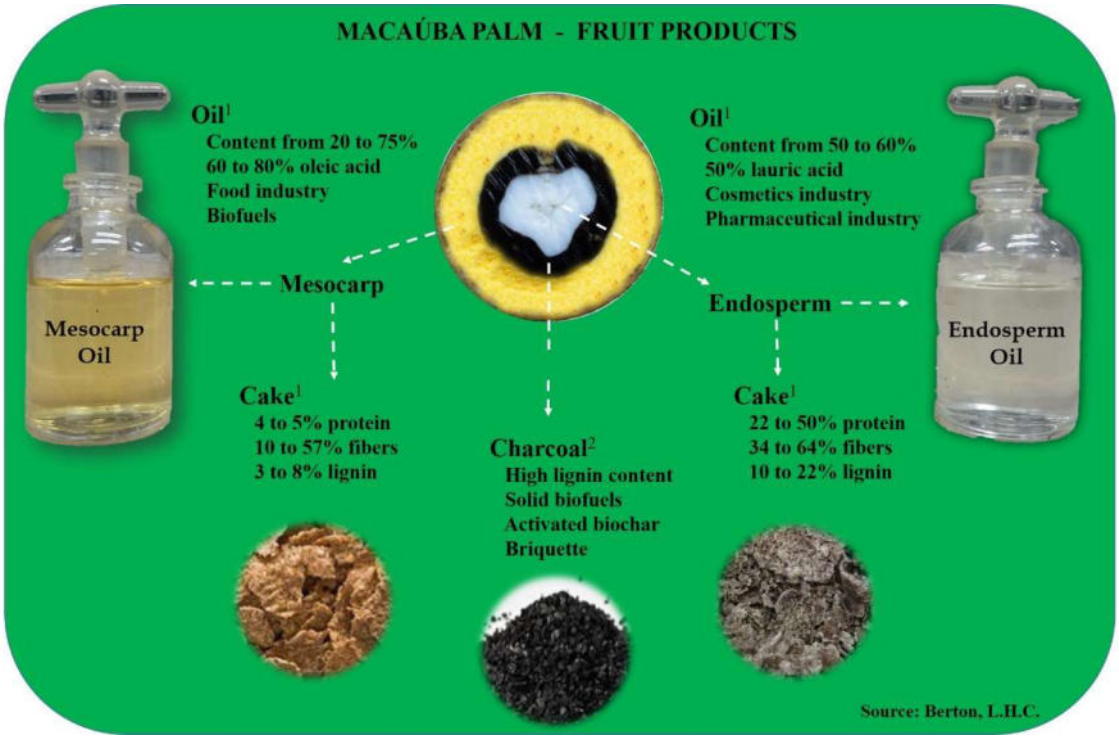


Fig. 3. Fractions and products of macauba fruit. Source: 1) Bora and Rocha, 2004; Beltrão and Oliveira, 2007; Ramos *et al.,* 2008; Berton, 2013; Hernández *et al.,* 2013; Trentini *et al.*, 2016. 2) Silva *et al.,* 1986; Vilas-Boas *et al.*, 2010; Silva and Caño Andrade, 2013.

From each kilogram of fruit subjected to cold oil extraction (hydraulic), is generated 500 g of residual biomass with high value of fibers, lipids and proteins that can be used for alternative feeding of animals. The use of macauba presscakes, for ruminant and non-ruminant feeding, has been successfully performed, according to studies by Costa Junior *et al.* (2015), for pigs, Santos *et al.* (2017) for sheep, Rufino *et al.* (2011) for goats, and Azevedo *et al.,* (2013) for cows.

In order to produce charcoal from this co-product, Vilas-Boas *et al.* (2010) showed that the carbon of the endocarp has an apparent density of 1.29 (g/ cm^2) and a calorific value of 8045.56 (kcal/kg) due to the high concentration of lignin in the endocarp. According to Silva *et al.,* (1986), the production of charcoal from the endocarp at the final temperature of 550ºC, whose product is superior to the eucalyptus charcoal. It is possible to be used in steelmaking, metallurgical operations and handicrafts, or for the production of charcoal with low ash, without sulfur and high density (Silva and Caño Andrade, 2013) or activated charcoal (Silva *et al.,* 1986).

According to Berton (2013) and Evaristo et al., (2016) (Fig. 4) considering the average proportions of fractions of the macauba fruit with 22.6% hull; 42.6% pulp; 28.4% endocarp and 6.4% kernel, the estimation of productivity in Brazil in conservative scenario and 400 trees per ha.year^{-1} would be of 2,500 kg oil.ha.year^{-1}. However, in optimistic scenario the production would be of 5,000 kg oil.ha.year^{-1} of pulp oil. This value represents the great potential of macauba palm, much higher than that palm oil today without plant breeding i.e. only taking in account the cultivation of seeds from wild plants.

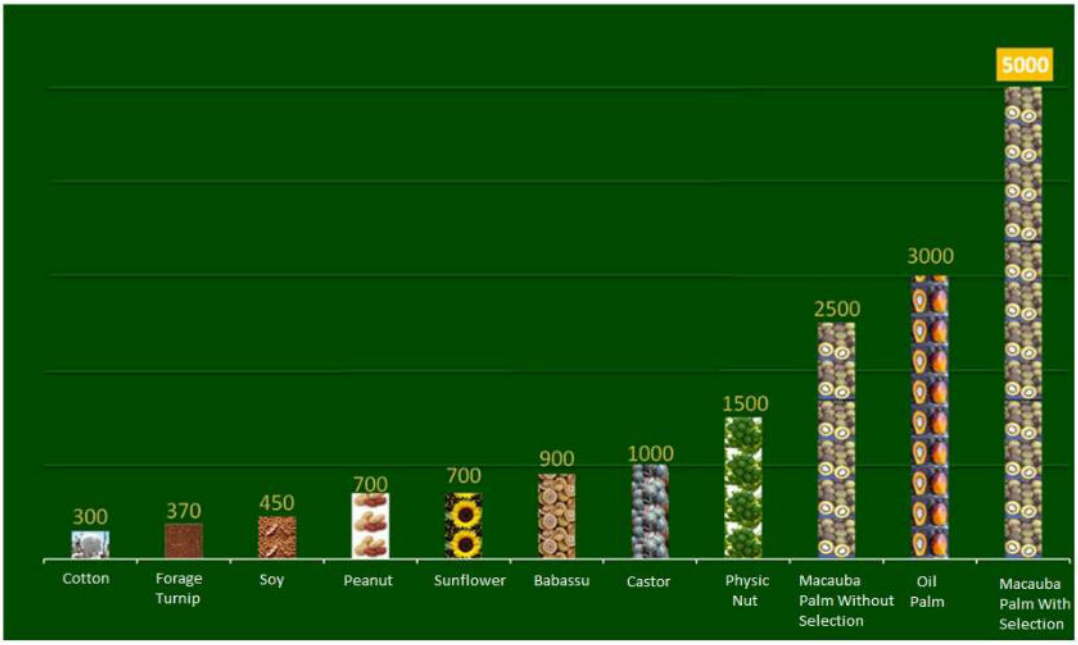


Fig. 4. Productivity of macauba in Brazil compared to other oilseeds in a conservative and optimistic scenario (400 plants. Ha^{-1}).

5 Conclusions

The world production and consumption of vegetable oil has been increasing every year accumulating almost 20% increase only in the last five years. Palm is the species that contributes most to the current 187 million tons of oil produced, with almost 38% of this total. This is the most productive species, reaching values of the order of 6 to 8 thousand kilos of oil per hectare. However,

the expansion of its production area is limited by its ecological limits, particularly in relation to its water requirements, in terms both of volume and distribution.

A consequence of the growing demand for vegetable oils is the diversification of more sustainable raw materials with greater social insertion, since both palm and soybeans, the second most important source of oil, have a low level of environmental sustainability.

In this context, the diversification of raw materials goes through the search for oleaginous plants that offer economic and sustainability guarantees. And these are the main arguments that have led hundreds of Brazilian scientists to develop research on biological and technological aspects of the macauba palm. The results obtained so far are quite encouraging. The productive capacity revealed by the commercial plantations indicates a productivity of 5,000 kilos of oil of the pulp and 1500 kilos of oil of the almond per hectare. Both the oil of the pulp and the oil of the almond are very similar to the one of the oil palm in terms of composition and both present a spectrum of industrial application quite diversified, in the alimentary area as of cosmetics and of energy. The oil extraction residues of both fractions of the fruit, mesocarp and endosperm, do not present anti-nutritional or toxicity factors, such as castor bean or jatropha pies, and the results of the application of both for ruminant and Ruminants are equally encouraging.

There is no evidence of the presence of pests or diseases in the natural populations visited by several researchers until now, nor in the commercial plantations already in production phase, mainly in the central region of Brazil, reinforcing the idea that this crop should present less phytosanitary problems than the jatropha has been presenting in its cultivated areas.

Similarly, seed germination, which initially appeared to be an obstacle to the production of seedlings due to seed dormancy, is now technically feasible.

Another important credential for the emergence of a large-scale production chain for macauba is related to its wide occurrence in the American continent, including subtropical areas with precipitation levels below 1000 mm per year and with seasonal water deficits. It is considered the most occurring palm tree in Brazil, being reported in practically all the Brazilian states, with absence only in the northeast and south coast of the country, occurring in greater abundance in cerrado regions, typically characterized by acidic, dystrophic and seasonal climate, with dry season of 3 to 5 months. The wide distribution of macauba palm can be verified also by the occurrence in diverse areas such as margins of highways and streams, in degraded areas, in consortium with annual crops and in open areas with pastures.

The macauba palm consortium option with pasture areas has been consolidating in Brazil as an important alternative for the aggregation of income to small and medium-sized ranchers. Besides the economic question, tropical and subtropical Brazilian soils with pastures present high rates of degradation. It is estimated that 80% of the 50-60 million hectares of pasture grown in central Brazil are in some state of degradation and inability to sustain the levels of production and quality demanded by the animals. One of the management options for remediation of these soils would be the intercropping with perennial species, which would introduce advantages such as carbon storage in the plant, ecosystem enrichment and maintenance of CO2 in the soil and the decrease of water flow from the fertile layer. In the case of macauba palm, in addition to all these advantages, it would also contribute its various co-products.

In view of the above, we believe that we are facing the emergence of a culture that should cause a significant and positive economic and environmental impact in the countries that adopt it for cultivation. In the case of Brazil, a country with a long tradition of agriculture, mainly of commodities, with a large arable area and a high-performance agricultural technology, macauba palm must find all the necessary conditions for the birth of a new oilseed crop for the tropics.

References

ABIOVE, Associação Brasileira da Indústrias de Óleos Vegetais. www.abiove.org.br., accessed may 20 th 2017.

Almeida SP, Proença CEB, Sano SM, Riveiro JF. 1998. Cerrado – Espécies vegetais úteis. Brasília: Embrapa Cerrados, 464p.

Andrade E. 2015. A cadeia produtiva da palma de óleo no Estado do Pará: uma avaliação crítica. Audiência Pública sobre o Programa Nacional de Produção Sustentável de óleo de alma. Brasília.

Azevedo RA, Dos Santos ACR, Ribeiro Jr CS, Bicalho FL, Bahiense RN, Araújo L, Geraseev LC. 2013. Ingestive behavior of dairy cows fed with macauba meal. Ciência Rural 43:1485-1488.

Bastos T, Muller A, Pacheco N, Neiva SM, Delgado AE. 2001. Climatic risk zoning for oil palm in Pará state. Rev. Bras. Agro. 9: 564-570.

Belén-Camacho DR, López I, García D, González M, Moreno-Álvarez MJ, Medina C. 2005. Evaluación fisico-química de la semilla y del aceite de corozo (*Acrocomia aculeata* Jacq.). Grasas Aceites. 56: 311-316.

Beltrão NEM, Oliveira MI. 2007. Oleaginosas Potenciais do Nordeste para a produção de Biodiesel. Documentos 177. Campina Grande – PB. EMBRAPA. 54p.

Berton, L.H.C. 2013. Avaliação de populações naturais, estimativas de parâmetros genéticos e seleção de genótipos elite de macaúba (*Acrocomia aculeata*). (Tese de Doutorado). Instituto Agronômico de Campinas – IAC. Campinas, SP, Brasil.

Bora PS, Rocha RVM. 2004. Macauba palm: fatty and amino acids composition of fruits. Ciência Tecnologia Alimentos 4: 158-162.

Byerlee D, Falcon WP, Naylor RL. 2016. The tropical oil crop revolution: food, feed, fuel and forests. Oxford University Press, 114p.

Carrasco, R.A.T. 2013. Evolution and Perspectives for Oil Palm in America. Palmas. Vol. 34 No. Especial, Tomo II, 2013.

Carvalho FAL. 2006. Etnogênese Mbayá-Guaykuru: Notas sobre emergência identitária, expansão territorial e resistência de um grupo étinico no vale do Rio Paraguai (*c.* 1650-1800). Rev. Hist. Est. Cult. 3: 1-20.

Cavalcante PB. 1991. Frutas comestíveis da Amazônia. 5.ed. Belém: CETUP/ CNPq/ Museu Paraense Emílio Goeldi, 279 p.

Cavalcanti-Oliveira ED, Silva PR, Rosa TS, Moura NML, Santos BCP, Carvalho DB, Sousa JS, Carvalhinho MTJE, Castro AM, Freire DMG, 2015. Methods to prevent acidification of Macauba (*Acrocomia aculeata*) fruit pulp oil: a promising oil for producing biodiesel. Ind. Crops Prod. 77: 703-707.

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 Crop Prod. 45:208-214.

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.

Conceição CA, Paula JE. 1986. Contribuição para o conhecimento da flora do Pantanal Mato-Grossense e sua relação com a fauna e o homem. In: Simpósio sobre recursos sócio econômicos e culturais do pantanal 1, 1984, Corumbá, MS. Anais... Brasília: EMBRAPA–DDT, 265p.

Corrêa MP. 1984. Dicionário das plantas úteis do Brasil e das exóticas cultivadas. Rio de Janeiro: IBDF. v.6.

Cruz PEN, Marques E.P, Amaya DR, Fáran JA. 1984. Macaúba, bacuri, inajá e tucumã. Caracterização química e nutricional destes frutos do Estado do Maranhão e os óleos respectivos. Rev. Quím. Ind. 278-281.

Evaristo AB, Grossi JAS, Carneiro ADO, Pimentel LD, Motoike SY, Kuki KN. 2016. Actual and putative potentials of macauba palm as feedstock for solid biofuel production from residues. Biomass Bioenerg. 8: 18-24,

Evaristo AB, Martino DC, Donato DB, Ferraez AH, Carneiro ACO, Grossi JAS. 2016. Energy potential of the macaw palm fruit residues and their use in charcoal production. Ci. Fl. 26: 571-577.

Hass MJ, Foglia TA. 2006. Alternative raw materials and technologies for the production of biodiesel. In: Knothe, G. et al. (Eds.) Manual do biodiesel. São Paulo: Edgard Blücher, p. 46-66.

Henderson A, Galeano G, Bernal R. 1995. Field guide to the palms of the Americas. Princeton University Press, Princepton, New Jersey. 352 pp.

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.

IPEA- Instituto de Pesquisa Econômica Aplicada. Biodiesel no Brasil: desafios das políticas públicas para a dinamização da produção. 2012.Comunicados n.º 137.

Kahn F, Moussa F. 1997. El papel de los grupos humanos en la distribución geográfica de algunas palmas en la Amazonía y su periferia. *In: Usos y recursos vegetales en el Ecuador* (Rios, M. & Pedersen, H.B., Eds.): 83-99.

Lescano CH, Oliveira IP, Silva LR, Baldivia DS, Sanjinez EJ, Argandoña, Arruda EJ, Moraes IC, Lima FF. 2015. Nutrients contente, characterization and oil extraction from *Acrocomia aculeata* (Jacq.) Lodd. Fruits. Afr. J. Food Sci. 9: 113-119.

Lévi-Strauss C. 1952. The use of wild plants in tropical South America. Econ. Bot. 6: 252-270.

Lorenzi GMAC. 2006. *Acrocomia aculeata* (Jacq.) Lodd. ex Mart. – Arecaceae: bases para o extrativismo sustentável. Tese de Doutorado. Universidade Federal do Paraná, Curitiba, 156f.

Lorenzi H, Noblick LR, Kahn F, Ferreira E. 2010. Flora brasileira: Arecaceae (Palmeiras). Instituto Plantarum, Nova Odessa. 382p.

Mattei L. 2017. Programa nacional para produção e uso do biodiesel no Brasil (PNPB): trajetória, situação atual e desafios. Rev. Econ. Nord. 41: 731-740.

Mcneil CL, Bumey DA, Burney LP. 2010. Evidence disputing deforestation as the cause for the colapse of the ancient Maya polity of Copan, Honduras. PNAS. 107: 1017-1022.

MME, Ministério de Minas e Energia. 2017. Boletim dos Biocombustíveis, edição nº 107 Jan/Fev.

Montoya SG, Motoike SY, Kuki K, Oliveira CM, Honório IG. 2015. Recording and damage characterization of coleopters in macao palm. Pesq. Flor. Bras. 35: 159-162.

Morcote-Rios G, Bernal R. 2001. Remains of palms (Palmae) at arqueological sites in the New World: A review. The Bot. Rev. v.67.

Navarro-Díaz HJ, Gonzalez SL, Irigaray B, Vieitez I, Jachmanián I, Hense H, Oliveira JV. 2014. Macauba oil as an alternative feedstock for biodiesel: Characterization and ester conversion by the supercritical method. Journal of Supercritical Fluids. 93: 130-137

Oliveira AL, Almeida E, Silva FBR, Filho VFN. 2006. Elemental contents in exotic Brazilian tropical fruits evaluated by energy dispersive x-ray fluorescence. Sci. Agric. 63: 82-84.

Pérez RCS, Canavaciolo VLG, Murillo RV, Rico CLM, Leyes EAR, Delange DM. 2015. Material vegetal seco y molido del fruto de *Acrocomia crispa* para la producción de D005. Rev. Cubana Plant. Med. 19: 246- 252.

Quero HJ. 1994. Flora de Veracruz. Palmae. Instituto de Ecologia, A.C. Xalapa, Veracruz, México, v. 81.

Ramos MIL, Ramos Filho MM, Hiane, PA, Braga Neto JA, Siqueira EMA. 2008. Nutritional quality of the pulp of bocaiuva *Acrocomia aculeata* (Jacq.) Lodd. Ciênc. Tecnol. Aliment. 90-94.

Rocha MS, Figueiredo RW, Araújo MAM, Áraújo RSR. 2013. Physical and chemical characterization and antioxidant activity (in vitro) of fruit of the Piaui savanna. Rev. Bras. Frutic. 35: 933-941.

Roig Mesa JT. 1945. Plantas medicinales, aromáticas o venenosas de Cuba. Havana: Editorial Guerrero Casamayor Y cía.

Rufino LMA, Barreto SMP, Duarte ER, Geraseev LC, Santos ACR, Jaruche YJ. 2011. Effects of inclusion of macaúba cake on ruminal protozoa population of goats. Rev. Bras. Zootec. 40: 899-903.

Santos ACR, Azevedo RA, Virginio Júnior GF, Rodriguez NM, Duarte ER, Geraseev LC. 2017. Effects of macauba cake on profile of rumen protozoa of lambs. Rev. Bras. Zootec.46:251-256.

Sauer IL. 2006. Energias renováveis: ações e perspectivas na Petrobras. Bahia Análise & Dados, Salvador. 16: 9-22.

Scariot AO, Lleras E, Hay JD. 1995. Flowering and fruiting phenology of the palm *Acrocomia aculeata*: patterns and consequences. Biotropica. 27: 168-173.

Silva GCR, Caño Andrade MH. 2013. Development and simulation of a new oil extraction process from fruit of macauba palm tree. J. Food Process. En, 36: 134-145.

Silva JC, Barrichelo LEG, Brito JO. 1986. Endocarpos de Macaúba e de Babaçu comparados a madeira de *Eucaliptus grandis* na produção de carvão vegetal. IPEF, Piracicaba. 34: 31-34.

Souza CTF, 2013. Desenvolvimento, maturação e sistemas de colheita de frutos da macaw palm (*Acrocomia aculeata*). Dissertação de Mestrado. Universidade Católica Dom Bosco, Campo Grande, MS.

Souza VHA, Santos LTS, Campos AF, Carolino J. 2016. "Um panorama do biodiesel no Brasil e no mundo: esforços para a ampliação do setor e desafios. Augustus. 21: 117-130.

Trentini CP, Oliveira DM, Zanette CM, Silva C. 2016. Low-pressure solvent extraction of oil from macaúba (*Acrocomia aculeata*) pulp: characterization of oil and defatted meal. Ciênc. Rural. 46: 725-731.

USDA. *United States Department of Agriculture.*www.fas.usda.gov. 2014, accessed may 20 th 2017.

USDA. *United States Department of Agriculture.*www.fas.usda.gov. 2017. accessed may 20 th 2017.

Vilas-Boas MA, Carneiro ACO, Vital BR, Carvalho AMML, Martins MA. 2010. Effect of carbonization temperature and the macauba residues in the production of charcoal. Scientia Forestalis. 38: 481- 490.

CONCLUSÕES GERAIS

Esta foi o primeiro trabalho que utilizou dados originados de NGS no gênero Acrocomia. A presente pesquisa mostrou a eficácia dos marcadores SNPs derivados de genotipagem por sequenciamento (GBS) para a caracterização da diversidade genética e da estrutura populacional de *A. aculeata* e outras espécies do gênero. Por ser uma abordagem que independe da necessidade de genoma de referência, representa uma eficiente ferramenta para estudos genômicos futuros.

Nosso estudo forneceu um panorama mais completo sobre a diversidade estrutura genética da de Acrocomia em relação a estudos prévios. Em *A. aculeata*, nossos resultados evidenciam pela primeira vez dois grupos genéticos associados a origem geográficas das populações (Norte-Sul do equador) com níveis diferentes de diversidade genômica. No Brasil, as populações apresentaram altos níveis de diversidade, sugerindo que há variação genética suficiente para ser explorada visando obter ganhos nos programa de melhoramento genético e avançar na seleção de plantas superiores sem restringir a base genética nas populações melhoradas. Em contraste, os países de Centro e Norte américa evidenciaram baixos níveis de diversidade genômica. Portanto, recomendamos avaliar os impactos da exploração local atual na diversidade genética da espécie, assim como realizar estudos em uma escala mais fina par se ter um panorama mais preciso da diversidade em cada país.

As informações fornecidas por este estudo são referência para estudos futuros e contribuem para a exploração sustentável de populações nativas de *A. aculeata* e de outras espécies do gênero mediante o desenvolvimento de estratégias de melhoramento genético e conservação específicos para cada espécie.

REFERENCIAS

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

Abreu IS, Carvalho CR, Carvalho GMA, Motoike SY. First karyotype, DNA C-value and AT/GC base composition of macaw palm (*Acrocomia aculeata*, Arecaceae)–a promising plant for biodiesel production. Aust J Bot. 2011;59(2):149-155.

Alfaro-Solís JD, Montoya-Arroyo A, Jiménez VM, Arnáez-Serrano E, Pérez J, Vetter W, Frank J, Lewandowski I 2020. *Acrocomia aculeata* fruits from three regions in Costa Rica: an assessment of biometric parameters, oil content and oil fatty acid composition to evaluate industrial potential. Agrofor Syst., 1-15. https://doi.org/10.1007/s10457-020-00511-8

Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. Nature Rev Genet. 11(10):697-709.

Amaral FP 2007. Estudo das características físico-químicas dos óleos da amêndoa e polpa da macaúba [*Acrocomia aculeata* (jacq.) Lodd. Ex mart]. Universidade Estadual Paulista "Júlio de Mequita Filho".

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

Asmussen CB, Dransfield J, Deickmann V, Barfod AS, Pintaud J-C, Baker WJ. 2006. A new subfamily classification of the palm family (Arecaceae): evidence from plastid DNA phylogeny. Bot J Linn Soc. 151(1):15-38.

Barbosa RJ. 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.

Balick MJ. 1990. Production of coyol wine from *Acrocomia mexicana* (Arecaceae) in Honduras. Econ Bot. 44(1):84-93.

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

Bailey LH. 1941. Acrocomia preliminary paper. Gentes Herbarum – Occasional Papers on the Kinds Plants.

Barrett SC, 2003. Mating strategies in flowering plants: the outcrossing–selfing paradigm and beyond. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 358, 991-1004. https://doi.org/10.1098/rstb.2003.1301

Berton L. 2013Avaliação de populações naturais, estimativas de parâmetros genéticos e seleção de genótipos elite de macaúba (*Acrocomia aculeata*). Campinas, SP. Brazil: Instituto Agronômico de Campinas - IAC.

Billotte N, Risterucci, A-M, Barcelos, E, Noyer, J-L, Amblard, P, Baurens, F-C, 2001. Development, characterisation, and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. Genome 44, 413-425. https://doi.org/10.1139/g01-017

Blanca J, Cañizares J, Cordero L, Pascual L, Diez MJ, Nuez F. 2012. Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato. PloS one. 7(10): e48198.

Bora PS, Rocha R. 2004. Macaiba palm: fatty and amino acids composition of fruits macaíba: composición de aminoácidos y ácidos grasos de frutos macaíba: composición de aminoácidos e ácidos graxos de froitos. CYTA-J Food. 4(3):158-162.

Boutet G, Carvalho SA, Falque M, Peterlongo P, Lhuillier E, Bouchez O, et al. 2016. SNP discovery and genetic mapping using genotyping by sequencing of whole genome genomic DNA from a pea RIL population. BMC genom. 17(1):1-14.

Breed MF, Christmas MJ, Lowe AJ, 2014. Higher levels of multiple paternities increase seedling survival in the long-lived tree *Eucalyptus gracilis*. Plos one 9, e90478. https://doi.org/10.1371/journal.pone.0090478

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

Byerlee D, Falcon WP, Naylor R., 2017. The tropical oil crop revolution: food, feed, fuel, and forests. Oxford University Press. https://doi.org/10.1111/joac.12272

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011. Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes genom genet. 1(3):171-182.

Cavers S, Degen B, Caron H, Lemes MR, Margis R, Salgueiro F, et al. . 2005. Optimal sampling strategy for estimation of spatial genetic structure in tree populations. Heredity 95(4):281-289.

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

Charlesworth D, Wright SI, 2001. Breeding systems and genome evolution. Curr Opin Genet Dev. 11, 685-690. https://doi.org/10.1016/S0959-437X(00)00254-9

Chen D, Zhang X, Kang H, Sun X, Yin S, Du H, et al. 2012. Phylogeography of Quercus variabilis based on chloroplast DNA sequence in East Asia: multiple glacial refugia and mainland-migrated island populations. PLoS One. 7(10):e47268.

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

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

Coimbra MC, Jorge N. 2011. Characterization of the pulp and kernel oils from Syagrus oleracea, *Syagrus romanzoffiana*, and *Acrocomia aculeata*. J Food Sci.;76(8):C1156-C1161.

Coimbra MC Jorge N, 2011. Proximate composition of guariroba (*Syagrus oleracea*), jerivá (*Syagrus romanzoffiana*) and macaúba (*Acrocomia aculeata*) palm fruits. Food Res Int. 44, 2139-2142. https://doi.org/10.1016/j.foodres.2011.03.032

Cole CT. 2003. Genetic variation in rare and common plants. Annu Rev Ecol Evol Syst. 34(1):213-237.

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

Conceição L, Junqueira NTV, Licurgo FdS, Antoniassi R, Wilhelm AE, Braga MF. 2012. Teor de óleo em frutos de diferentes espécies de macaubeira (Acrocomia spp.). Embrapa Cerrados-Artigo em anais de congresso (ALICE); 2012: In: Congresso brasileiro de fruticultura, 22., Bento Gonçalves. Anais

Cook OF. A synopsis of the palms of Puerto Rico. The Botanical Club. 281901. p. 565-567.

Cornejo OE, Yee M-C, Dominguez V, Andrews M, Sockell A, Strandberg E, et al. Population genomic analyses of the chocolate tree, *Theobroma cacao* L., provide insights into its domestication process. Communications biology. 2018;1(1):1-12.

Costa CJ, Marchi ECS. 2008.Germinação de sementes de palmeiras com potencial para produção de agroenergia. Embrapa Cerrados-Documentos (INFOTECA-E).

Couvreur TL, Hahn W, Granville J-Jd, Pham J-L, Ludena B, Pintaud J-C. 2007. Phylogenetic relationships of the cultivated Neotropical palm *Bactris gasipaes* (Arecaceae) with its wild

relatives inferred from chloroplast and nuclear DNA polymorphisms. Syst Bot. 32(3):519-530.

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.

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 César A, de Azedias Almeida F, de Souza RP, Silva GC, Atabani A. 2015. The prospects of using *Acrocomia aculeata* (macaúba) a non-edible biodiesel feedstock in Brazil. Renew Sust Energ Rev. 49:1213-1220.

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

Davey JW, Blaxter ML. 2010. RADSeq: next-generation population genetics. Brief Funct Genomics. 9(5-6):416-423.

Díaz-Fuentes VH, Avendaño-Arrazate CH, Gálvez-Marroquín LA, Ruíz-Cruz PA., Iracheta-Donjuán L, Nájera-Domínguez W, Martínez-Valencia BB, 2019. Caracterizacion morfoagronomica de coyol (*Acrocomia aculeata* Jacq.) para determinar su potencial productivo de aceite para biodiesel. AGROProductividad 12, 97-105.

Diaz FVH, Cruz AP, Matinez VBB, Moises A-B. 2020. Caracterización morfológica del coyol *Acrocomia aculeata* (Jacq.), y su potencial productivo de aceite vegetal. In: Instituto Nacional de Investigaciones Forestales AyPCdIRPSCERI, editor. Folleto Técnico Tuxtla Chico, Chiapas, México p. 54.

Domiciano GP, Alves AA, Laviola BG, Conceição LDHCSd, 2015. Parâmetros genéticos e diversidade em progênies de Macaúba com base em características morfológicas e fisiológicas. Cienc Rural 45, 1599-1605. https://doi.org/10.1590/0103-8478cr20140909

dos Reis EF, Pinto JFN, da Assunção HF, da Costa Netto AP, da Silva DFP, 2019. Characteristics of 137 macaw palm (*Acrocomia aculeata*) fruit accessions from Goias, Brazil. Comun Sci. 10, 117-124. https://doi.org/10.14295/cs.v10i1.2389

Dransfield J, Uhl NW, Lange CBA, Baker WJ, Harley MM, Lewis CE. Genera Palmarum: the evolution and classification of palms: Kew Publishing; 2008.

Dray S, Dufour AB, Chessel D. 2007. The ade4 package-II: Two-table and K-table methods. R news. 7(2):47-52.

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 DF. 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

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. Focus. 12(13):39-40.

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

Eiserhardt WL, Pintaud J-C, Asmussen-Lange C, Hahn WJ, Bernal R, Balslev H, et al. 2011. Phylogeny and divergence times of Bactridinae (Arecaceae, Palmae) based on plastid and nuclear DNA sequences. Taxon. 60(2):485-498.

Ellis J, Pashley C, Burke J, McCauley D. 2006.High genetic diversity in a rare and endangered sunflower as compared to a common congener. Mol Ecol 15(9):2345-2355.

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

Eguiarte LE, Perez-Nasser N, Piñero D, 1992. Genetic structure, outcrossing rate and heterosis in *Astrocaryum mexicanum* (tropical palm): implications for evolution and conservation. Heredity 69, 217. https://doi.org/10.1038/hdy.1992.119

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

Flanagan SP, Jones AG. 2017. Constraints on the F ST–heterozygosity outlier approach. J Hered. 108(5):561-573.

Franks SJ. 2010. Genetics, evolution, and conservation of island plants. J Plant Biol. 53(1):1-9.

Frichot E, Schoville SD, Bouchard G, François O. 2013.Testing for associations between loci and environmental gradients using latent factor mixed models. Mol Biol Evol. 30(7):1687-1699.

Frichot E, François O. 2015. LEA: An R package for landscape and ecological association studies. Methods Ecol Evol. 6(8):925-929.

Furches MS, Small RL, Furches A. 2013. Genetic diversity in three endangered pitcher plant species (Sarracenia; Sarraceniaceae) is lower than widespread congeners. Am J Bot. 100(10):2092-2101.

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

Gunn BF. 2004. The phylogeny of the Cocoeae (Arecaceae) with emphasis on *Cocos nucifera*. Ann Missouri Bot Gard. 505-522.

Hamrick J, Godt M. 1989Allozyme diversity in plant species. In 'Plant population genetics, breeding and genetic resources'. (Eds AHD Brown, MT Clegg, AL Kahler, BS Weir) pp. 43–63. Sinauer Associates: Sunderland, MA.

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

Hahn WJ. 2002. A phylogenetic analysis of the Arecoid line of palms based on plastid DNA sequence data. Mol Phylogenet Evol. 23(2):189-204.

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.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 25:965–1978.

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

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. 2010. Population genomics of parallel adaptation in three spine stickleback using sequenced RAD tags. PLoS genet. 6(2):e1000862.

Holsinger KE, Weir BS. 2009. Genetics in geographically structured populations: defining, estimating, and interpreting F ST. Nat Rev Genet. 10(9):639-650.

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.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 23(14):1801-1806.

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.

Kalia RK, Rai MK, Kalia S, Singh R, Dhawan A, 2011. Microsatellite markers: an overview of the recent progress in plants. Euphytica 177, 309-334. https://doi.org/10.1007/s10681-010-0286-9

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

Kalinowski ST. 2005. Hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Molecular ecology notes. 5(1):187-189.

Kamvar ZN, Brooks JC, Grünwald NJ. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Frontiers in genetics. 6:208.

Kantar MB, Betts K, Michno J-M, Luby JJ, Morrell PL, Hulke BS, Stupar RM, Wyse DL, 2014. Evaluating an interspecific *Helianthus annuus* × *Helianthus tuberosus* population for use in a perennial sunflower breeding program. Field Crops Res. 155, 254-264. https://doi.org/10.1016/j.fcr.2013.04.018

Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA. 2013. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods in ecology and evolution 4(8):782-788.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Molecular ecology resources. 15:1179-1191.

Kubota S, Kameyama Y, Hirao AS, Ohara M, 2008. Adaptive significance of self-fertilization in a hermaphroditic perennial, *Trillium camschatcense* (Melanthiaceae). Am J Bot. 95, 482-489. https://doi.org/10.3732/ajb.95.4.482

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 ECM, Motoike, SY, Kuki KN, Resende MDV, Caixeta ET. 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

Lefèvre F, 2004. Human impacts on forest genetic resources in the temperate zone: an updated review. For Ecol Manag. 197, 257-271. https://doi.org/10.1016/j.foreco.2004.05.017

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

Lentz DL. 1990. *Acrocomia mexicana*: Palm of the ancient Mesoamericans. J Ethnobiol. 10(2):183-194.

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.

Levy E, Byrne M, Coates D, Macdonald B, McArthur S, Van Leeuwen S. 2016. Contrasting influences of geographic range and distribution of populations on patterns of genetic diversity in two sympatric Pilbara Acacias. PLoS One. 11(10):e0163995.

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

Liu X, Tang H, Li D, Hou L, 2011. Genetic diversity of coconut cultivars in China by microsatellite (SSR) markers. Mol Plant Breed. 2. https://doi.org/10.5376/mpb.2011.02.0012

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.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. 2003. The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet. 2003;4(12):981-994.

Luu K, Bazin E, Blum MG. 2017. pcadapt: an R package to perform genome scans for selection based on principal component analysis. Mol Ecol Resour. 17(1):67-77.

Machado SR. 2003.Variaciones en la morfología polínica de Arecaceae en Cuba: abertura tricotomosulcada y estratificación de la exina. Revista Jard Bot Nac Univ Habana. 71-79.

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

Martins H, Caye K, Luu K, Blum MG, François O. 2016. Identifying outlier loci in admixed and in continuous populations using ancestral population differentiation statistics. Molecular ecology. 25(20):5029-5042.

Martius CFPV. Historia Naturalis Palmarum. Leipzig, TC Weigel. 1824:285-286.

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.

Meerow AW, Noblick L, Salas-Leiva DE, Sanchez V, Francisco-Ortega J, Jestrow B, et al. 2015. Phylogeny and historical biogeography of the cocosoid palms (Arecaceae, Arecoideae, Cocoseae) inferred from sequences of six WRKY gene family loci. Cladistics. 31(5):509-534.

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

Mengistu FG, Motoike SY, Caixeta ET, Cruz CD, Kuki KN, 2016. Cross-species amplification and characterization of new microsatellite markers for the macaw palm, *Acrocomia aculeata* (Arecaceae). Plant Genet Resour. 14, 163-172. https://doi.org/10.1017/S1479262115000179

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

Mitsui Y, Chen S-T, Zhou Z-K, Peng C-I, Deng Y-F, Setoguchi H. 2008. Phylogeny and biogeography of the genus Ainsliaea (Asteraceae) in the Sino-Japanese region based on nuclear rDNA and plastid DNA sequence data. Ann Bot. 101(1):111-124.

Moore HE. 1968. *Gastrococos crispa* (Kunth). Principes. 11.

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.

Morin PA, Martien KK, Taylor BL. 2009. Assessing statistical power of SNPs for population structure and conservation studies. Mol Ecol Resour. 9(1):66-73.

Murawski D, Hamrick J, 1991. The effect of the density of flowering individuals on the mating systems of nine tropical tree species. Heredity 67, 167-174. https://doi.org/10.1038/hdy.1991.76

Murawski D, Hamrick J, 1992. Mating System and Phenology *Ceiba Pentadra* (Bombacaeae) in Central Panama. J Hered. 83, 401-404. https://doi.org/10.1093/oxfordjournals.jhered.a111241

Nakamura K, Denda T, Kokubugata G, Huang C-J, Peng C-I, Yokota M. 2015. Phylogeny and biogeography of the Viola iwagawae-tashiroi species complex (Violaceae, section Plagiostigma) endemic to the Ryukyu Archipelago, Japan. Plant Syst Evol. 301(1):337-351.

Nascimento AR. 2010. Riqueza e etnobotânica de palmeiras no Território Indígena Krahô. Tocantins. Brasil. Floresta 40:209-220. http://dx.doi.org/10.5380/rf.v40i1.17112

Nazareno AG, Reis MSd, 2012. Linking phenology to mating system: exploring the reproductive biology of the threatened palm species Butia eriospatha. J Hered. 103, 842-852. https://doi.org/10.1093/jhered/ess070

Neiva DS, Melo Júnior AF, Oliveira DA, Royo VA, Brandão 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

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89(3):583-590.

Noblick LR, Hahn WJ, Griffith MP. 2013. Structural cladistic study of Cocoseae, subtribe Attaleinae (Arecaceae): Evaluating taxonomic limits in Attaleinae and the neotropical genus Syagrus. Brittonia. 65(2):232-261.

Nucci SM, 2007. Desenvolvimento, caracterização e análise da utilidade de marcadores microssatélites em genética de população de macaúba. Master dissertation. Intituto Agrônomico, Campinas, SP.

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

OCDE/FAO, 2019. OCDE-FAO Perspectivas Agrícolas 2019-2028, OECD Publishing, París/Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), Roma https://doi.org/10.1787/7b2e8ba3-es

Oliveira DA, Melo JAF, Brandão 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.

Palmweb. Acrocomia Mart. Hist. Nat. Palm. n. d [Available from: http://www.palmweb.org/.

Park S, 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. University of Dublin. Ph. D. Thesis. [thesis], pp 254.

Patiño RVM. 2002. Historia y dispersión 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, Bordács S, Burg K, Coart E, Cottrell J, van Dam B, Deans JD, Dumolin-Lapègue S, Fineschi S, Finkelday R, Gillies A, Gla I, Goicoechea PG, Jensen JS, König A, Lowe AJ, Madsen SF, Mátyás 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

Picanço-Rodrigues D, Astolfi-Filho S, Lemes MR, Gribel R, Sebbenn AM, Clement CR, 2015. Conservation implications of the mating system of the Pampa Hermosa landrace of peach palm analyzed with microsatellite markers. Genet Mol Biol. 38, 59-66. https://doi.org/10.1590/S1415-475738120140022

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, Mejía K, Millán B, Moraes M, Noblick L, Stauffer FW, Kahn F 2008. Las palmeras de América del Sur: diversidad, distribución e história evolutiva. Rev Peru Biol 15:7-29.

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

Poland JA, Rife TW. 2012. Genotyping-by-sequencing for plant breeding and genetics. The Plant Genome 5(3):92-102.

Poland JA, Brown PJ, Sorrells ME, Jannink J-L. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS one 7(2):e32253.

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

Proctor GR. Arecaceae (Palmae). P Acevedo-Rodríguez, M T Strong. 2005. Monocots and Gymnosperms of Puerto Rico and the Virgin Islands.: Smithsonian Institution: Contributions of the United States National Herbarium.

Pugesgaard S, Schelde K, Larsen SU, Lærke PE, Jørgensen U, 2015. Comparing annual and perennial crops for bioenergy production–influence on nitrate leaching and energy balance. Gcb Bioenergy 7, 1136-1149. https://doi.org/10.1111/gcbb.12215

Qi X-S, Yuan N, Comes HP, Sakaguchi S, Qiu Y-X. 2014. A strong 'filter'effect of the East China Sea land bridge for East Asia's temperate plant species: inferences from molecular phylogeography and ecological niche modelling of Platycrater arguta (Hydrangeaceae). BMC Evol Biol 14(1):41.

Qiu YX, Sun Y, Zhang XP, Lee J, Fu CX, Comes HP. 2009. Molecular phylogeography of East Asian Kirengeshoma (Hydrangeaceae) in relation to Quaternary climate change and landbridge configurations. New Phytol. 183(2):480-495.

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/

Ramírez Hernández BC, Zañudo Hernández J, García de Alba Verduzco JE, Délano Frier JP, Pimienta Barrios E, García Martínez MÁ. 2013. Importancia agroecológica del coyul (*Acrocomia mexicana* Karw. ex Mart.). Estudios sociales (Hermosillo, Son). 21(41):95-113.

Ramírez-Rodríguez R, Tovar-Sánchez E, Jiménez RJ, Vega FK, Rodríguez 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

Ramos MIL, Ramos Filho MM, Hiane PA, Braga Neto JA, Siqueira EMA. 2008.Qualidade nutricional da polpa de bocaiúva *Acrocomia aculeata* (Jacq.) Lodd. Food Sci Technol; 28:90-94.

Ritland K. 1988. The genetic-mating structure of subdivided populations II. Correlated mating models. Theor Popul Biol. 34, 320-346. https://doi.org/10.1016/0040-5809(88)90027-5

Ritland K, 2002. Extensions of models for the estimation of mating systems using n independent loci. Heredity 88, 221-228. https://doi.org/10.1038/sj.hdy.6800029

Ritland K,El-Kassaby Y. 1985. The nature of inbreeding in a seed orchard of Douglas fir as shown by an efficient multilocus model. Theor Appl Genet 71, 375-384. https://doi.org/10.1007/BF00251176

Ritland K, Jain S. 1981. A model for the estimation of outcrossing rate and gene frequencies using n independent loci. Heredity 47, 35-52. https://doi.org/10.1038/hdy.1981.57

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

Roncal J, Kahn F, Millan B, Couvreur TL, Pintaud J-C. 2013. Cenozoic colonization and diversification patterns of tropical American palms: evidence from Astrocaryum (Arecaceae). Bot J Linn Soc. 171(1):120-139.

Rossi M, Bitocchi E, Bellucci E, Nanni L, Rau D, Attene G, et al. 2009. Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L. Evol Appl. 2(4):504-522.

Sakaguchi S, QIU YX, LIU YH, QI XS, KIM SH, Han J, et al. 2012. Climate oscillation during the Quaternary associated with landscape heterogeneity promoted allopatric lineage divergence of a temperate tree *Kalopanax septemlobus* (Araliaceae) in East Asia. Mol Ecol. 21(15):3823-3838.

Salis SM, Mattos PP. 2009. Floração e Frutificação da Bocaiúva (*Acrocomia aculeata)* e do Carandá (*Copernicia alba*) no Pantanal. Corumbá: Embrapa Pantanal. Comunicado Técnico 78.

Sanjinez-Argandoña EJ, Machado CCA. 2011. Caracterização biométrica, física e química 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

Schönswetter P, Solstad H, García 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

Sebbenn A. 2006. Sistema de reprodução em espécies arbóreas tropicais e suas implicações para a seleção de árvores matrizes para reflorestamentos ambientais. In: Higa A., R., Sulva, L., D. (Eds.) Pomares de sementes de espécies florestais nativas. Curitiba: FUPEF, 193-198.

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 Àrvore 41:1-9. http://dx.doi.org/10.1590/1806-90882017000500008

Simiqueli GF, de Resende, MDV, Motoike SY, Henriques E. 2018. Inbreeding depression as a cause of fruit abortion in structured populations of macaw palm (*Acrocomia aculeata*): Implications for breeding programs. Ind Crops Prod. 112, 652-659. https://doi.org/10.1016/j.indcrop.2017.12.068

Slovák M, Kucera J, Marhold K, Zozomová-Lihová 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

Souza L, Kageyama PY, Sebbenn AM. 2003. Sistema de reprodução em população natural de *Chorisia speciosa* A. St.-Hil.(Bombacaceae). Rev Bras Bot 26, 113-121. https://doi.org/10.1590/S0100-84042003000100012

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

Stuessy TF, Takayama K, López-Sepúlveda P, Crawford DJ. 2014. Interpretation of patterns of genetic variation in endemic plant species of oceanic islands. Bot J Linn Soc. 174(3):276-288.

Templeton AR. The reality and importance of founder speciation in evolution. BioEssays. 2008;30(5):470-479.

The Plant List. Version 1.1 2013 [Available from: http://www.theplantlist.org/

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

USDA, 2020. United States Department of Agriculture, www.fas.usda.gov

Whitehead MR, Lanfear R, Mitchell RJ, Karron JD. 2018. Plant mating systems often vary widely among populations. Front Ecol Evol. 6, 38. https://doi.org/10.3389/fevo.2018.00038

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

Velazquez MAM. 2013. Caracterización fitoquímica, molecular y evaluación de respuestas del cultivo in vitro de coyol (*Acrocomia mexicana* Kraw ex Marti)/por Antonio Magdiel Velazquez Mendez: Universidad Autónoma Chapingo.

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, Campos-Rocha A. Acrocomia in Flora do Brasil 2020 em construção: Jardim Botânico do Rio de Janeiro; 2020. [Disponivel em: <http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB15662.

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, Colombo C, editors. 2013.Distribuição geográfica de *Acrocomia aculeata* (jacq,) lodd ex mart.(arecaceae) em sua região de ocorrência. I Congresso brasileiro de macaúba.

Vianna SA, Hiane PA, Jordão 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.

Villanueva RAM, Chen ZJ. 2019. ggplot2: Elegant graphics for data analysis. Taylor & Francis.

Vinson C, Mangaravite E, Sebbenn A, Lander T. 2018. Using molecular markers to investigate genetic diversity, mating system and gene flow of Neotropical trees. Braz J Bot. 41(2):481-496.

Welch J, Ventura Santo R, Flowers NM, Coimbra Jr CEA. 2013. Na primeira margem do rio: território e ecologia do povo Xavante de Wedezé. Museu do Índio-FUNAI. Rio de Janeiro, Brazil.

Whittaker RJ, Fernández-Palacios JM. 2007. Island biogeography: ecology, evolution, and conservation: Oxford University Press.

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

Wright S. 1931. Evolution in Mendelian populations. Genetics 16(2):97.

Wright S, 1940. Breeding structure of populations in relation to speciation. The American Naturalist 74, 232-248. https://www.jstor.org/stable/2457575

Wright S. 1943. Isolation by distance. Genetics 28(2):114.

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

Zehdi S, Trifi M, Billotte N, Marrakchi M, Christophe Pintaud J. 2004. Genetic diversity of Tunisian date palms (Phoenixdactylifera L.) revealed by nuclear microsatellite_polymprohism. Hereditas 141, 278-287. https://doi.org/10.1111/j.1601-5223.2004.01855.x

ANEXO I. Declaração de Bioética e/ou Biossegurança

UNICAMP
COORDENADORIA DE PÓS-GRADUAÇÃO
INSTITUTO DE BIOLOGIA
Universidade Estadual de Campinas
Caixa Postal 6109. 13083-970, Campinas, SP, Brasil
Fone (19) 3521-6378. email: cpgib@unicamp.br
IB

DECLARAÇÃO

Em observância ao **§5º do Artigo 1º da Informação CCPG-UNICAMP/001/15**, referente a Bioética e Biossegurança, declaro que o conteúdo de minha Tese de Doutorado, intitulada "***Genômica e estrutura populacional da macaúba [Acrocomia aculeata (Jacq.) Lodd. ex Mart.], visando subsidiar à domesticação da espécie.***", desenvolvida no Programa de Pós-Graduação em Genética e Biologia Molecular do Instituto de Biologia da Unicamp, não versa sobre pesquisa envolvendo seres humanos, animais ou temas afetos a Biossegurança.

Assinatura: _____
Nome do(a) aluno(a): Brenda Gabriela Díaz Hernández

Assinatura: _____
Nome do(a) orientador(a): Maria Imaculada Zucchi

Data: 05/01/2021

ANEXO II. Direitos autorais

Declaração

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **Genômica e estrutura populacional da macaúba [Acrocomia aculeata (Jacq.) Lodd. ex Mart.], visando subsidiar à domesticação da espécie.**, não infringem os dispositivos da Lei n.° 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 05/01/2021

Assinatura : _____

Nome do(a) autor(a): **Brenda Gabriela Díaz Hernández**

RG n.° V843463-O

Assinatura : _____

Nome do(a) orientador(a): **Maria Imaculada Zucchi**

RG n.° 19810213-6