

ÉDER CRISTIAN MALTA DE LANES

**VARIABILIDADE MOLECULAR E SISTEMA DE REPRODUÇÃO DE
MACAÚBA (*Acrocomia aculeata*)**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de Doctor Scientiae.

VIÇOSA
MINAS GERAIS – BRASIL
2014

**Ficha catalográfica preparada pela Biblioteca Central da
Universidade Federal de Viçosa - Câmpus Viçosa**

T

L267v
2014 Lanes, Éder Cristina Malta de, 1983-
Variabilidade molecular e sistema de reprodução de
macaúba (*Acrocomia aculeata*) / Éder Cristina Malta de
Lanes. - Viçosa, MG, 2014.
xiv, 112f. : il. ; 29 cm.

Inclui apêndices.

Orientador : Sérgio Yoshimitsu Motoike.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Macaúba - Variedades. 2. *Acrocomia aculeata*.
3. Macaúba - Melhoramento genético. 4. Microssatélites.
I. Universidade Federal de Viçosa. Departamento de
Fitotecnia. Programa de Pós-graduação em Genética e
Melhoramento. II. Título.

CDD 22. ed. 631.52

ÉDER CRISTIAN MALTA DE LANES

**VARIABILIDADE MOLECULAR E SISTEMA DE REPRODUÇÃO DE
MACAÚBA (*Acrocomia aculeata*)**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de Doctor Scientiae.

APROVADA: 16 de dezembro de 2014

Eveline Teixeira Caixeta
(Coorientadora)

Leonardo Duarte Pimentel

Kacilda Naomi Kuki

Elisa Ferreira Moura Cunha

Sérgio Yoshimitsu Motoike
(Orientador)

“A grande diferença entre o realista e o otimista, é a informação”.

(Charles Clement, 2012)

“O conhecimento adquirido quando não exercitado constantemente, se perde, “atrofia”. Está é uma prova de que Lamarck não estava totalmente errado”.

(Éder Lanes, 2013)

“O livro formado por conceitos, teorias e hipóteses da própria experiência prática, será sempre mais relevante do que uma obra construída pela simples união de ideias fragmentadas e alheias”.

(Éder Lanes, 2014)

“Diante da incontável intervenção humana sobre o meio ambiente, resta-nos conservar a variabilidade genética existente e continuarmos a rezar para que nenhuma espécie selvagem entre em Equilíbrio de Hardy-Weinberg”.

(Éder Lanes, 2014)

“Um germoplasma sem identificação na BAG é o mesmo que um livro em branco em uma biblioteca”

(Renato Veiga, 2012)

Ofereço

*A minha esposa, Julieta Lanes
E ao meu filho, Heitor Lanes*

Dedico

*Aos meus queridos pais, João Lanes e Sueli Lanes
Ao meu irmão e amigo, Everton Lanes
A toda a minha família,
E principalmente a Deus pela proteção constante.*

AGRADECIMENTOS

Os agradecimentos é sem dúvida um processo muito complicado, pois, mesmo sem querer cometemos injustiças. Isso acontece porque somos traídos pelos os lapsos da memória recente, e da amnésia profunda, que nos faz acreditar que apenas as etapas de graduação e pós-graduação seriam os únicos estágios relevantes na formação de um profissional. Mesmo assim, ciente desta enorme injustiça e da incapacidade de expressar minha verdadeira gratidão, irei realiza-la de maneira superficial.

Acima de tudo, gostaria de agradecer primeiramente a Deus pela vida, força, coragem, saúde, paciência, paz e discernimento para enfrentar os desafios durante toda essa caminhada.

Ao meu orientador e amigo Sérgio Yoshimitsu Motoike, agradeço pela oportunidade, confiança, ensinamento, ótima convivência e estímulo em buscar sempre mais conhecimento. Agradeço também pelos conselhos e compreensão nos momentos mais difíceis.

Aos meus co-orientadores, Dra. Eveline Teixeira Caixeta e Dr. Marcos Deon Vilela de Resende pela amizade, humildade, competência, e disponibilidade em ajudar nas metodologias de biologia molecular e análises estatísticas.

À Dra. Kacilda Naomi Kuki, pelo enorme incentivo e confiança, mas principalmente, pela admirável disponibilidade em ajudar e ensinar, competência, empenho e pela amizade, pelos conselhos e pela orientação em todas as etapas deste trabalho.

À Dra. Elisa Ferreira Moura Cunha e o Dr. Leonardo Duarte Pimentel agradeço pelo olhar crítico e construtivo, e também por ter aceitado participar da banca de tese.

À CAPES — Coordenação de Aperfeiçoamento de Pessoal de Nível Superior e ao CNPq — Conselho Nacional de Desenvolvimento Científico e Tecnológico pelo o financiamento da bolsa de doutorado (nº 143353/2011-0), o qual foi fundamental para o desenvolvimento deste projeto. À PETROBRAS — Petróleo Brasileiro S.A. pelo apoio financeiro (nº 0050.0061.571.109), que tornaram possíveis as viagens de campo e análises de laboratório.

À Universidade Federal de Viçosa, Departamento de Fitotecnia, Programa de Pós-Graduação em Genética e Melhoramento pela disponibilidade da infraestrutura, laboratórios e funcionários, fundamentais para a realização do projeto. À todos os professores da Universidade Federal de Viçosa, em especial ao Prof. José Marcelo Soriano Viana, pelas sugestões e palavras de incentivo durante o doutoramento. E

também pela alegria e descontração no futebol de sábado. Ao Prof. Carlos Nick Gomes pelas as sugestões e momentos de descontração nas viagens de trabalho. Aos funcionários da secretaria do Programa de Genética e Melhoramento, onde fui sempre bem recebido.

Aos amigos da Rede Macaúba de Pesquisa — REMAPE/UFV em especial ao Samuel Goulart, Angelica Rueda, Ana Paula Andrade, Renata Freitas, Thais Corrêa, Sebastián Montoya, Anderson Evaristo, Rafael Santos, Manuela Granja, Elma Souza, Thiago Pires, Gustavo Silveira, Alberto Boldt e Fekadu Mengistu pela a agradável convivência. E aos funcionários do Laboratório de Biotecnologia e Melhoramento Vegetal (LBMV), em especial a Telma Miranda e Francisco de Assis pelo estímulo, a agradável convivência e a disponibilidade em ajudar e ensinar.

Ao meu grande amigo desde a segunda série Marciano Oliveira. E também as grandes amizades que realizei durante a pós-graduação, pessoas que tenho um enorme apreço, respeito e confiança Paulo Mafra, Ciro Maia e Paulo Coutinho. Faço um agradecimento especial ao amigo Paulo Mafra pela humildade, ética, um exemplo de profissional a ser seguido, e que contribuiu muito na elaboração deste trabalho.

A minha mãe Sueli Aparecida Malta de Lanes, e ao meu pai João Alves de Lanes pelo constante incentivo, pelos esforços dedicados a minha educação e formação cristã, pelo amor e fonte de luz nesta longa caminhada, pelas preces e pelas lições de vida. Ao meu irmão Everton Lanes, pelas palavras de apoio e carinho, pela amizade, pelo amor, à cumplicidade e união. Aos meus primos ou irmãos Wilson Amorim e Pedro Amorim, e a minha tia Marly Barbosa por ter me recebido e acolhido em sua casa durante todo o período de graduação e parte do ensino médio.

Aos meus avós José Lanes e Felizarda Lanes pelo exemplo de vida e de fé. E também à memória de meus avós falecidos: Valentin Oliveira e Nelza Oliveira.

Aos meus sogros Jairo Silveira e Regina Silveira, pela a compreensão e incentivo, pelas palavras de apoio e pelas preces nos momentos difíceis.

A minha esposa Julieta Lanes pela paciência, compreensão, companheirismo, incentivo, fé, dedicação, amor e carinho. E ao meu filho Heitor Lanes que me proporcionou a oportunidade de conhecer o verdadeiro amor.

E a todos aqueles que direta ou indiretamente, contribuíram para a realização deste trabalho, bem como todas as pessoas que desejam o meu sucesso ou que de alguma forma fizeram parte da minha trajetória acadêmica. Muito Obrigado!

BIOGRAFIA

ÉDER CRISTIAN MALTA DE LANES, filho de João Alves de Lanes e Sueli Aparecida Malta de Lanes, nasceu em 06 de setembro de 1983, em Juiz de Fora, Minas Gerais.

Em julho do ano de 2007, graduou-se em Ciências Biológicas pelo o Centro de Ensino Superior de Juiz de Fora - CES/JF - PUC MINAS, Juiz de Fora – MG e Magister Scientiae em Genética e Melhoramento de Plantas, em julho de 2010, pela a Universidade Federal de Viçosa, Viçosa – MG.

Em agosto de 2010, iniciou o curso de Doutorado em Genética e Melhoramento pela a Universidade Federal de Viçosa, com a orientação do Professor Sérgio Yoshimitsu Motoike, defendendo a tese em dezembro de 2014.

SUMÁRIO

LISTA DE FIGURAS	ix
LISTA DE TABELAS	xii
RESUMO	xiii
ABSTRACT	xiv
1. INTRODUÇÃO GERAL	1
2. REVISÃO DE LITERATURA	2
2.1. Acrocomia aculeata - características gerais e distribuição geografia	2
2.1.1. Aspectos reprodutivo da <i>Acrocomia aculeata</i>	3
2.1.2. Potencialidades do fruto da <i>Acrocomia aculeata</i>	4
2.1.3. Sustentabilidade da cadeia produtiva da <i>Acrocomia aculeata</i>	5
2.1.4. Banco ativo de germoplasma de <i>Acrocomia aculeata</i>	7
2.2. Biologia molecular no melhoramento e conservação genética	8
2.2.1 Marcadores Moleculares e suas aplicações.....	8
2.2.2 Random Amplification of Polymorphic DNA (RAPD).....	9
2.2.3 Simple Sequence Repeats (SSR).....	9
2.3. Análise e importância do sistema de reprodução	10
2.3.1. Software Multilocus Mating System Program - MLTR	12
2.4. Genética de populações	12
2.4.1. Estrutura genética de população.....	12
2.4.2. Estatística F (Wright, 1951)	13
2.4.3. Coeficientes de coancestria (Cockerhan, 1969).....	14
2.4.4. Software Structure	15
2.4.5. Tamanho efetivo populacional	15
2.4.6. Depressão por endogamia	16
2.4.7. Diversidade gênica (Nei, 1972)	17
2.4.8. Diversidade genética intrapopulacional	17
2.5.8.1. Frequência alélica	17
2.5.8.2. Porcentagem de locos polimórficos (P).....	18
2.5.8.3. Número médio de alelos por loco polimórfico (N_a).....	18
2.5.8.4. Número efetivo de alelos (N_e).....	18
2.5.8.5. Conteúdo de informação polimórfica (PIC)	18
2.5.8.6. Heterozigosidade.....	19
3. OBJETIVOS.....	20
4. ESTRUTURA DA TESE	21
5. REFERÊNCIAS	22
CAPÍTULO 1	
Molecular characterization and population structure of the macaw palm, <i>Acrocomia aculeata</i> (Arecaceae), ex situ germplasm collection using microsatellites markers	32
Abstract	33

Introduction	34
Materials and Methods	36
Plant material.....	36
Genomic DNA extraction and genotyping microsatellites	37
Data analysis	40
Results	42
Allelic diversity of microsatellite loci.....	42
Population structure and genetic diversity	42
Discussion	47
Funding	51
Acknowledgments	51
References	52
CAPÍTULO 2	
Sistema de cruzamento e estimativa de depressão por endogamia em macaúba, <i>Acrocomia aculeata</i> (Arecaceae) utilizando dados quantitativos e moleculares	58
Resumo	59
Introdução	60
Material and Métodos	62
Estabelecimento do banco de germoplasma e coleta de amostras	62
Extração de DNA genômico e análise de microssatélites	63
Análise da diversidade genética	63
Análise de paternidade e depressão por endogamia.....	66
Resultados	67
Diversidade genética	67
Taxa de fecundação cruzada e parâmetros do sistema de cruzamento	69
Análise de paternidade e depressão por endogamia.....	71
Discussão	72
Diversidade genética	72
Sistema de cruzamento da <i>A. aculeata</i>	73
Cruzamento correlacionados e entre indivíduos aparentados.....	74
Depressão endogâmica entre e dentro de famílias	76
Implicações para o melhoramento e a conservação genética	78
Agradecimentos	80
Material Suplementar	81
Referências	84
 CONCLUSÕES GERAIS	 96
APÊNDICES	97

LISTA DE FIGURAS

REVISÃO DE LITERATURA

- Figura 1.** Distribuição geográfica da *Acrocomia aculeata* no Brasil. Mapas (A) elaborado por Berton (2013); e (B) por consultoria de Harri Lorenzi do Instituto Plantarum (Um Pé de Quê?/ Instituto Plantarum - Harri Lorenzi, 2014)2
- Figura 2.** Diversidade morfológica da *Acrocomia aculeata*. Plantas com 4 anos de idade, oriundas dos Estados de Minas Gerais (esquerda) e Pernambuco (direita). Local: Laboratório de Biotecnologia e Melhoramento DFT – Universidade Federal de Viçosa, 2014 (Fonte: Acervo pessoal de Motoike, S.Y.).....3
- Figura 3.** Análise comparativa do potencial de produção de óleo vegetal por ha.ano⁻¹ entre as principais oleaginosas. Macaúba**: Trata-se do potencial de produção de óleo do mesocarpo por ha.ano⁻¹ do melhor genótipo selecionado em condições naturais (adaptado de Berton, 2013).....4
- Figura 4.** Esquema de aproveitamento das partes do fruto da macaúba e as respectivas aplicações dos co-produtos após o processamento (adaptado de Melo, 2012).....5
- Figura 5.** Banco Ativo de Germoplasma de Macaúba (BAG-Macaúba) localizada na cidade de Araponga - MG. (A) Imagem aérea da Gleba I com 54 famílias maternas plantadas em fevereiro de 2009; (B) Imagem aérea da Gleba II com 34 famílias maternas plantadas em novembro de 2009; (C) Coleção viva conservada ex situ.....8
- Figura 6.** Degradação do habitat natural da *Acrocomia aculeata* no município de Betim – MG. (A) Populações nativas; (B e C) Perda do habitat natural da espécie pela a ação antrópica; (D) Semente em processo de germinação natural presente no ambiente sob fragmentação; (E e F) Loteamento imobiliário da área fragmentada 11
- Figura 7.** Possíveis respostas do sistema de acasalamento após a perturbação do habitat. Árvores geneticamente semelhantes são representadas por cores idênticas. A falha na polinização é indicada por linhas pontilhadas coloridas seguidas pelo símbolo (X); A redução na polinização é indicada por linhas pontilhadas coloridas; e a polinização normal, por linhas de cor sólida; um aumento relativo da polinização é indicado por linhas coloridas com um símbolo (+). As respostas (b, aumento da autofecundação), (c, aumentou de cruzamentos endogâmicos) e (d, diminuição da diversidade de pólen) não são mutuamente exclusivos para espécies hermafroditas e autocompatíveis (adaptado de Breed et al., 2012) 11
- Figura 8.** Porcentagem média da captura alélica por número de acessos na coleção ex situ da palma *Leucothrinax morrisii* em relação há população original (adaptado de Namoff et al., 2010) 16

CAPÍTULO 1

Figure 1. Brazilian map depicting; (A) Provenance of *Acrocomia aculeata* accessions collected from different Brazilian states to compose the Macaw Palm Germplasm Bank (BAG – Macauba), with latitude and longitude in decimal degrees; (B) Proposed map of the diversity center for *Acrocomia aculeata*, in the state of Minas Gerais/Brazil, based on the examined accessions grouped into three subpopulations inferred in the investigations. Diverse color spots represent individuals from different gene pools37

Figure 2. Population structure of *Acrocomia aculeata* from the genebank BAG – Macauba. Mean posterior probabilities $\text{LnP}(D)$ and estimated Δk over fifteen repeats of Structure analysis; (A) Optimal population structure of 77 accessions of *A. aculeata* ($K=2$) and (B) sub-structuring ($K=3$) of Pop2 assessed by Structure. Each individual is represented by a thin vertical line, which is partitioned into five colored segments that represent the individual estimated membership (EM) to the five clusters.....43

Figure 3. Neighbor-joining (NJ) analysis of 77 *Acrocomia aculeata* accessions based on Nei's genetic distance. (A) Unrooted neighbor-joining tree of 77 accessions; each colored branch represents one accession collected from corresponding subpopulation inferred by Structure. (B) NJ tree of the four inferred subpopulation44

Figure 4. Principal coordinate analysis of the 77 *Acrocomia aculeata* accessions. Diversity between the accessions according to the values of the first two principal coordinates with clusters identified by their provenance (Brazilian States) (A) and population structure analysis (B)45

CAPÍTULO 2

Figura 1S. Esquema do estudo da *Acrocomia aculeata*. O diagrama mostra as fontes de plantas e o tipo de tecido vegetal usado para a extração de DNA. Árvores maternas ou Matrizes: São as árvores adultas nativas, doadoras de semente que deram origem as coleção viva da BAG-Macaúba; Progenie: Plantas jovens que compõem o Banco Ativo de Germoplasma da Macaúba. O termo "família materna" refere-se aos indivíduos descendentes de uma mesma árvore materna. (A) Para a pesquisa de diversidade genética e teste de paternidade foram utilizadas 27 árvores adultas nativas e 157 progênies (obtidas de 19 árvores matrizes); (B) O sistema de cruzamento foi analisado utilizando 157 progênies juntamente com as respectivas informações genotípicas de suas árvores matrizes ($n=19$); (C) E a depressão por endogamia foi estimada nas progênies de 13 famílias maternas, uma vez que, nestas famílias foram identificadas progênies de autofecundação pelo o teste de paternidade.....81

APÊNDICES

Apêndice A – Croqui do Banco de Germoplasma de Macaúba (BAG-Macaúba), glebas I, II, III. São indicados na cor vermelha os acessos (um indivíduo de cada família materna) de *Acrocomia aculeata* selecionados para o capítulo 1 da tese. A seleção do material genético foi baseada em dois critérios: (1) representatividade da coleção viva mantida na BAG-Macaúba e (2) os estados e regiões do Brasil onde foram realizadas as coletas.....98

Apêndice B – Esquematização do critério adotado para a seleção das famílias maternas de *Acrocomia aculeata* do capítulo 2 da tese. (A) Seleção do material genético com base nos resultados de diversidade genética de Manfio (2010); (B) Croqui do Banco de Germoplasma de Macaúba indicando na cor vermelha as famílias selecionadas; (C) Coleta do tecido vegetal das árvores maternas selvagens das respectivas famílias selecionadas na BAG-Macaúba99

Apêndice C – Racionalização para o melhoramento da *Acrocomia aculeata*. As informações geradas por este estudo permitiu a resposta para algumas etapas do programa e são indicadas em negrito na cor azul. Respostas em negrito na cor vermelha são possíveis procedimentos que podem ser incorporados no programa de melhoramento. CGC: Capacidade Geral de combinação; CEC: Capacidade Específica de combinação; SRR: Seleção Recorrente Recíproca; SRI: Seleção Recorrente Intrapopulacional; SRRI: Seleção Recorrente Recíproca Individual.....100

Apêndice D – Método de Seleção Recorrente Recíproca ou Interpopulacional (SSR), sugerido para o programa de melhoramento da <i>Acrocomia aculeata</i>	101
Apêndice E – Método de Seleção Recorrente Recíproca Individual (SSRI), sugerido para o programa de melhoramento da <i>Acrocomia aculeata</i>	102
Apêndice F – Figure 1. A. Macaw palm plant in its natural habitat. B. Exposed trunk after removal of debris. C. Stipe sample harvest with cylindrical cork cutter. D. and E. Stipe sample in Eppendorf tube. F. Storage of samples in liquid nitrogen	105
Apêndice F – Figure 2. Genomic DNA extracted from specimens of <i>Acrocomia aculeata</i> , using Doyle and Doyle protocols modified. A. Stipe tissues, B. Fresh leaf tissues, and C. lyophilized leaf tissues. Lanes 1-4 of A. , B. and C. correspond to their respective tissues. Lane M = 200-bp DNA ladder	108
Apêndice F – Figure 3. Standard SSR bands generated with primers Aacu12 (A) and Aacu07 (B). ST = Stipe tissues (Lanes 1-4); FL = fresh leaf tissues (Lanes 5-8); LL = lyophilized leaf tissues (Lanes 9-12). M = 200-bp DNA ladder.....	109

LISTA DE TABELAS

CAPÍTULO 1

Table 1. Identification of 77 <i>Acrocomia aculeata</i> accessions assessed from the Active Germplasm Bank of Macauba (BAG – Macauba) with the correspondent bank code and provenance (geographic and climatic characteristics)	38
Table 2. Microsatellite loci, along with their respective allele amplitude, number of alleles (A), major allele frequency (MAF), expected and observed heterozygosity (H), and polymorphism information content (PIC) in the 77 accessions of <i>Acrocomia aculeata</i>	40
Table 3. Molecular diversity of model-based subpopulations of <i>Acrocomia aculeata</i> as inferred by Structure: sample size (SS), average number of alleles per locus (Na), average effective number of alleles per locus (Ne), average number of different alleles with a frequency greater or equal to 0.05 (N_{Fa}), gene diversity (GD), observed heterozygosity (H_o), number of rare alleles (N_r), number of private alleles per subpopulations (AP) and percentage of polymorphic loci (P)	46
Table 4. Genetic distances among four subpopulations of <i>Acrocomia aculeata</i> identified by Structure	46
Table 5. Analysis of molecular variance (AMOVA) of <i>Acrocomia aculeata</i> accessions among the four subpopulations identified by Structure	47

CAPÍTULO 2

Tabela 1. Estimativas de diversidade genética e do índice de fixação de árvores adultas e progênes de <i>Acrocomia aculeata</i> : k é o número de alelos; H_o e H_e são as heterozigosidades observada e esperada; F e F_{of} são os índices de fixação de adultos e progênes; P_{ID} é a probabilidade de não-exclusão para identidade; P_{exc2} é a probabilidade de não-exclusão do segundo parental.....	68
Tabela 2. Estimativas de parâmetros do sistema de reprodução, endogamia e parentesco em <i>Acrocomia aculeata</i>	69
Tabela 3. Estimativas do sistema de reprodução e parâmetros genéticos de famílias maternas de <i>Acrocomia aculeata</i>	70
Tabela 4. Estimativas de depressão por endogamia (δ) para os caracteres vegetativos de famílias maternas de <i>Acrocomia aculeata</i> com progênie de autofecundação identificadas pelo teste de análise de paternidade: altura da planta (AP), diâmetro do estipe (DE), número de folhas (NF), comprimento da folha (CF), largura do ráquis (LR) e número de folíolos (NFA).....	72
Tabela 1S. Valores fenotípicos dos caracteres vegetativos de cada progênie híbrida e de autofecundação de <i>Acrocomia aculeata</i> : altura de planta (AP), diâmetro de estipe (DE), número de folhas (NF) comprimento de folha (CF), largura do ráquis (LR), e número de folíolos na folha avaliada (NFA)	82

APÊNDICES

Apêndice F – Table 1. Quantification and purity of genomic DNA using 200 mg from leaf (lyophilized and fresh) and stipe tissues of <i>Acrocomia aculeata</i>	108
---	-----

RESUMO

LANES, Éder Cristian Malta de, M.S., Universidade Federal de Viçosa, Dezembro de 2014. **Variabilidade molecular e sistema de reprodução de macaúba (*Acrocomia aculeata*)**. Orientador: Sérgio Yoshimitsu Motoike, Co-orientadores: Marcos Deon Vilela de Resende e Eveline Teixeira Caixeta.

A macaúba [*Acrocomia aculeata* (Jacq.) Lodd. ex Martius] uma planta endêmica das Américas tropical e subtropical vem despertando grande interesse da comunidade científica devido às inúmeras potencialidades do fruto, e principalmente pelo seu grande potencial como fonte de energia renovável. Mais do que uma fonte de energia inesgotável, trata-se de um “recurso genético ou fitogenético modelo” por oferecer uma cadeia produtiva efetivamente sustentável. A fim de contribuir para o conhecimento e a domesticação da espécie *A. aculeata*, o estudo foi estruturado em dois capítulos, cuja aplicação de marcadores microssatélites foi destinada a atender aos seguintes objetivos: i) Determinar a estrutura populacional e os níveis de variabilidade genética dos acessos de *A. aculeata* mantidos no Banco de Germoplasma de Macaúba (BAG-Macaúba); e ii) Analisar os parâmetros do sistema de cruzamento, bem como os níveis de depressão por endogamia em caracteres vegetativos de famílias maternas de *A. aculeata* conservados ex situ. Os resultados indicaram uma abundante variação de formas alélicas entre os acessos de *A. aculeata*, possivelmente devido à larga representatividade da coleção. Por meio do software Structure observou-se a presença de quatro pools gênicos com uma ampla variabilidade genética a ser explorada pelo programa de melhoramento genético da espécie. As análises do sistema reprodutivo indicaram que *A. aculeata* é uma espécie alógama ($\hat{t}_m=0,986$), com sistema parcialmente auto-incompatível. As altas taxas de cruzamentos biparentais ($\hat{r}_{p(m)}=0,564$) e entre indivíduos aparentados ($\hat{t}_m - \hat{t}_s=0,383$) foram responsáveis pelos desvios da pressuposição de panmixia, sendo a maior fração das progênies compostas por irmãos-completos. A depressão por endogamia afetou de forma significativa todos os caracteres vegetativos avaliados por este estudo na *A. aculeata*, embora a magnitude depressiva tenha variado amplamente entre as famílias maternas. Os resultados deste estudo vão permitir predições mais realistas sobre tamanhos amostrais para a conservação genética, o uso de modelos matemáticos mais precisos para estimar os diversos parâmetros genéticos e critérios mais eficientes de seleção no programa de melhoramento.

ABSTRACT

LANES, Éder Cristian Malta de, M.S., Universidade Federal de Viçosa, Dezembro de 2014. **Molecular variability and mating system of macaw palm (*Acrocomia aculeata*)**. Orientador: Sérgio Yoshimitsu Motoike, Co-orientadores: Marcos Deon Vilela de Resende e Eveline Teixeira Caixeta.

The macauba [*Acrocomia aculeata* (Jacq.) Lodd. Ex Martius], an endemic plant of the tropical and subtropical American, has been arousing great interest from the scientific community because of the numerous potentialities of fruit, and especially by its great potential as a source of renewable energy. More than a source of inexhaustible energy, it is a "genetic resource or phylogenetic model" by offering a effective sustainable production chain. In order to contribute for knowledge and domestication of *A. aculeata*, the study was divided into two chapters, whose application of microsatellite markers was designed to meet the following objectives: i) to determine the population structure and the level of genetic variability of *A. aculeata* accessions maintained in the Active Germplasm Bank of Macauba (BAG-Macauba); and ii) to analyze the parameters of the mating system, well as levels of inbreeding depression in vegetative characters maternal families of *A. aculeata* conserved in ex situ collections. The results indicated an abundant variation of allelic forms in accession of *A. aculeata*, possibly due to the large representation of the collection. In addition to this, the software Structure indicated the presence of four gene pools with a wide genetic variability to be exploited by the breeding program of the species. The analyzes on reproductive system indicates that the species is predominantly allogamous ($\hat{t}_m=0.986$), with system partially self-incompatible. The high rates of biparental crossing ($\hat{r}_{p(m)}=0.564$) and mating among relatives ($\hat{t}_m - \hat{t}_s=0.383$) were responsible for the deviations on the assumption of panmixy, with the largest fraction of the progenies composed of full siblings. Inbreeding depression affected significantly vegetative characters of *A. aculeata*, although depressive magnitude was diferente between maternal families. The results of this study will allow more realistic predictions on sample sizes for genetic conservation, the use of mathematical models more accurate to estimate the several genetic parameters, and criteria more efficient selection the breeding program of the species.

1. INTRODUÇÃO GERAL

A crescente demanda mundial por energia e a busca por alternativas para substituir os combustíveis derivados do petróleo tem propiciado a busca por fontes limpas e renováveis de energia. Dentre as diversas plantas oleaginosas, a macaúba (*Acrocomia aculeata* (Jacq.) Lodd. ex Martius) é uma das mais promissoras fontes de óleo vegetal das Américas. As estimativas de produtividade da *A. aculeata* estão acima de 6200 kg/ha⁻¹ ano de óleo (Wandeck & Justo, 1988).

A *A. aculeata* é dispersa por quase todo o território nacional (espécie nativa), com alta densidade nas regiões Centro-Oeste e Sudeste do Brasil (Henderson et al. 1995). A produtividade atual da macaúba ainda é baixa, resultado da exploração extrativista, onde a coleta dos frutos é praticada em populações de plantas heterogêneas, de idade, stand e qualidade variável, o que dificulta o seu uso em escala comercial (Motta et al., 2002). Sendo assim, a domesticação da espécie é essencial para se estabelecer a exploração racional em sistemas agrícolas. O desenvolvimento de cultivares de macaúba de produtividade e qualidade comprovada traz segurança e previsibilidade ao investimento para a sua exploração agrícola, garantindo o sucesso de projetos de produção.

O uso de coleções vivas ex situ com o auxílio de ferramentas de biologia molecular podem fornecer informações cruciais para o delineamento de estratégias de melhoramento e conservação genética. O Banco Ativo de Germoplasma de Macaúba (BAG-Macauba) da Rede Macaúba de Pesquisa da Universidade Federal de Viçosa (REMAPE/UFV) destaca-se como um dos maiores bancos de acessos de *A. aculeata* da América Latina e do mundo. O conjunto de acessos desta coleção possui uma ampla variabilidade fenotípica, e constitui um importante recurso para se iniciar os esforços de melhoramento genético da espécie (Manfio, 2010; Manfio et al., 2012).

Portanto, a aplicação de marcadores moleculares nos acessos mantidos na coleção de germoplasma (BAG-Macauba) neste estudo, visa contribuir para o conhecimento e a domesticação da espécie *A. aculeata*. Dessa forma, esse estudo disponibiliza informações sobre a variabilidade genética da coleção ex situ, o sistema de reprodução, estimativas dos parâmetros de cruzamento, coancestria e endogamia. Além disso, a magnitude dos efeitos depressivos da endogamia sobre os caracteres vegetativos foram analisados.

2. REVISÃO DE LITERATURA

2.1. *Acrocomia aculeata* - características gerais e distribuição geográfica

A macaúba, *Acrocomia aculeata* (Jacq.) Lodd. ex Martius, é uma palmeira perenifólia da família Arecaceae, subtribo Bactridinae, e nativa do continente Americano. Esta espécie oleaginosa possui um genoma de 2841 Mb distribuídas em 15 pares de cromossomos ($2n = 30$) (Röser et al., 1997; Abreu et al., 2011). Quando adultas podem ultrapassar os 15m de altura, e sua estipe é geralmente recoberta por espinhos e resquícios dos pecíolos foliares. Os frutos, dispostos em cachos, são drupas com diâmetro variando de 25–60 mm composto de epicarpo fibroso e quebradiço quando maduro, mesocarpo carnoso e fibroso, rico em glicerídeos e ácidos graxos. Protegido por um endocarpo rígido e lignificado, o endosperma sólido possui alto teor de lipídios, proteína, fibra e ácidos graxos insaturados (Scariot & Lleras 1995; Lorenzi et al., 2004; Bora & Rocha 2004; Hiane et al., 2006; Motoike & Kuki 2009; Von Lisingen & Cervi, 2009; Ciconini et al., 2013).

A espécie é comumente encontrada nas regiões tropicais e subtropicais das Américas. No Brasil, *A. aculeata* está dispersa por quase todo o território nacional, ocupando uma área de 11,6 milhões de hectares (Melo, 2012). A espécie é encontrada nas cinco regiões brasileiras, mas com alta densidade no Centro-Oeste e Sudeste (Henderson et al., 1995), povoando regiões de cerrado, transição cerrado – caatinga, floresta estacional semidecidual e floresta tropical úmida (Figura 1).

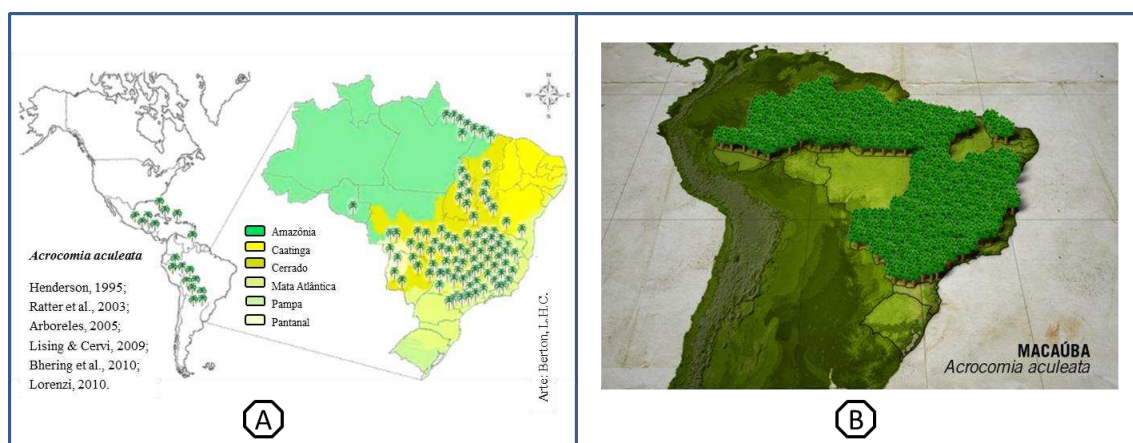


Figura 1. Distribuição geográfica da *Acrocomia aculeata* no Brasil. Mapas (A) elaborado por Berton (2013); e (B) por consultoria de Harri Lorenzi do Instituto Plantarum (Um Pé de Quê?/ Instituto Plantarum - Harri Lorenzi, 2014).

Reconhecidamente existem apenas duas espécies do gênero *Aculeata* (Henderson et al., 1995): a *Acrocomia aculeata* (Jacq.) Lodd. ex Mart, de maior porte e amplamente distribuída nas regiões secas da América Tropical, e a *Acrocomia hassleri* (Barb. Rodr.)

W. J. Hahn, de menor porte e endêmica do centro-oeste brasileiro. Porém, a grande variabilidade morfológica exibida pela espécie *A. aculeata* tem gerado uma grande gama de sinonímia, como *A. sclerocarpa*, *A. intumescens*, *A. media*, *A. totai* e etc (<http://www.kew.org>, <http://www.palmweb.org>, <http://www.splink.org.br>, <http://www.nybg.org/science> e <http://www.tropicos.org>).

O caráter heliófilo da *A. aculeata* a enquadra como pioneira no contexto sucessional florestal, e sua resiliência a permite colonizar habitats com distintas características edafoclimáticas. Essa alta capacidade de ocupação e adaptação para ambientes heterogêneos é acompanhada de notória diferenças morfológicas, diversidade esta que pode ser verificada entre populações, como também dentro dos maciços naturais de *A. aculeata* (Ciconini et al., 2013) (Figura 2). Dentre os fatores ambientais que exercem pressão de seleção, os componentes abióticos, como a temperatura, radiação solar e as condições edáficas estão dentre os elementos que podem promover a variação clinal nas plantas (Saldaña et al., 2007; Haraguchi et al., 2009).



Figura 2. Diversidade morfológica da *Acrocomia aculeata*. Plantas com 4 anos de idade, oriundas dos Estados de Minas Gerais (esquerda) e Pernambuco (direita). Local: Laboratório de Biotecnologia e Melhoramento DFT – Universidade Federal de Viçosa, 2014 (Fonte: Acervo pessoal de Motoike, S.Y.).

2.1.1. Aspectos reprodutivo da *Acrocomia aculeata*

A *A. aculeata* é uma arbórea monoica, onde as inflorescências são compostas por flores amareladas unissexuais reunidas em espigas e protegidas por uma espata de até 2m de comprimento. As numerosas flores masculinas e femininas estão dispostas em partes distintas da mesma inflorescência (ou seja, andrógenas). A inflorescência

apresenta uma marcada protoginia, em que as flores femininas entram em antese antes das flores masculinas (Scariot et al., 1991).

Os principais agentes polinizadores da *A. aculeata* são coleópteros, enquanto que o vento exerce papel secundário na polinização (Scariot et al., 1991). As espécies de coleópteros polinizadores da *A. aculeata* variam entre as diferentes regiões. Em populações naturais do Distrito Federal (Bioma Cerrado), Scariot et al. (1991) relataram as espécies *Andranthobius* sp. (Curculionidae), *Mystrops* cf *mexicana* (Nitidulidae) e *Cyclocephala forsteri* (Scarabaeidae); enquanto que, nas populações naturais do município de Acaiaca, na região da Zona da Mata de Minas Gerais (Bioma Mata Atlântica), Brito (2013) descreveu as espécies *Andranthobius* aff. *Bondari* (Curculionidae), *Phyllotrox tatianae* (Curculionidae), *Mystrops dalmasi* (Nitidulidae) e *Mystrops debilis* (Nitidulidae), como os principais polinizadores efetivos da espécie.

2.1.2. Potencialidades do fruto da *Acrocomia aculeata*

As estimativas de produtividade da *A. aculeata* estão acima dos 6200 kg/ha⁻¹ano de óleo (Wandeck & Justo, 1988). Porém, o estudo realizado por Berton (2013) indica um potencial de produção de até 13000 kg/ha⁻¹ ano de óleo a partir do genótipo mais produtivo selecionado de populações naturais do estado de São Paulo (ver Figura 3).

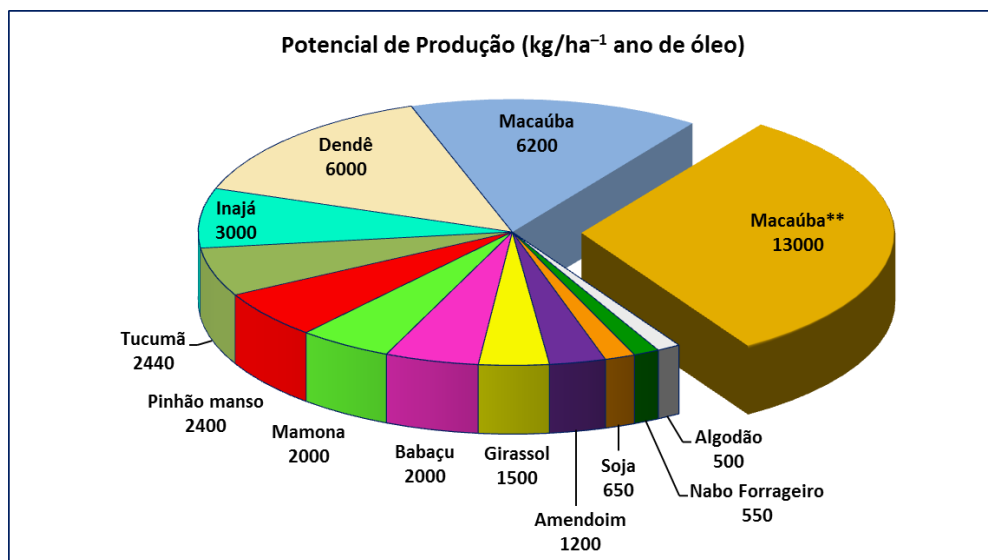


Figura 3. Análise comparativa do potencial de produção de óleo vegetal por ha.ano⁻¹ entre as principais oleaginosas. Macaúba**: Trata-se do potencial de produção de óleo do mesocarpo por ha.ano⁻¹ do melhor genótipo selecionado em condições naturais (adaptado de Berton, 2013).

A rentabilidade desta espécie reverbera especialmente pela multitude de usos de seus frutos. O óleo extraído do mesocarpo, com alto teor de ácido oleico, é comparável ao azeite de oliva, enquanto que na amêndoa predomina o ácido linoleico, atendendo o setor de fármaco-cosméticos (Hiane et al., 2005; Bora & Rocha 2004; Ciconini et al.,

2013; Berton et al., 2014). Contudo, a predominância de ácidos graxos insaturados no óleo do mesocarpo (Bora & Rocha 2004) o torna uma matéria prima de excelente qualidade para a produção de biodiesel (Kusdiana & Saka, 2001). Os resíduos sólidos, oriundos da extração dos óleos podem ter diferentes destinos: (i) o endocarpo lignificado pode ser utilizado como fonte de energia calorífica na forma de carvão vegetal (Silva et al., 1986); (ii) a biomassa do mesocarpo e do exocarpo pode ser utilizada tanto na produção de etanol de segunda geração (Gonçalves et al., 2013), como na cadeia produtiva de bioquerosene para a aviação (Lanes et al., 2014a); (iii) as tortas fibro - protéicas do mesocarpo e endosperma podem ser empregadas na nutrição humana ou animal, além de também mostrarem eficiência na remoção de resíduos de corante de efluentes químicos industriais (Vieira et al., 2012). Como observado, a destinação destes co-produtos da macaúba atende a diversos nichos de mercado, permitindo agregar valor a produção (Figura 4).

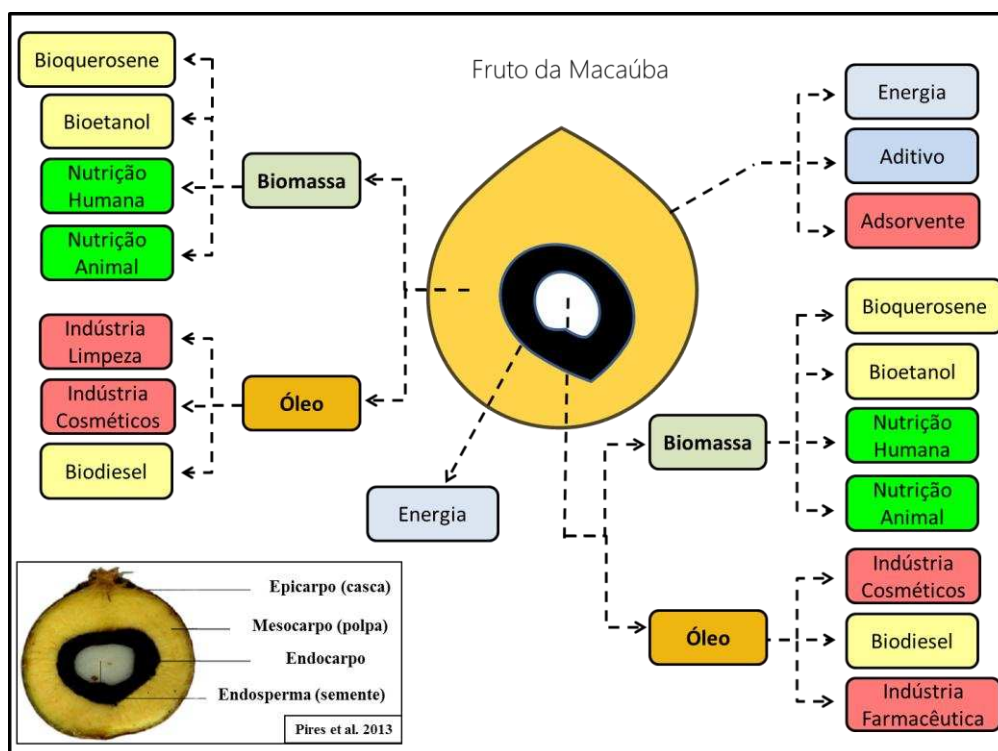


Figura 4. Esquema de aproveitamento das partes do fruto da macaúba e as respectivas aplicações dos co-produtos após o processamento (adaptado de Melo, 2012).

2.1.3. Sustentabilidade da cadeia produtiva da *Acrocomia aculeata*

O tema sustentabilidade é algo imprescindível nos dias atuais, não sendo mais permitido se falar em desenvolvimento econômico sem que este esteja atrelado com as questões ambientais. O planeta possivelmente chegou ao seu estágio crítico, onde qualquer modelo que não envolva uma cadeia produtiva efetivamente sustentável pode

representam sérios riscos para as presentes e futuras gerações. Fatos recentes ocorridos no Brasil como a intensa seca na região Sudeste e tornados registrados na região Centro-Oeste (Mato Grosso do Sul e Brasília) podem estar relacionados com as mudanças climáticas provocadas pela intervenção humana.

O grande desafio atual é como conciliar a expansão de área agrícola para a produção de alimentos, plantas oleaginosas para geração de energia e ao mesmo tempo promover a preservação e/ou restabelecimento de áreas florestais (Berton et al., 2013). Neste contexto, a *A. aculeata* se destaca de outras oleaginosas, como por exemplo, o dendê, por ser uma palmeira rústica, tolerante à escassez hídrica, além de adaptar-se a um amplo gradiente de temperatura, luminosidade e tipos de solos (Arkcoll 1990; Motoike & Kuki, 2009). É uma espécie que pode ser integrada em diversos sistemas de produção, até mesmo no modelo agrossilvipastoril (Ciconini et al., 2013). A inserção desta espécie apenas em pastagens já existentes pode ajudar a conter a expansão de novas áreas agrícolas, protegendo as florestas, sequestrando carbono (CO₂), enriquecendo o solo com matéria orgânica e restaurando a biodiversidade de áreas degradadas (Lanes et al., 2014a).

De acordo com Melo (2012), se apenas 1% do território brasileiro (~ 8,5 milhões de hectares) for cultivado com a *A. aculeata* pode-se produzir:

- 34 bilhões de litros de óleos vegetais para a produção de biodiesel e material prima para a indústria química;
- 40 milhões de toneladas de tortas, matéria prima para rações para a pecuária;
- 3,9 milhões de toneladas de farelo de castanha de alto teor protéico para consumo humano;
- 12,3 milhões de toneladas de carvão;
- Geração de renda para 850.000 famílias;

Além do mais, o estoque de carbono em uma área de regeneração natural de *A. aculeata* corresponde ao estoque de CO₂ equivalente de 12,412 MgCO₂ eq.ha⁻¹, que convertidos em créditos de carbono podem gerar uma renda bruta de R\$ 3.682,14 . ha⁻¹.ano⁻¹ (Tolêdo, 2010). É importante ressaltar que, de acordo com o último Censo Agropecuário Brasileiro em 2006 (IBGE, 2006), foram registrados uma área total de pastagens (naturais e plantadas) no Brasil de 172,3 milhões de hectares. Sendo assim, o cultivo desta espécie apenas nestas áreas, exerceria um inestimável serviço ambiental, social e econômico. Para a viabilidade deste cultivo, entretanto, é necessário estudos de

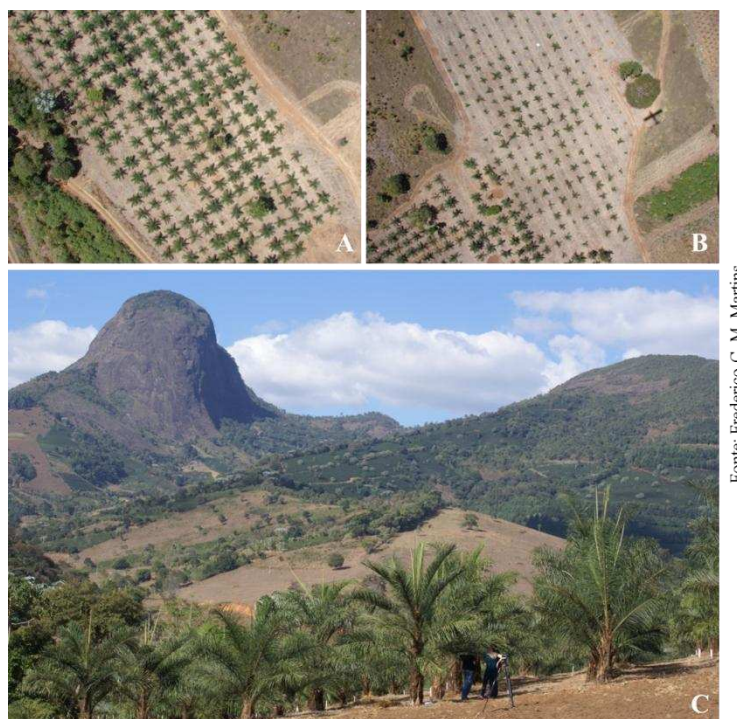
base genética que permitam o surgimento de cultivares produtivas e adaptadas a condições locais.

2.1.4. Banco ativo de germoplasma de *Acrocomia aculeata*

A intensa busca pelos melhores genótipos (seleção dos alelos favoráveis) como forma de se obter culturas mais uniformes tem conduzido uma redução na base genética (erosão gênica) de inúmeras espécies vegetais. A erosão gênica provocada pelo melhoramento genético restringe a disponibilidade de genes para o melhoramento futuro (Barbieri, 2003). Fato que é agravado pela eliminação das espécies selvagens.

Como forma de contornar essas questões, os Bancos de Germoplasma (ou banco de alelos), unidades conservadoras de genótipos representativos da variabilidade genética de uma determinada espécie, tem como objetivo suprir a necessidade de novos genes (ou novas formas alélicas) para incrementar a base genética por intermédio do melhoramento genético. A conservação genética *ex-situ* (conservação fora do seu local nativo), tem sido a primeira etapa para dar suporte a um programa de melhoramento e de conservação genética de uma espécie.

O Banco Ativo de Germoplasma de Macaúba (BAG-Macaúba), Registro nº 084 - 2013/CGEN/MMA, da Rede Macaúba de Pesquisa da Universidade Federal de Viçosa (REMAPE/UFV), destaca-se como um dos maiores bancos de acessos de *A. aculeata* da América Latina e do mundo. Implantado em 2009 na Fazenda Experimental de Araponga/MG (20°40'1"S, 42°31'15"W), a coleção possui atualmente 253 famílias maternas, composto aproximadamente por 1500 acessos representativos de quase todas as regiões brasileiras (Figura 5) (Silveira et al., 2013). O conjunto de acessos da coleção viva da BAG-Macaúba possui uma ampla variabilidade fenotípica, e constitui um importante recurso para se iniciar os esforços de melhoramento genético, e consequentemente, a domesticação da espécie (Manfio, 2010; Manfio et al., 2012), como também, delinear estratégias de conservação das plantas nativas da *A. aculeata* na região sudeste do Brasil.



Fonte: Frederico C. M. Martins

Figura 5. Banco Ativo de Germoplasma de Macaúba (BAG-Macaúba) localizada na cidade de Araponga - MG. (A) Imagem aérea da Gleba I com 54 famílias maternas plantadas em fevereiro de 2009; (B) Imagem aérea da Gleba II com 34 famílias maternas plantadas em novembro de 2009; (C) Coleção viva conservada ex situ.

2.2. Biologia molecular no melhoramento e conservação genética

2.2.1 Marcadores Moleculares e suas aplicações

A biologia molecular tem sido uma ferramenta cada vez mais frequente no auxílio de programas de conservação, caracterização e uso de germoplasma (Faleiro, 2011). Os marcadores moleculares oferecem uma alternativa estável e confiável para a caracterização de bancos de germoplasma fornecendo uma medida mais direta da diversidade genética entre os acessos pelo polimorfismo genético ao nível do DNA (Aradhya et al., 2010).

O uso de marcadores moleculares para fins de caracterização de coleções de germoplasma tem permitido determinar o modo como são constituídas as recombinações gênicas a cada evento reprodutivo (Karasawa et al., 2007), os padrões de dispersão do pólen (Silva et al., 2011), conhecer a estrutura e diversidade genética das populações (Lanes et al., 2014b), estabelecer as relações genéticas entre acessos (Stajner et al., 2008; Bracco et al., 2009) e formar coleções nucleares (Odong et al., 2011). As informações ao nível de DNA também tem possibilitado o entendimento do processo histórico de domesticação, centro de diversidade e origem de várias espécies, incluindo

palmáceas como coco e o dendê (Corley & Tinker, 2003; Maizura et al., 2006; Gunn et al., 2011).

Em pesquisas realizadas recentemente na espécie *A. Aculeata* foram aplicados marcadores microssatélites, também conhecidos como Simple Sequence Repeats (SSR) e os Random Amplification of Polymorphic DNA (RAPD) para avaliar a diversidade genética, sistema de cruzamento, estrutura e diferenciação genética de populações naturais (Nucci et al., 2008; Oliveira et al., 2012; Abreu et al., 2012).

2.2.2 Random Amplification of Polymorphic DNA (RAPD)

A técnica conhecida como Random Amplification of Polymorphic DNA (RAPD; Williams et al., 1990), ou Arbitrarily Primed – Polymerase Chain Reaction (AP-PCR; Welsh & McClelland, 1990) ou DNA Amplification Fingerprinting (DAF; Caetano-Anollés et al., 1991) consiste na amplificação de DNA genômico em Polymerase Chain Reaction (PCR) utilizando um único primer de sequência arbitrária, normalmente com 10 nucleotídeos (Lacerda et al., 2002). O primer se liga à segmentos complementares em fitas opostas do DNA alvo entre uma distancia de no máximo 3 a 4 mil pares de bases, amplificando in vitro o segmento de DNA entre dois primers adjacentes com o auxílio da enzima Taq DNA polimerase (Ferreira & Grattapaglia, 1998; Caixeta, 2009). Dentre as vantagens da técnica RAPD destacam-se a simplicidade de execução, obtenção de dados em curto espaço de tempo, custo reduzido e uso de primers universais. Apesar das limitações da técnica, as mais importantes referem-se à característica dominante do marcador RAPD (impossibilidade da distinção de indivíduos heterozigotos) e a baixa reprodutibilidade das informações genéticas obtidas (Lacerda et al., 2002; Faleiro, 2007; Caixeta, 2009).

2.2.3 Simple Sequence Repeats (SSR)

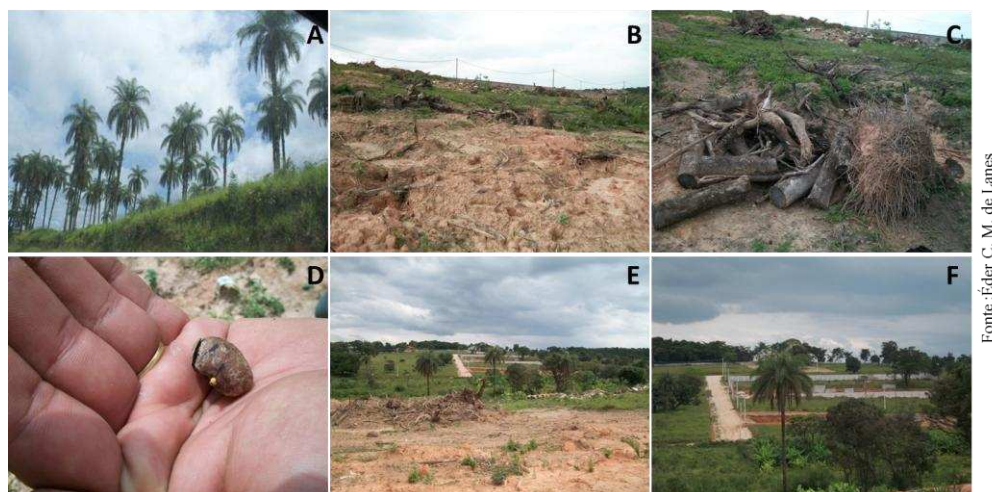
Os microssatélites ou Simple Sequence Repeats (SSR) ou STR (Short Tandem Repeats) são regiões de repetição de pares de bases do DNA, constituídos por 1 a 6 nucleotídeos repetidos em tandem. Regiões contendo sequências simples repetidas são amplificadas, via PCR, utilizando-se um par de primers específicos (de 20 a 30 bases) complementares as sequências únicas que flanqueiam os SSR em questão (Ferreira & Grattapaglia, 1998; Faleiro, 2007). O polimorfismo em um loco SSR é resultado do número de unidades repetitivas dentro da estrutura do microssatélite, por exemplo, a quantidade que um dinucleotídeo (AG)_n se repete para um determinado loco. Dentre as principais vantagens, estes se caracterizam por serem codominantes, multialélico,

elevado conteúdo de informação de polimorfismo (PIC), aparentemente distribuído por todo o genoma, alta reprodutibilidade das marcas e baseados na especificidade e rapidez da PCR. Quanto às limitações, destacam-se o alto custo e mão-de-obra requerido no desenvolvimento dos primers SSR, levando em consideração todas as etapas do seu desenvolvimento, incluindo seleção de biblioteca genômica, seleção de clones positivos, desenho e teste dos primers (Buso et al., 2003; Faleiro, 2007; Caixeta, 2009).

2.3. Análise e importância do sistema de reprodução

O sistema de cruzamento desempenha um papel primordial na composição genética das populações, uma vez que determina os padrões pelos quais os gametas são transmitidos para formar a próxima geração. O modo como as recombinações gênicas são constituídas pelos indivíduos influencia na extensão e distribuição da variabilidade genética dentro e entre populações, incluindo os seus níveis de endogamia. Sendo assim, o sistema reprodutivo não somente determina a associação entre as progênes e o grau de parentesco entre os descendentes dentro destas progênes, como também afeta os parâmetros genéticos populacionais (Gaiotto et al., 2003; Silva et al., 2011). De acordo com Loveless & Hamrick (1984), o sistema reprodutivo pode ser considerado como o fator mais influente na estrutura genética das populações.

O padrão ou mesmo as variações do sistema de cruzamento dentro de uma espécie é determinada por fatores reprodutivos e ecológicos, como por exemplo, o comportamento dos polinizadores (Núñez-Avellaneda et al., 2005), fenologia do florescimento (Khanduri et al., 2013) e a densidade populacional (Murawski & Hamrick, 1991). Os fatores ecológicos são possivelmente os mais afetados pela ação antrópica, o que conseqüentemente, altera os padrões do sistema de acasalamento de uma espécie. A alteração e/ou perda de habitat resultam em isolamento de populações, diminuição do número efetivo de árvores de uma população, diminuição do número de polinizadores e ausência de dispersores de sementes (Laurance et al. 2002; Laurance et al. 2003; Portela & Santos 2014), representado um sério risco a diversas espécies pertencentes à família Arecaceae (Nazareno et al., 2014; Portela & Santos, 2014), inclusive a *A. aculeata*, conforme demonstrando pela a Figura 6.



Fonte: Éder C. M. de Lanes

Figura 6. Degradação do habitat natural da *Acrocomia aculeata* no município de Betim – MG. (A) População nativa; (B e C) Perda do habitat natural e de indivíduos da espécie pela ação antrópica; (D) Semente em processo de germinação natural presente no ambiente sob fragmentação; (E e F) Loteamento imobiliário da área fragmentada.

Breed et al. (2012) descreveram possíveis respostas do sistema de cruzamento após a fragmentação do habitat, incluindo: (a) insuficiência de pólen/polinizadores; (b) aumento da autofecundação; (c) aumento de cruzamentos endógamos; e (d) diminuição da diversidade de pólen (Figura 7).

Pré-perturbação do habitat	Padrão de respostas do acasalamento após a perturbação do habitat	Efeito sobre o desempenho	Consequências e principais observações
 Equilíbrio (marco-zero)	Insuficiência de pólen/polinizadores  (a)	Sem fertilização por causa da deficiência de pólen	➤ Nenhuma reprodução se realiza a não ser nas cleistogâmicas
	Mais autofecundação  (b)	Alelos deletérios recessivos causando depressão por endogamia	➤ Aumento da taxa de autofecundação ➤ Limitado pela auto-incompatibilidade, cruzamento preferencial e diócia
	Mais pólen de aparentados  (c)	Alelos deletérios recessivos causando depressão por endogamia	➤ Aumento da reprodução entre pais aparentados ➤ Possível para espécies autoincompatíveis e dióicas
	Redução da diversidade de pólen  (d)	Baixa diversidade de pólen; resulta em uma qualidade inferior da descendência	➤ Aumento da correlação de paternidade ➤ Possível para espécies autoincompatíveis e dióicas

Figura 7. Possíveis respostas do sistema de acasalamento após a perturbação do habitat. Árvores geneticamente semelhantes são representadas por cores idênticas. A falha na polinização é indicada por linhas pontilhadas coloridas seguidas pelo símbolo (X); A redução na

polinização é indicada por linhas pontilhadas coloridas; e a polinização normal, por linhas de cor sólida; um aumento relativo da polinização é indicado por linhas coloridas com um símbolo (+). As respostas (b, aumento da autofecundação), (c, aumento de cruzamentos endogâmicos) e (d, diminuição da diversidade de pólen) não são mutuamente exclusivos para espécies hermafroditas e autocompatíveis (adaptado de Breed et al., 2012).

2.3.1. Software Multilocus Mating System Program - MLTR

Os modelos de cruzamento misto e correlacionado (Ritland & Jain, 1981; Ritland, 1989) implementado pelo software Multilocus Mating System Program - MLTR (Ritland, 2002) é um dos mais citados na literatura para inferir sobre o sistema de cruzamento de uma espécie. As estimativas são obtidas por máxima verossimilhança, e admite que as populações se reproduzam por autofecundação a uma taxa s e por cruzamentos aleatórios a uma taxa t . As pressuposições básicas do modelo são: (i) o conjunto de pólen é homogêneo para os cruzamentos com todos os genótipos maternos; (ii) que os alelos de diferentes locos segregam independentemente; e (iii) que os locos não são afetados pela seleção ou mutação entre o evento reprodutivo e a análise (Ritland & Jain, 1981). O MLTR fornece estimativas dos seguintes parâmetros de cruzamento: índice de fixação das árvores maternas (F_m); taxa de cruzamento multiloco (t_m); taxa de autofecundação ($s=1-t_m$); taxa de cruzamento uniloco (t_s); taxa de cruzamento entre indivíduos aparentados (t_m-t_s); correlação multiloco de paternidade (cruzamento biparental) ($r_{p(m)}$); correlação de paternidade uniloco ($r_{p(s)}$).

2.4. Genética de populações

A genética de populações visa investigar a dinâmica dos genes em populações biológicas, e a elucidação dos mecanismos que alteram a composição gênica ou apenas a frequência genotípica. Para prever e investigar a ocorrência das mudanças das frequências gênicas e alélicas, os geneticistas populacionais fazem uso de modelos matemáticos, tendo como base o entendimento da genética molecular, as leis de Mendel sobre a herança e a teoria moderna de evolução. O foco da genética de população é estudar o fenômeno hereditário a nível populacional (Hartl & Clark, 2007; Beiguelman, 2008).

2.4.1. Estrutura genética de população

A estrutura genética pode ser definida como a distribuição da variabilidade genética entre e dentro de populações, ou mesmo, como a distribuição não aleatória de alelos ou de genótipos no espaço e no tempo (Loveless & Hamrick, 1984). A estrutura genética populacional é resultado da ação e das interações de uma série de mecanismos

evolutivos e ecológicos, promovendo tanto a diferenciação como a homogeneização genética em uma população natural (Hamrick, 1989). Essas forças violam as condições de equilíbrio de Hardy-Weinberg em favor da evolução. As forças evolucionárias como a mutação, seleção natural, migração (sementes), fluxo gênico (pólen) e deriva genética são responsáveis por promoverem alteração no conjunto gênico de uma população (Hamrick, 1989; Finkeldey, 1998). Detalhes sobre os fatores que afetam a estrutura genética de populações e seus efeitos podem ser encontrados na revisão de Loveless & Hamrick (1984). As análises de estrutura genética de populações através de marcadores codominantes podem ser realizada pela estatística F (Wright, 1951), diversidade gênica (Nei, 1973), e pelo coeficiente de coancestria (Cockerhan, 1969). Além destes, a análise de estrutura de população também vem sendo realizada por métodos estatísticos Bayesianos, principalmente pelas abordagens implementadas pelo software Structure (Pritchard et al., 2000; Falush et al., 2003).

Ao analisar a estrutura genética de onze populações naturais de *A. aculeata* do estado de São Paulo e Minas Gerais, Abreu et al. (2012) relataram a presença de 2 à 3 pools gênicos com base na inferência Bayesiana por meio de locos microssatélites. Quanto ao estudo de estrutura populacional da *A. aculeata* a partir de marcadores RAPD, Oliveira et al. (2012) identificaram níveis elevados de variabilidade genética dentro das populações da região norte de Minas Gerais, utilizando-se de medidas de diversidade genética intrapopulacional e do coeficiente de diversidade relativa entre grupos (G_{ST}) proposta por Nei (1978) para a diferenciação genética. É importante ressaltar que essa medida, matematicamente, equivale à estatística F_{ST} de Wright.

2.4.2. Estatística F (Wright, 1951)

A estatística F (Wright 1951) pode ser definida como a correlação entre dois alelos que se unem na formação de um zigoto (indivíduos). Essa estatística é utilizada para quantificar o efeito do endocruzamento na subdivisão populacional (Hartl & Clark, 2007). Os valores do índice de fixação gênica que variam de -1 a $+1$, e é um indicador útil de diferenciação genética, fornecendo níveis médios de fixação dos alelos para o total de populações, dentro das populações e a divergência genética entre as populações, com base em medidas de probabilidade de identidade por descendência (Wright, 1965; Nei, 1977). Quando o índice de fixação difere de zero, significa que está ocorrendo desvio do EHW e, neste caso, a população pode atingir outro estado de equilíbrio, denominado Equilíbrio de Wright (Equilíbrio com Endogamia). Considerando a

população como um todo (H_T), suas subpopulações (H_S) e seus indivíduos (H_I), temos os seguintes parâmetros:

F_{IT} = índice de fixação ou coeficiente de endogamia para o conjunto das populações (devido ao sistema reprodutivo e subdivisão),

$$F_{IT} = H_T - H_I / H_T$$

F_{IS} = índice de fixação ou coeficiente de endogamia intrapopulacional (devido ao sistema reprodutivo),

$$F_{IS} = H_S - H_I / H_S$$

F_{ST} = índice de fixação ou divergência genética entre populações (devido á subdivisão populacional).

$$F_{ST} = H_T - H_S / H_T$$

Segundo Wright (1951), os três parâmetros se relacionam por: $F_{ST} = (F_{IT} - F_{IS}) / (1 - F_{IS})$

2.4.3. Coeficientes de coancestria (Cockerhan, 1969)

O coeficiente de coancestria (θ) de Cockerhan (1969) é uma medida de diferenciação entre populações e é consequência da correlação entre genes de diferentes indivíduos na mesma população (Weir & Cockerhan, 1984). O estimador θ permite investigar o padrão de variabilidade das frequências alélicas em diferentes níveis de hierarquia, com genes estratificados em nível de indivíduo, dentro de família e dentro de populações (Dias, 2006). Este estimador permite ainda, obter estimativas dos níveis de endogamia a partir de uma base não viesada. Considerando indivíduos diplóides, amostrados a partir de várias populações, as três medidas propostas por Cockerhan (1969) correspondem às estatísticas F de Wright da seguinte forma:

F (F_{IT} de Wright): é a correlação entre alelos dentro de indivíduos em todas as populações, representando o coeficiente de endogamia dos indivíduos em relação ao conjunto de populações;

θ (F_{ST} de Wright): é a correlação de todos os alelos de diferentes indivíduos na mesma população. É o coeficiente de coancestralidade e mede a divergência entre populações;

f (F_{IS} de Wright): é a correlação dos alelos dentro de indivíduos dentro da população a qual estes indivíduos pertencem, ou seja, representa grau de endogamia dentro de populações e pode ser expresso como $f = (F - \theta) / (1 - \theta)$.

2.4.4. Software Structure

O software Structure (Pritchard et al., 2000; Falush et al., 2003) é baseado nas pressuposições que cada população esteja em equilíbrio de Hardy-Weinberg e que dentro de cada população os locos estejam em equilíbrio de ligação. De maneira mais simples e aplicada, estima-se a frequência alélica das populações (conjunto de genótipos multilocos dos indivíduos amostrados). Dado essas estimativas, por meio do algoritmo Bayesiano do modelo, atribui-se probabilisticamente os indivíduos a K populações. Os indivíduos da amostra analisada são, então, distribuídos probabilisticamente para cada K, ou são atribuídos para mais de um K, caso os genótipos indiquem que sejam híbridos, ou seja, indivíduos que herdaram alguma fração do genoma de mais de uma população. Dentre as diferentes propostas na literatura para se determinar o melhor valor de K, destaca-se a metodologia de Evanno et al. (2005), o qual sugerem inferir K a partir de uma estatística ad hoc ΔK .

2.4.5. Tamanho efetivo populacional

O tamanho efetivo de população (N_e) representa o tamanho de uma população ideal onde a redução das variâncias nas frequências alélicas (variabilidade genética), seja por endogamia ou por deriva genética, ocorre de maneira similar a aquela observada em uma população de referência, panmítica, de tamanho finito N (Crow & Kimura, 1970). O tamanho efetivo depende do número de indivíduos que efetivamente estão envolvidos no processo reprodutivo e de suas contribuições relativa para a geração seguinte (Crossa & Vencovsky, 1994). Segundo Wright (1931), o tamanho ideal de uma população é aquela em que cada indivíduo contribui igualmente para o conjunto gamético. Portanto, o tamanho efetivo populacional relaciona-se com a representatividade das populações (Vencovsky, 1987).

A manutenção da estrutura genética de uma população ao longo das gerações é diretamente influenciada pelo tamanho efetivo populacional. Neste contexto, o N_e é de grande importância para estabelecer uma população mínima viável para conservação in situ de uma espécie. Além deste, o conhecimento do N_e permite delinear a melhor estratégia de conservação ex situ, uma vez que, a coleta de sementes para uma coleção de germoplasma poderá ser realizada de maneira a se capturar o máximo de representatividade genética (Sebbenn, 2001; Hardy et al., 2006). Segundo Namoff et al. (2010), a conservação genética ex situ de um total de 15 indivíduos coletados de três árvores matrizes (famílias de meios-irmãos) da palmeira *Leucothrinax morrisii*, assegura a captura de mais de 80% da diversidade alélica da população original (Figura

8). Conforme Namoff et al. (2010), os maiores ganhos em porcentagem média de captura alélica é obtido pelo o aumento de 1-2 acessos (captura alélica >10%), quando comparado com o aumento de 10-11 acessos (captura alélica <1%, Figura 8).

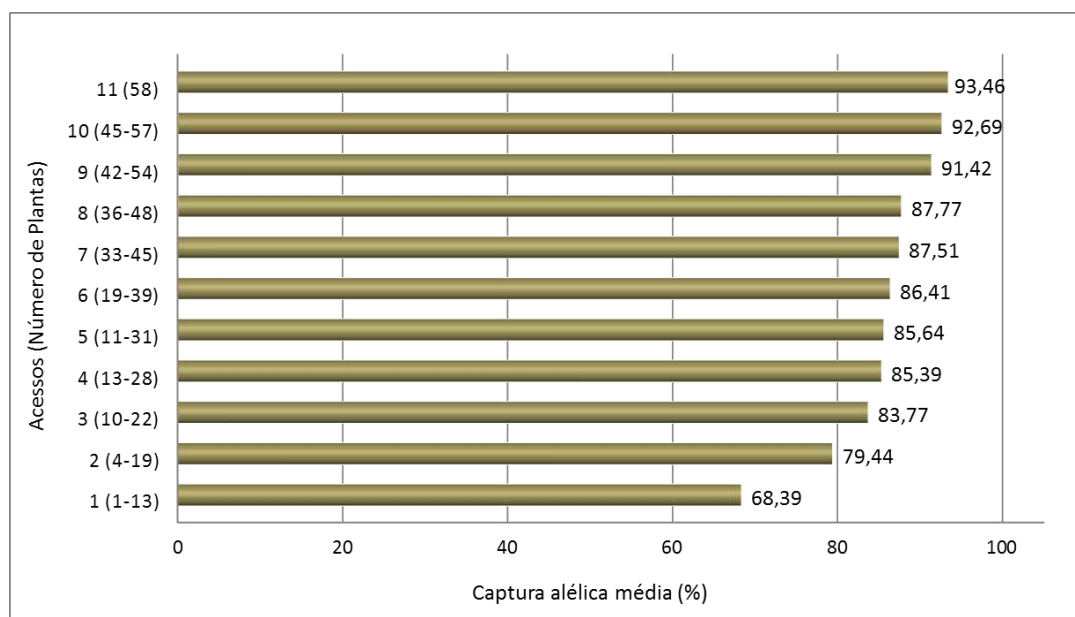


Figura 8. Porcentagem média da captura alélica por número de acessos na coleção ex situ da palma *Leucothrinax morrisii* em relação a população original (adaptado de Namoff et al., 2010).

2.4.6. Depressão por endogamia

A depressão por endogamia é caracterizada pelo declínio do desempenho dos descendentes endogâmicos comparado às progênies de cruzamento. A magnitude depressiva da endogamia atua de forma diferente entre as características, variando conforme o sistema de cruzamento, a fase do ciclo de vida de uma espécie, e particularmente pelas informações implícitas no controle genético do carácter; como, o grau médio de dominância e o número e tipo de alelos recessivos deletérios que estão expostos no estado homozigoto (Lande & Schemske, 1985; Charlesworth et al., 1990; Husband & Schemske, 1996; Thiele et al., 2010).

Para fins de melhoramento genético, informações a respeito da base genética da depressão por endogamia de caracteres quantitativos ampliam a eficiência do processo seletivo de genótipos superiores. Caracteres que não estão sob o controle da sobredominância ou dominância completa podem ser usados como critério de seleção dentro das populações segregantes. Além disso, estes caracteres quando correlacionados com componentes de produtividade, representam menores custos e tempo no processo seletivo de plantas perenes, uma vez que permite a seleção indireta ainda nos primeiros anos da fase juvenil. Em palmáceas, níveis expressivos de depressão por endogamia em caracteres vegetativos e componentes de produção foram registrados em espécies como

Elaeis guineense e em *Cocos nucifera* L. (Corley & Tinker, 2003; Luyindula et al., 2005; Pandin, 2009).

2.4.7. Diversidade gênica (Nei, 1972)

A distância genética (D) de Nei (1972) é uma metodologia alternativa para a análise de populações subdivididas que não requer o conhecimento das frequências genotípicas, podendo ser estimadas a partir das frequências alélicas em termos de heterozigosidade. Esta metodologia fornece a proporção da variabilidade genética contida entre e dentro das populações e baseia-se nas identidades e diversidades gênicas. A estatística (D) pressupõe uma medida de identidade genética (I), expressa pela probabilidade de que dado alelo de um loco, tomado ao acaso em duas diferentes populações, seja idêntico em relação à probabilidade de que dois alelos do mesmo loco, tomados também ao acaso em cada população, sejam também idênticos. A probabilidade J_{PQ} de que dois alelos tomados ao acaso nas populações P e Q sejam idênticos é $J_{PQ} = \sum p_i q_i$. De modo análogo, J_P e J_Q são as probabilidades de que dois alelos tomados ao acaso dentro de cada uma das populações P e Q, respectivamente, sejam idênticos. Nesse caso, $J_P = \sum p_i^2$ e $J_Q = \sum q_i^2$ e, a identidade genética (I) é expressa da seguinte forma (Dias, 2006):

$$I = J_{PQ} / \sqrt{J_P J_Q}$$

p_i e q_i são as frequências do alelo i nas populações P e Q

A distância genética (D) foi obtida pelo cologaritmo neperiano de I, calculada como:

$$D = -\ln(I)$$

2.4.8. Diversidade genética intrapopulacional

A caracterização da variabilidade genética dentro de populações (ou subpopulações) pode ser efetuada a partir de medidas de diversidade genética com base na variação alélica em locos estruturais (Weir, 1996), tais como a frequência alélica, a porcentagem de locos polimórficos, o número médio de alelos por loco, o número efetivo de alelos, o conteúdo de informação polimórfica, a heterozigosidade média observada e a heterozigosidade média esperada segundo as expectativas do EHW.

2.5.8.1. Frequência alélica

A frequência alélica ou gênica é a frequência de um alelo particular em uma população, este pode ser estimado contando o número de vezes que cada alelo é

encontrado, e dividindo-se pelo número total de alelos no loco: $p_{ij} = n_{ij}/n_j$ sendo, p_{ij} a frequência do alelo i na população j ; n_{ij} é o número de ocorrência do alelo i na população j , e n_j é o número total de alelos amostrados na população j .

2.5.8.2. Porcentagem de locos polimórficos (P)

Um loco é considerado como polimórfico, segundo Cole (2003), quando este atende um dos três critérios mais comumente utilizados: (i) loco exibindo polimorfismo em pelo menos um indivíduo da amostra; (ii) loco em que o alelo mais comum possui frequência menor que 99%; e (iii) loco em que o alelo mais comum possui frequência menor que 95%. Desta forma, o P é obtido pela divisão do número de locos polimórfico pelo número de locos analisados.

2.5.8.3. Número médio de alelos por loco polimórfico (N_a)

O N_a fornece uma ideia da distribuição dos alelos na diferentes populações analisadas. De acordo com Brown & Weir (1983), trata-se de uma medida que caracteriza um componente da diversidade, que é a riqueza alélica. O número médio de alelos por loco em cada população é obtido pela razão entre o número total de alelos (A) e o número de locos (L) analisados ($N_a = A/L$).

2.5.8.4. Número efetivo de alelos (N_e)

O número efetivo de alelos (N_e) mede a igualdade das frequências alélicas, ou seja, a uniformidade das frequências alélicas mais comuns em determinado loco, uma vez que, alelos raros contribuem pouco na soma. De forma ainda mais simplificada, trata-se de uma medida que quantifica o número de alelos com frequência significativa dentro da população. O N_e foi introduzido por Kimura & Crown (1964) e é definido como o recíproco da homozigosidade esperada em cada loco por $A_e = 1/\sum p_i^2$, em que p_i é igual às frequências alélicas de cada loco i , calculando-se em seguida a média aritmética para todos os locos.

2.5.8.5. Conteúdo de informação polimórfica (PIC)

O conteúdo de informação polimórfica pode ser definido como um indicador de qualidade do marcador, ou seja, fornece uma estimativa do poder discriminatório do marcador molecular. Este estimador considera não somente o número de alelos por loco, mas também a frequência relativa desses alelos (Cruz et al., 2011).

$$PIC = 1 - \sum_{i=1}^k p_i^2 - \sum_{i,j=1}^k \sum_{(i \neq j)} p_i^2 p_j^2$$

Botstein et al. (1980)

Marcadores com valores de PIC superiores a 0,50 são considerados muito informativos, valores entre 0,25 e 0,50 mediantemente informativos e valores inferiores a 0,25 pouco informativos (Botstein et al., 1980).

2.5.8.6. Heterozigosidade

A heterozigosidade possivelmente é uma das mais importantes medidas de variabilidade genética de uma população, pois sumariza a variação genética de uma população em uma única estatística (Berg & Hamrick, 1997). A heterozigosidade individual descreve a proporção de locos heterozigotos em um indivíduo, e a heterozigosidade média, reflete a proporção de indivíduos heterozigotos numa população medida por vários locos. A heterozigosidade média pode ser determinada de duas formas:

- (a) Heterozigosidade média observada (H_o), definida como a proporção média de indivíduos heterozigotos detectados na população de estudo.
- (b) Heterozigosidade média esperada (H_e), também designada por Nei (1987) como diversidade gênica (gene diversity) é definida como a proporção média de heterozigotos por loco em uma população de cruzamentos aleatórios, sendo igual à proporção esperada em Equilíbrio de Hardy-Weinberg (HWE) de locos heterozigóticos em um indivíduo escolhido ao acaso. Ou mesmo, como a probabilidade de dois alelos aleatoriamente amostrados numa população infinita, diferirem entre si. A H_e para um dado loco é definida por:

$$H_e = 1 - \sum_{i=1}^m p_i^2 \quad (\text{Nei, 1987})$$

onde m é o número de alelos e p_i é a frequência do alelo i . A H_e de uma população é a média dos valores da H_e de todos os seus locos. Se $H_e \neq H_o$, a população está em desequilíbrio de HWE, o qual pode ser medido pelo índice de fixação F de Wright (1951).

3. Objetivos

Pelo exposto e, portanto, visando contribuir para o conhecimento e a domesticação da espécie *A. aculeata*, a aplicação de marcadores microssatélites neste trabalho teve como principais objetivos:

- (i) Determinar a estrutura populacional e os níveis de variabilidade genética dos acessos de *A. aculeata* mantidos no Banco de Germoplasma de Macaúba (BAG-Macaúba).
- (ii) Analisar os parâmetros do sistema de cruzamento, bem como os níveis de depressão por endogamia em caracteres vegetativos de famílias maternas da *A. aculeata* conservados ex situ.

O estudo teve como base as seguintes hipóteses:

- (i) A grande variabilidade morfológica dos acessos conservados ex situ é traduzida em uma abundante variação alélica, devido à larga representatividade da coleção.
- (ii) A determinação dos pools gênicos dos acessos da coleção ex situ podem coincidir com suas origens geográficas, como consequência da pressão exercida por cada região eco-geográfica, aliada ao fato da condição selvagem que se encontra a espécie *A. aculeata*.
- (iii) Os estudos sobre o sistema de cruzamento da espécie *A. aculeata* realizados com populações naturais apresentam limitações, e podem ter conduzindo a uma interpretação inadequada do sistema de reprodução da espécie.
- (iv) O processo de dispersão de semente (migração) e pólen (fluxo gênico) promovem altos índices de cruzamentos biparentais e entre indivíduos aparentados.

4. Estrutura da Tese

Esta tese está estruturada em capítulos, os quais são apresentados de forma independente.

Capítulo 1:

Molecular characterization and population structure of the macaw palm, *Acrocomia aculeata* (Arecaceae), ex situ germplasm collection using microsatellites markers

O capítulo 1 responde ao primeiro objetivo da tese que tem como base a primeira e a segunda hipótese. A seleção de material genético (acessos) para este estudo foi baseada em dois critérios: (1) representatividade da coleção mantida no Banco de Germoplasma de Macaúba, e (2) os estados e regiões do Brasil onde foram realizadas as coletas. A estrutura genética populacional foi inferida pelo o software Structure, e a caracterização da variabilidade genética dos diferentes “pools gênicos” detectados dentro da coleção viva foram determinados por diferentes medidas de diversidade genética intrapopulacional.

Capítulo 2:

Sistema de cruzamento e estimativa de depressão por endogamia em macaúba, *Acrocomia aculeata* (Arecaceae) utilizando dados quantitativos e moleculares

A proposta do capítulo 2 é atender ao segundo objetivo da tese que tem como base a terceira e a quarta hipótese. Para a genotipagem com locos SSR foram coletados tecidos vegetais de árvores maternas georreferenciadas (árvores nativas adultas) e de seus descendentes (BAG-Macaúba: coleção viva). O total de 19 famílias maternas (média de 8,3 indivíduos por árvore materna) foi utilizado neste estudo. Foram disponibilizadas informações sobre a diversidade genética das diferentes fases ontogenéticas, parâmetros de sistema de cruzamento, coancestria e do tamanho efetivo da variância, além destes, foram estimados os níveis de depressão por endogamia sobre caracteres vegetativos na fase juvenil da *Acrocomia aculeata*. As informações geradas neste capítulo são importantes para a conservação genética ex situ e in situ, bem como a adoção de critérios mais eficientes de seleção em programas de melhoramento genético.

5. Referências

- Abreu IS, Carvalho CR, Carvalho GMA, Motoike SY. 2011. 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.* 59:149–155.
- 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* (Arecaceae). *Genet Mol Biol.* 35:119–121.
- Aradhya MK, Stover E, Velasco D, Koehmstedt A. 2010. Genetic structure and differentiation in cultivated fig (*Ficus carica* L.). *Genetica.* 138:681–694.
- Arkcoll D. 1990. New crops from Brazil. In: Janick, J., Simon, E.J. (Eds.) *Advances in New Crops*. Timber Press, Portland, OR, p. 367–371.
- Barbieri RL. 2003. Conservação e uso de recursos genéticos vegetais. In: Freitas LB, Bered F (eds.). *Genética e evolução vegetal*. Porto Alegre: UFRGS, p. 403-414.
- Berg EE, Hamrick JL. 1997. Quantification of diversity at allozyme loci. *Canadian Journal Forest Research* 27: 415–424.
- Beiguelman B. 2008. *Genética de populações humanas*. Ribeirão Preto: SBG 293p. Disponível em: < <http://www.sbg.org.br/ebooks.html> > (acessado em 04/11/2014).
- Berton, LHC. 2013. Avaliação de populações naturais, estimativas de parâmetros genéticos e seleção de genótipos elite de macaúba (*Acrocomia aculeata*). Tese (Doutorado em Agricultura Tropical e Subtropical) – Instituto Agronômico de Campinas, Campinas – SP, 154p.
- Berton LHC, Azevedo Filho JA, Siqueira WJ, Colombo CA. 2013. Seed germination and estimates of genetic parameters of promising macaw palm (*Acrocomia aculeata*) progenies for biofuel production. *Industrial Crops and Products* 51: 258–266.
- Bora PS, Rocha RVM. 2004. Macaúba palm: fatty and amino acids composition of fruits. *Cienc Tecnol Aliment.* 4:158–162.
- Botstein D, White RL, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms, *American Journal of Human Genetics*, 32: 314–331.

- Bracco M, Lia VV, Gottlieb AM, Hernández JC, Poggio L. 2009. Genetic diversity in maize landraces from indigenous settlements of Northeastern Argentina. *Genetica*. 135:39–49.
- Breed MF, Gardner MG, Ottewell KM, Navarro CM, Lowe AJ. 2012. Shifts in reproductive assurance strategies and inbreeding costs associated with habitat fragmentation in Central American mahogany. *Ecology Letters* 15:444–452.
- Brito AC. 2013. Biologia reprodutiva de macaúba : floração, polinizadores, frutificação e conservação de pólen. Tese (Doutorado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa – MG, 47p.
- Brown AHD, Weir BS. 1983. Measuring genetic variability in plant population. In: Tanksley SD, Orton TJ. (Eds). *Isozymes in plants genetics and breeding*. Amsterdam: Elsevier, p. 219–239.
- Buso GSC, Ciampi AY, Moretzsohn MC, Amaral ZPS, Brondani RV. 2003. Marcadores Microssatélites em Espécies Vegetais. *Biotecnologia Ciência e Desenvolvimento* 7:46–50.
- Caetano-Anollés, Bassam BJ, Gresshoff PM. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Biotechnology* 9: 553–557.
- Caixeta ET, Oliveira ACB, Brito GG, Sakiyama NS. 2009. Tipos de Marcadores Moleculares In: Borém A, Caixeta ET (eds.) *Marcadores Moleculares*. 2ª Edição Viçosa: UFV, p.11–93.
- Charlesworth D, Morgan MT, Charlesworth B. 1990. Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. *Evolution* 44: 1469–1489.
- 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.
- Cockerham CC. 1969. Variance of Gene Frequencies. *Evolution* 23:72–84.
- Cole CT. 2003. Genetic variation in rare and common plants. *Annual Reviews Ecology Systems* 34: 213–237.
- Corley RHV, Tinker PB. 2003. *The oil palm*, 4th edition. Blackwell Science, Oxford, United Kingdom. 562p.

- Crossa J, Vencovsky R. 1994. Implications of the variance of effective population size on the genetic conservation of monoecious species. *Theoretical and Applied Genetic* 89: 936–942.
- Crow JF, Kimura MA. 1970. A introduction to population genetics theory. New York: Harper & Row, 591p.
- Cruz CD, Ferreira FM, Pessoni LA. 2011. Biometria aplicada ao estudo da diversidade genética. Visconde do Rio Branco – MG: Suprema, 620p.
- Dias LAS. 2006. Análises multidimensionais. In: Alfenas AC (Ed). Eletroforese e marcadores bioquímicos em plantas e microrganismos. Viçosa: UFV. p. 405–475.
- 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.
- Faleiro, FG. 2007. Marcadores moleculares aplicados a programas de conservação e uso de recursos genéticos. Planaltina-DF: Embrapa Cerrados, 102 p.
- Faleiro, FG. 2011. Aplicações de marcadores moleculares como ferramenta auxiliar em programas de conservação, caracterização e uso de germoplasma e melhoramento genético vegetal. In: Faleiro FG, Andrade SRM, Reis Junior FB (Ed). Biotecnologia estado da arte e aplicações na agropecuária. Planaltina-DF: Embrapa Cerrados, 730 p.
- Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*. 164:1567–1587.
- Ferreira ME, Grattapaglia D. 1998. Introdução ao uso de marcadores moleculares em análise genética. 3. ed. Brasília: Embrapa-Cenargen, 220 p.
- Finkeldey R. 1998. An Introduction to Tropical Forest Genetics. Lecture Notes. Georg-August University Göttingen, Institute of Forest Genetics and Forest Tree Breeding, 241 p.
- Gaiotto FA, Grattapaglia D, Vencovsky R. 2003. Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart). *Journal of Heredity* 94:399–406.
- Gonçalves DB, Batista AF, Rodrigues MQ, Nogueira KM, Santos VL. 2013. Ethanol production from macaúba (*Acrocomia aculeata*) presscake hemicellulosic

- hydrolysate by *Candida boidinii* UFMG14. *Bioresource Technology* 146:261–266.
- Gunn BF, Baudouin L, Olsen KM. 2011. Independent Origins of Cultivated Coconut (*Cocos nucifera* L.) in the Old World Tropics. *PLoS ONE* 6:e21143.
- Hamrick JL. 1989. Isozymes and analysis of genetic structure in plant populations. In: Soltis, D. E.; Soltis, P. (Ed.), *Isozymes and the analysis of genetic structure in plant populations*. New York: Chapman and Hall, p. 87–105.
- Haraguchi A, Li B, Matsuki S, Nagata O, Suzuki J.-I, Hara T. 2009. Variation and plasticity of photosynthesis and respiration in local populations of fat-hen *Chenopodium album* in northern Japan. *Plant Species Biology* 24:189–201.
- Hardy OJ, Maggia L, Bandou E, Breyne P, Caron H, Chevallier MH, Doligez A, Dutech C, Kremer A, Latouche-Halle C, Troispoux V, Veron V, Degen B. 2006. Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. *Molecular Ecology* 15:559–571.
- Hartl DL, Clark AG. 2007. *Principles of population genetics*. 4th edn. Sunderland, MA: Sinauer Associates.
- Henderson A, Galeano G, Bernal R. 1995. ‘Field guide to the palms of the Americas.’ (Princeton University Press: New Jersey).
- Hiane PA, Ramos Filho MM, Ramos MIL, Macedo MLR. 2005. Bocaiuva, *Acrocomia aculeata* (Jacq.) Lodd., pulp and kernel oils: Characterization and fatty acid composition. *Braz J Food Technol*. 8:256–259.
- Hiane PA, Baldasso PA, Marangoni S, Macedo MLR. 2006. Chemical and nutritional evaluation of kernels of bocaiúva, *Acrocomia aculeata* (Jacq.) Lodd. *Cienc Tecnol Aliment*. 26:683–689.
- Husband B, Schemske D. 1996. Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution* 50: 50–74.
- Karasawa MMG, Vencovsky R, Silva CM, Zucchi MI, Oliveira GCX, Veasey EA. 2007. Mating system of Brazilian *Oryza glumaepatula* populations studied with microsatellite markers. *Annals of Botany* 99: 245–253.
- Khanduri VP, Sharma CM, Kumar KS, Ghildiyal SK. 2013. Annual Variation in Flowering Phenology, Pollination, Mating System, and Pollen Yield in Two

- Natural Populations of *Schima wallichii* (DC.) Korth. The Scientific World Journal 2013: 350157.
- Kimura M, Crow JF. 1964. The measurement of the effective population number. *Evolution* 17:279–288.
- Kusdiana D, Saka S. 2001. Methyl esterification of free fatty acids of rapeseed oil as treated in supercritical methanol. *J Chem Eng Jpn.* 34:383–387.
- IBGE – Instituto Brasileiro de Geografia e Estatística. 2006. Censo agropecuário 2006: Resultados Preliminares. Rio de Janeiro: IBGE. Disponível em: < <http://www.ibge.gov.br/home/estatistica/economia/agropecuaria/censoagro/2006/agropecuaria.pdf> > (acessado em 20/10/2014).
- Lacerda DR, Acedo MDP, Lemos Filho JP, Lovato MB. 2002. A técnica de RAPD: uma ferramenta molecular em estudos de conservação de plantas. *Lundiana* 3:87–92.
- Lande R, Schemske DW. 1985. The evolution of self-fertilization and inbreeding depression in plants I: genetic models. *Evolution* 39: 24–40.
- Lanes, ECM, Costa PMA, Motoike SY. 2014a. Alternative fuels: Brazil promotes aviation biofuels. *Nature* 511: 31–31.
- Lanes ECM, Viana JMS, Paes GP, Paula MFB, Maia C, Caixeta ET, Miranda GV. 2014b. Population structure and genetic diversity of maize inbreds derived from tropical hybrids. *Genetics and Molecular Research* 13:7365–7376.
- Laurance WF, Lovejoy TE, Vasconcelos HL, Bruna EM, Didham RK, Stouffer PC, Gascon C, Bierregaard RO, Laurance SG, Sampaio E. 2002. Ecosystem decay of Amazonian forest fragments: a 22-year investigation. *Conservation Biology* 16: 605–618.
- Laurance WF, Merona JMR, Andrade A, Laurance SG, D'Angelo S, Lovejoy TE, Vasconcelos HL. 2003. Rain-forest fragmentation and the phenology of Amazonian tree communities. *Journal of Tropical Ecology* 19: 343–347.
- Lorenzi H, Souza HM, Costa JTM, Cerqueira LSC, Ferreira E. 2004. *Palmeiras brasileiras e exóticas cultivadas*. [Brazilian palm trees and exotic cultivated]. Instituto Plantarum, Nova Odessa.
- Loveless MD, Hamrick JL. 1984. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* 15: 65–95.

- Luyindula N, Mantantu N, Dumortier F, Corley RHV. 2005. Effects of inbreeding on growth and yield of oil palm. *Euphytica* 143: 9-17.
- Maizura I, Rajanaidu N, Zakri AH, Cheah SC. 2006. Assessment of genetic diversity in oil palm (*Elaeis guineensis* Jacq.) using Restriction Fragment Length Polymorphism (RFLP). *Genet Resour Crop Evol.* 53:187–195.
- Manfio CE. 2010. Análise genética no melhoramento da macaúba. Tese (Doutorado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa – MG, 52 p.
- Manfio CE, Motoike SY, Resende MDV, Santos CEM, Sato AY. 2012. Avaliação de progênies de macaúba na fase juvenil e estimativas de parâmetros genéticos e diversidade genética. *Pesquisa Florestal Brasileira* 32:63–69.
- Melo PG. 2012. Produção e caracterização de biodieseis obtidos a partir da oleaginosa macaúba (*Acrocomia aculeata*). Dissertação (Mestrado em Química) – Universidade Federal de Uberlândia, Uberlândia – MG, 80 p.
- Motta PE, Curi N, Oliveira-Filho AT, Gomes JBV. 2002. Ocorrência de macaúba em Minas Gerais: relação com atributos climáticos pedológicos e vegetacionais. *Pesquisas Agropecuária Brasileira* 37:1023–1031.
- Motoike SY, Kuki KN. 2009. The potential of macaw palm (*Acrocomia aculeata*) as source of biodiesel in Brazil. *IRECHE.* 1:632–635.
- Murawski DA, Hamrick JL. 1991. The effect of the density of flowering individuals on the mating systems of nine tropical tree species. *Heredity* 67: 167–174.
- Namoff S, Husby CE, Francisco-Ortega J, Noblick LR, Lewis CE, Griffith MP. 2010. How well does a botanical garden collection of a rare palm capture the genetic variation in a wild population? *Biological Conservation* 143: 1110–1117.
- Nazareno AG, Reis MS. 2014. At Risk of Population Decline? Na Ecological and Genetic Approach to the Threatened Palm Species *Butia eriospatha* (Arecaceae) of Southern Brazil. *Journal of Heredity* 105:120–129.
- Nei M. 1972. Genetic distance between populations. *American Naturalist* 106:283–292.
- Nei M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* 70:3321–3323.
- Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 586-590.

- Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press, 512 p.
- Nucci SM, Azevedo-Filho JA, Colombo CA, Priolli RHG, Coelho RM, Mata TL, Zucchi MI. 2008. Development and characterization of microsatellites markers from the macaw. *Mol Ecol Resour.* 8:224–226.
- Núñez-Avellaneda LR, Bernal R, Knudsen JT. 2005. Diurnal palm pollination by mytropical beetles: is it weather-related? *Plant Systematics and Evolution* 254: 149–171.
- Odong TL, Heerwaarden J, Jansen J, Hintum TJL, Eeuwijk FA. 2011. Determination of genetic structure of germplasm collections: are traditional hierarchical clustering methods appropriate for molecular marker data? *Theor Appl Genet.* 123:195–205.
- Oliveira DA, Melo Júnior AF, Brandão MM, Rodrigues LA, Menezes EV, Ferreira PRB. 2012. Genetic diversity in populations of *Acrocomia aculeata* (Arecaceae) in the northern region of Minas Gerais, Brazil. *Gen Mol Res.* 11:531–538.
- Pandin DS. 2009. Inbreeding depression analysis based on morphological characters in four generations of selfed mapanget tall coconut N°32 (*Cocos nucifera* L.). *Indonesian Journal of Agriculture* 2: 110-114.
- Portela RCQ, Santos FAM. 2014. Impact of forest fragment size on the population structure of three palm species (Arecaceae) in the Brazilian Atlantic Rainforest. *Revista de Biología Tropical* 64: 433–442.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure from multilocus genotype data. *Genetics.* 155:945–959.
- Ritland K, Jain S. 1981. A model for the estimation of outcrossing rate and gene frequencies using independent loci. *Heredity* 47: 35–52.
- Ritland K. 1989. Correlated matings in the partial selfer *Mimulus guttatus*. *Evolution* 43:848–859.
- Ritland K. 2002. Extensions of models for the estimation of mating systems using independent loci. *Heredity* 88: 221–228.
- Röser M, Johnson MAT, Hanson L. 1997. Nuclear DNA amounts in palms (Arecaceae). *Bot Acta.* 110:79–89.

- Saldaña, A. Lusk CH, Gonzáles WL, Gianoli E. 2007. Natural selection on ecophysiology traits of a fern species in a temperate rainforest. *Evolutionary Ecology* 21: 651-662.
- Scariot A, Lleras E, Hay JD. 1991. Reproductive biology of the palm *Acrocomia aculeata* in Central Brazil. *Biotropica*. 23:12–22.
- Scariot A, Lleras E. 1995. Flowering and fruiting phenologies of the palm *Acrocomia aculeata*: patterns and consequences. *Biotropica*. 27:168–173.
- Sebbenn AM. 2001. Distribuição da variação genética de populações de jequitibá-rosa [*Cariniana legalis* (Mart.) O. Ktze] por caracteres quantitativos e isoenzimas. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba – SP, 210 p.
- Silva JC, Barrichelo LEG, Brito JO. 1986. Endocarpos de babaçu e macaúba comparados a madeira de *Eucalyptus grandis* para a produção de carvão vegetal. *IPEF*. 34:31–34.
- Silva CRS, Albuquerque PSB, Ervedosa FR, Mota JWS, Figueira A, Sebbenn AM. 2011. Understanding the genetic diversity, spatial genetic structure and mating system at the hierarchical levels of fruits and individuals of a continuous *Theobroma cacao* population from the Brazilian Amazon. *Heredity* 106: 973–985.
- Silveira G, Lopes FA, Motoike SY, Souza FS, Henriques E. 2013. Banco de Germoplasma da Macaúba: base para o melhoramento genético. In: I Congresso Brasileiro de Macaúba: Consolidação da Cadeia Produtiva; 2013 Nov. UNIPAM. Patos de Minas – MG. 4 p.
- Stajner N, Satovic Z, Cerenak A, Javornik B. 2008. Genetic structure and differentiation in hop (*Humulus lupulus* L.) as inferred from microsatellites. *Euphytica*. 161:301–311.
- Thiele J, Hansen T, Siegismund HR, Hauser TP. 2010. Genetic variation of inbreeding depression among floral and fitness traits in *Silene nutans*. *Heredity* 104: 52–60.
- Tolêdo DP. 2010. Análise técnica, econômica e ambiental de macaúba e de pinhão-mansão como alternativas de agregação de renda na cadeia produtiva de biodiesel. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Uberlândia, Uberlândia – MG, 92 p.

- Um Pé de Quê?/ Instituto Plantarum - Harri Lorenzi. 2014. Disponível em: <
http://www.umpedeque.com.br/bkp/site_umpedeque/public/img/mapa/MACAUBA_mapa_gde.jpg> (acessado em 16/10/2014).
- Vencovsky R. 1987. Tamanho efetivo populacional na coleta e preservação de
 germoplasma de espécies alógamas. Instituto de Pesquisas Florestais 35:79–84.
- Vieira SS, Magriots ZM, Santos NAV, Cardoso MG, Saczk AA. 2012. Macauba palm
 (*Acrocomia aculeata*) cake from biodiesel processing: An efficient and low cost
 substrate for the adsorption of dyes. Chemical Engineering. Journal 183:152–16.
- Von Lisingen L, Cervi AC. 2009. *Acrocomia aculeata* (Jacq.) Lodd ex Mart., nova
 ocorrência para a flora do Estado do Paraná. Acta Biol Parana. 38:187–192.
- Wandeck FA, Justo PG. 1988. A macaúba, fonte energética e insumo industrial: sua
 significação econômica no Brasil. In: Simposio sobre o Cerrado, Savana, Brasília.
 EMBRAPA, CPAC, p. 541–577.
- Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analyses of population
 structure. Evolution 38: 1358–1370.
- Weir, B. S. 1996. Genetic data analysis II: methods for discrete population genetic data.
 Sunderland, Massachusetts: Sinauer Associates, 455 p.
- Welsh J, McClelland M. 1990. Fingerprinting genomes using PCR with arbitrary
 primers. Nucleic Acids Research 18: 7213–7218.
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. 1990. DNA
 polymorphisms amplified by arbitrary primers are useful as genetic markers.
 Nucleic Acids Research 18: 6531–6535.
- Wright S. Evolution in Mendelian populations. Genetics, 16, 97–159, 1931.
- Wright S. 1951. The genetical structure of populations. Annals of Eugenics 15: 395–420.
- Wright S. 1965. The interpretation of population structure by F- statistics with special
 regard to systems of mating. Evolution 19: 395–420.
- <http://apps.kew.org/wcsp/qsearch.do> (acessado em 05/11/2014).
- http://www.palmweb.org/cdm_dataportal/taxon/d58c0b44-29f8-4071-aa49-32baa1852
 (acessado em 05/11/2014).
- <http://www.splink.org.br> (acessado em 05/11/2014).

<http://sciweb.nybg.org/Science2/hcol/lists/arecaceae.html> (acessado em 05/11/2014).

<http://www.tropicos.org/Name/2401428?tab=synonyms> (acessado em 05/11/2014).

CAPÍTULO 1

Este manuscrito encontra-se publicado no periódico Journal of Heredity

Molecular characterization and population structure of the macaw palm, *Acrocomia aculeata* (Arecaceae), ex situ germplasm collection using microsatellites markers

Éder C. M. Lanes, Sérgio Y. Motoike, Kacilda N. Kuki, Carlos Nick, and Renata D. Freitas

Laboratory of Biotechnology and Plant Breeding, Federal University of Viçosa, 36570-000, Viçosa, MG, Brazil (Lanes, Motoike and Freitas); and the Department of Plant Science, Federal University of Viçosa, 36570-000, Viçosa, MG, Brazil (Motoike, Kuki and Nick)

Address correspondence to Éder C. M. Lanes at the address above, or e-mail: edercml@gmail.com

Doi: 10.1093/jhered/esu073

Abstract

The *Acrocomia aculeata* is one of the most promising plants for sustainable production of renewable energy. In order to understand patterns of the distribution of the allelic diversity of *A. aculeata* ex situ germplasm collection, the present study investigated the hypothesis that the genetic variability of the accessions may match their geographical origin. A genotypic analysis of 77 *A. aculeata* accessions was conducted with six simple sequence repeat (SSR) markers. A high degree of molecular diversity among the accessions was found, with an average of nine alleles per locus and a polymorphic information content with a mean of 0.76. A total of four clusters was identified by the Bayesian analysis of population structure. The highest subpopulation diversity was identified in Pop1, mainly formed by accessions from State of Mato Grosso do Sul. The populations Pop2A, Pop2B and Pop2C, all from the State of Minas Gerais, showed high genetic variability as determined by a higher F_{st} , and a wide genetic variance, which were identified within and among the population by AMOVA. Based on our results and on Vavilov's theory on crop origins, one possible diversity center for *A. aculeata* is proposed to be in a region in southeast Brazil.

Key words: biofuel, diversity center, domestication, gene pools, speciation

Introduction

The macaw palm (*Acrocomia aculeata* (Jacq.) Lodd. ex Martius), also known as macauba, is a perennial palm that is commonly found in tropical and subtropical regions of the Americas. In Brazil, *A. aculeata* is profusely dispersed throughout the country, with the highest densities in the Midwest and Southeast regions (Henderson et al. 1995). This arboreal oleiferous species is monoecious with marked protogyny and a mixed mating system with predominance of outcrossing, although self-pollination seems to be a frequent occurrence (Scariot et al. 1991; Abreu et al. 2012). Its genome has 2841 Mb distributed in 15 pairs of chromosomes ($2n = 30$) (Röser et al. 1997; Abreu et al. 2011). Adult plants can exceed 15 m in height, and their single stem is usually covered with thorns and remnants of leaf petiole. Androgynous inflorescences, with yellow flowers grouped in spikes, are protected by a spathe up to 2 m in length. Drupes, arranged in bulky bunches, range from 25 to 60 mm in diameter and when ripped are composed of a fibrous and brittle epicarp, a fibrous and fleshy mesocarp rich in fatty acids and glycerides, a thick lignified endocarp and a solid endosperm (kernel) with a high content of lipids (52–69%), proteins (17.06%), fiber (15.8%) and unsaturated fatty acids (20%) (Scariot and Lleras 1995; Lorenzi et al. 2004; Bora and Rock 2004; Hiane et al. 2006; Motoike and Kuki 2009; Von Lisingen and Cervi 2009; Ciconini et al. 2013).

The search for renewable energy sources as an alternative to petroleum fuels has resulted in interest in the macaw palm as a potential raw material for biofuel production. Because of its high yield (above 25 t fruits / ha⁻¹), the palm can generate up to 4500 liters of oil per ha⁻¹ that is predominantly composed of unsaturated fatty acids (81%), which is an important trait for biodiesel (Kusdiana and Saka 2001; Bora and Rocha 2004; Motoike and Kuki 2009; Moura et al. 2010; Pires et al. 2013). This estimate exceeds the oil yield production of annual crops, such as soybeans (420 liters), sunflower (700 liters), peanuts (900 liters) and castor (660 liters), and perennial crops, such as oil palm (3740 liters per hectare / year) (Abdalla et al. 2008; Sumathi et al. 2008).

However, the profitability of this species is amplified especially due to the multitude of uses for its fruit as illustrated by Pires et al. (2013). Based on the fatty acid profile of the oil from different parts of the fruit, the macaw palm can be aimed at specific industries other than the biofuel sector. Kernel oil, which has a high content of lauric acid (13–45%) (Bora and Rocha 2004; Hiane et al. 2005), is suitable for the cosmetics industry, whereas the oil from the pulp, which contains up to 72.59% oleic

acid (comparable to olive oil), is appropriated for the food industry (Bora and Rocha 2004; Hiane et al. 2005; Ciconini 2012). The co-products and by-products generated from the processing of the fruits for biofuel are highly valued, which facilitates additional profitability from its production. For example, brans from kernel or pulp can be used in animal nutrition, and the lignified endocarp, when turned into charcoal, may be used as a source of heat energy (Silva et al. 1986).

Nevertheless, prospecting for new sources of biological raw materials should not be based only on the potential of the products and revenue that the given plant species can offer. Equally relevant, bioprospecting should investigate the possible biological variations of the species in question (Jeffery 2002). The precaution ensures knowledge of the genetic diversity among plant materials and simultaneously paves the best way for their domestication. The formation of a germplasm bank and its genetic characterization corresponds to the first steps.

The Active Germplasm Bank of Macauba (BAG – Macauba) at the Universidade Federal de Viçosa (UFV – Brazil) is the first official repository registered by the Brazilian Board of Management of Genetic Heritage (# 084/2013–SECEX/CEGEN) and stands out as one of the largest genebank of *A. aculeata* in Latin America and the world. Currently, the repository holds a living collection of 253 maternal families representing almost all Brazilian regions. The set of accessions in this ex situ collection has wide morphological variability and is an important genetic resource for the initiation of plant breeding programs.

The genetic diversity of populations is the basis of the evolutionary potential of the species to respond to environmental changes (Toro and Caballero 2005). Knowing the patterns of genetic variability among and within populations can help in the efficient management, operation and maintenance of germplasm collections (Renau-Morata et al. 2005; Belaj et al. 2012). Studies using molecular markers to estimate genetic structure in ex situ collections and the diversity of populations can provide crucial information for associative mapping (Yang et al. 2011), linkage disequilibrium (Wang et al. 2012), determination of genetic relationships (Stajner et al. 2008; Bracco et al. 2009) and the establishment of core collections (Odong et al. 2011). The use of molecular biology for the purpose of characterization of the germplasm collections has allowed the understanding of the historical process of domestication, the center of diversity and the origin of palms such as coconut and oil palm (Corley and Tinker 2003; Maizura et al. 2006; Gunn et al. 2011). In the Neotropics, while evidence supports that human groups have played an important role in the dispersal of *A. aculeata*, no references indicate the

existence of a probable center of diversity for the species, although South America is highly quoted (Morcote-Rios and Bernal 2001). The center of diversity is defined as the geographic region wherein the species exhibits the highest degree of genetic variability, represented by the highest number of either cultivated types or wild relatives. On the other hand, the center of origin of a crop plant is normally associated with the geographical origin of its wild and weedy relatives, in other words, is the location where it first appeared (Harlan 1971).

Molecular markers offer a stable and reliable alternative for the characterization of germplasm banks (Aradhya et al. 2010), providing a direct measure of genetic diversity among accessions by means of genetic polymorphism at the DNA level. Microsatellites, or Simple Sequence Repeats (SSR), and Random Amplification of Polymorphic DNA (RAPD) have recently been used in the assessment of genetic diversity, mating systems, and structure and differentiation of natural populations of *A. aculeata* (Nucci et al. 2008; Oliveira et al. 2012; Abreu et al. 2012). Despite researchers' efforts, the information about *A. aculeata* is still incipient. Furthermore, no reports exist on the genetic variability of the macaw palms that are conserved in ex situ collections.

As part of an ongoing effort to characterize the germplasm of BAG – Macauba, the present study investigated the hypothesis that the genetic variability of macaw palm in the ex situ collection may match the accessions' geographical origin. To test it, we examined the population structure in the genebank by evaluating the genetic diversity and differentiating among gene pools based on microsatellite markers.

Materials and Methods

Plant material

A panel of 77 *A. aculeata* accessions was obtained from the Active Germplasm Bank of Macauba (BAG – Macauba) situated on the Araponga Experimental Farm (20° 40'1"S, 42°31'15"W) in the municipality of Araponga, State of Minas Gerais, Brazil. For more details, see Pires et al. (2013). This living collection is maintained by the Genetic Improvement Program for Macaw Palm at the Plant Science Department of the Universidade Federal de Viçosa (UFV – Brazil). The climate of the region is of rainy summers and dry winters, Cwb type by Köppen classification. The rain precipitation reaches annual average of 1339 mm, and the average annual temperature is 18 °C. The selection of genetic material for this study was based on two criteria: (1) representativeness from the collection maintained at the Germplasm Bank and (2) the States and regions of Brazil where the samples were collected (Figure 1A).

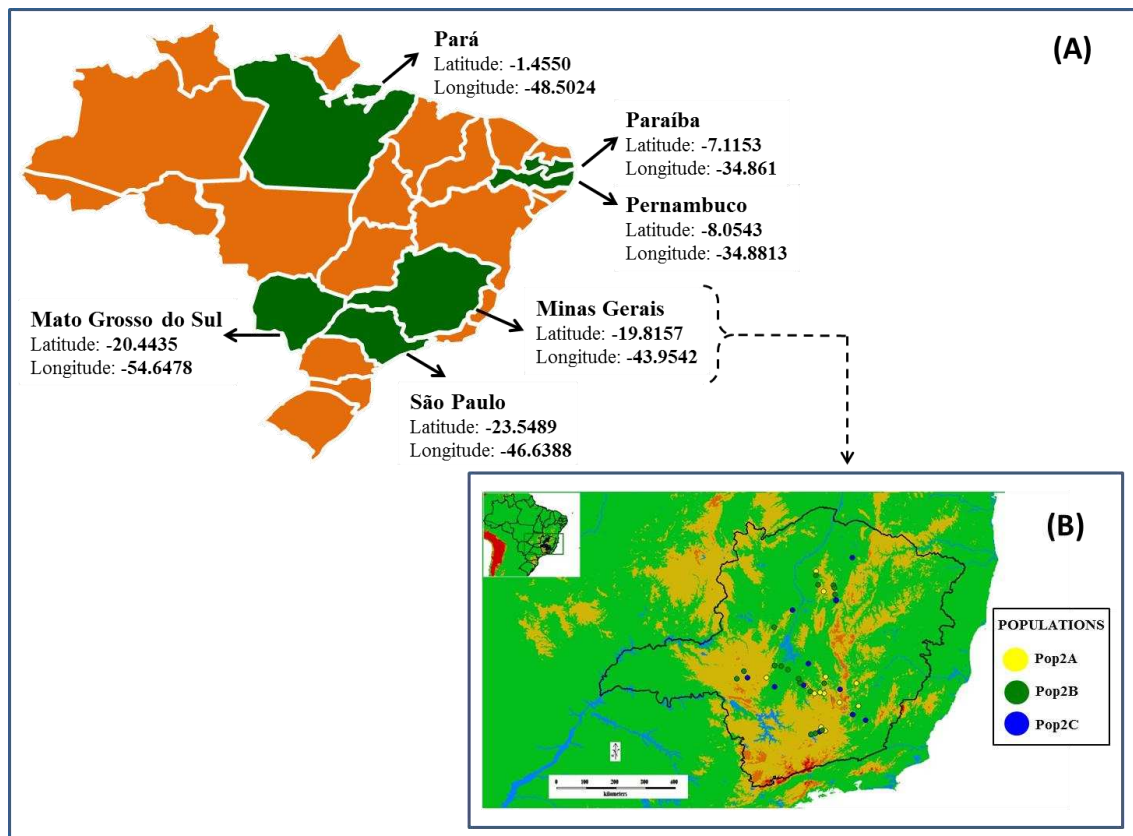


Figure 1. Brazilian map depicting; (A) Provenance of *Acrocomia aculeata* accessions collected from different Brazilian states to compose the Macaw Palm Germplasm Bank (BAG – Macauba), with latitude and longitude in decimal degrees; (B) Proposed map of the diversity center for *Acrocomia aculeata*, in the state of Minas Gerais/Brazil, based on the examined accessions grouped into three subpopulations inferred in the investigations. Diverse color spots represent individuals from different gene pools.

The geographic information and climatic characteristics of 77 accessions' provenance are available in Table 1. The data were collected from the weather stations of each mesoregion and relates to the historical series of 1961 – 2014, made available by the Brazilian National Institute of Meteorology (INMET).

Genomic DNA extraction and genotyping microsatellites

After choosing the individual plants, leaflets were collected from non-senescent healthy fronds. The leaflets were excised from the rachis and immediately placed in liquid nitrogen and sent to the Laboratory of Biotechnology and Plant Breeding of the Plant Science Department – DFT/ UFV, in Viçosa, MG. The samples were lyophilized for 72 h and then stored in the freezer at -20°C. Genomic DNA was extracted from leaf tissue using hexadecyltrimethylammonium bromide (CTAB) following a modified protocol that was previously described by Lanes et al. (2013). The integrity and concentration of each DNA sample was determined by spectrophotometry (absorbance at 260 and 280 nm) and gel electrophoresis 0.8% agarose, respectively. Working stocks containing 70 µL were prepared at a concentration of 10 ng µL⁻¹.

Table 1. Identification of 77 *Acrocomia aculeata* accessions assessed from the Active Germplasm Bank of Macauba (BAG – Macauba) with the correspondent bank code and provenance (geographic and climatic characteristics).

Code	Geographical Region					Climatic factors		
	Lt (S)	Ln (W)	Elv	States	Mesorregions	Tm	Rn	Ur
BGP42	21.15	51.11	393	SP	Araçatuba	23.1	111.2	67.4
BGP35	21.16	51.14	411	SP	Araçatuba			
BGP51	21.53	48.74	509	SP	Araraquara	23.1	108.2	69.5
BGP39	21.59	49.04	477	SP	Araraquara			
BGP34	22.42	50.58	511	SP	Assis	23.1	111.2	67.4
BGP47	22.49	50.78	399	SP	Assis			
BGP24	21.24	44.45	1049	MG	Campos das Vertentes	18.2	116.7	80.4
BGP6	21.19	44.32	924	MG	Campos das Vertentes			
BGP3	21.19	44.33	925	MG	Campos das Vertentes			
BGP1	21.27	44.59	957	MG	Campos das Vertentes			
BGP14	21.19	44.33	925	MG	Campos das Vertentes			
BGP68	21.05	44.28	969	MG	Campos das Vertentes			
BGP70	21.15	44.24	929	MG	Campos das Vertentes			
BGP15	21.17	44.14	906	MG	Campos das Vertentes			
BGP58	19.78	45.68	692	MG	Central Mineira	22.8	88.9	68.4
BGP54	19.08	44.65	672	MG	Central Mineira			
BGP31	18.33	45.11	733	MG	Central Mineira			
BGP40	19.56	44.96	658	MG	Central Mineira			
BGP27	19.16	45.69	851	MG	Central Mineira			
BGP16	19.27	45.29	620	MG	Central Mineira			
BGP50	19.17	45.49	612	MG	Central Mineira			
BGP37	19.68	44.91	674	MG	Central Mineira			
BGP2	19.89	43.69	992	MG	Metropolitana de Belo Horizonte	21.5	127.9	68.0
BGP69	19.76	44.80	675	MG	Metropolitana de Belo Horizonte			
BGP65	19.89	43.92	860	MG	Metropolitana de Belo Horizonte			
BGP26	19.94	44.6	989	MG	Metropolitana de Belo Horizonte			
BGP48	20.00	44.47	823	MG	Metropolitana de Belo Horizonte			
BGP4	19.70	44.18	876	MG	Metropolitana de Belo Horizonte			
BGP33	19.77	44.76	668	MG	Metropolitana de Belo Horizonte			
BGP13	20.00	44.18	761	MG	Metropolitana de Belo Horizonte			
BGP11	20.30	43.71	942	MG	Metropolitana de Belo Horizonte			
BGP36	20.00	44.18	761	MG	Metropolitana de Belo Horizonte			
BGP78	19.70	43.19	774	MG	Metropolitana de Belo Horizonte	24.2	110.7	66.0
BGP30	19.97	44.30	759	MG	Metropolitana de Belo Horizonte			
BGP29	19.49	44.13	748	MG	Metropolitana de Belo Horizonte			
BGP74	17.96	45.70	712	MG	Noroeste de Minas			
BGP63	16.22	44.43	725	MG	Norte de Minas			
BGP10	16.74	43.86	684	MG	Norte de Minas			
BGP49	17.12	43.82	686	MG	Norte de Minas			
BGP79	16.64	44.36	774	MG	Norte de Minas			
BGP66	16.95	43.86	1008	MG	Norte de Minas		22.9	65.3
BGP56	16.66	43.90	684	MG	Norte de Minas			

Table 1. Continued

Code	Geographical Region					Climatic factors		
	Lt (S)	Ln (W)	Elv	States	Mesorregions	Tm	Rn	Ur
BGP67	17.43	45.15	695	MG	Norte de Minas	19.8	101.5	80.5
BGP19	15.80	43.32	510	MG	Norte de Minas			
BGP91	16.85	44.20	870	MG	Norte de Minas			
BGP87	16.36	44.43	730	MG	Norte de Minas			
BGP9	20.66	43.31	679	MG	Zona da Mata de Minas			
BGP44	20.40	43.13	574	MG	Zona da Mata de Minas	20.9	128.0	74.1
BGP45	20.84	42.91	594	MG	Zona da Mata de Minas			
BGP64	19.52	46.53	990	MG	Triângulo Mineiro e Alto			
BGP71	19.33	46.64	861	MG	Triângulo Mineiro e Alto			
BGP12	19.55	46.85	917	MG	Triângulo Mineiro e Alto			
BGP80	19.52	45.94	953	MG	Triângulo Mineiro e Alto	25.6	80.2	72.9
BGP23	-	-	-	MG	-			
BGP112	21.15	55.83	219	MS	Pantaneais Sul-Mato-Grossenses			
BGP114	20.88	55.86	219	MS	Pantaneais Sul-Mato-Grossenses			
BGP113	21.47	56.17	259	MS	Pantaneais Sul-Mato-Grossenses			
BGP93	21.70	57.84	90	MS	Pantaneais Sul-Mato-Grossenses			
BGP116	20.84	55.91	219	MS	Pantaneais Sul-Mato-Grossenses			
BGP117	20.50	55.31	142	MS	Pantaneais Sul-Mato-Grossenses			
BGP105	21.70	57.84	90	MS	Pantaneais Sul-Mato-Grossenses			
BGP103	20.50	55.31	142	MS	Pantaneais Sul-Mato-Grossenses			
BGP104	21.48	56.17	259	MS	Pantaneais Sul-Mato-Grossenses			
BGP119	20.50	55.31	142	MS	Pantaneais Sul-Mato-Grossenses			
BGP118	21.48	56.17	259	MS	Pantaneais Sul-Mato-Grossenses			
BGP94	20.50	55.31	142	MS	Pantaneais Sul-Mato-Grossenses			
BGP95	21.70	57.84	90	MS	Pantaneais Sul-Mato-Grossenses			
BGP92	21.48	56.17	259	MS	Pantaneais Sul-Mato-Grossenses			
BGP106	21.48	56.17	259	MS	Pantaneais Sul-Mato-Grossenses			
BGP107	20.84	55.91	219	MS	Pantaneais Sul-Mato-Grossenses			
BGP101	21.70	57.89	90	MS	Pantaneais Sul-Mato-Grossenses			
BGP102	20.47	55.78	142	MS	Pantaneais Sul-Mato-Grossenses			
BGP100	20.51	55.63	142	MS	Pantaneais Sul-Mato-Grossenses			
BGP115	20.46	55.78	142	MS	Pantaneais Sul-Mato-Grossenses			
BGP98	-	-	-	PA	-	-	-	-
BGP124	-	-	-	PB	-	-	-	-
BGP82	-	-	-	PE	-	-	-	-

Abbreviations: Geographic: Lt = latitude (decimal degrees). Ln = longitude (decimal degrees). Elv = Elevation (m); Brazilian States: São Paulo = SP, Minas Gerais = MG, Mato Grosso do Sul = MS, Pará = PA, Paraíba = PB, Pernambuco = PE; Climatic factors of the mesorregions: Tm = average annual temperature (°C), Rn = average annual rainfall (mm), Ur = average annual relative humidity (%).

A total of 77 microsatellite loci available for *A. aculeata* were tested; however, only 12 loci were successfully amplified by the examined accessions from the BAG – Macauba. Half of these loci were discarded, four due to difficulty assessing the amplified PCR products and two others for being monomorphic. Therefore, only six

SSR markers (Aacu 7, Aacu 45, Aacu 12, Aacu 26, Aacu 10 and Aacu 30), which were developed by Nucci et al. (2008), were effectively used. The Polymerase Chain Reaction (PCR) was conducted using an Applied Biosystems Veriti™ thermal cycler (Applied Biosystems). The sequences of the primers are listed in Table 2.

Table 2. Microsatellite loci, along with their respective allele amplitude, number of alleles (A), major allele frequency (MAF), expected and observed heterozygosity (H), and polymorphism information content (PIC) in the 77 accessions of *Acrocomia aculeata*.

Locus	Motif	Allelic size range (bp)	A	MAF	H _E	H _O	PIC
Aacu7	(GA) ₁₃	144-175	10	0.23	0.82	0.61	0.79
Aacu45	(CGAC) ₅	244-312	10	0.54	0.63	0.23	0.58
Aacu12	(TC) ₂₀	166-205	10	0.24	0.84	0.40	0.82
Aacu26	(AC) ₁₃ (AG) ₁₄	262-301	6	0.37	0.77	0.22	0.73
Aacu10	(AG) ₁₆	164-188	10	0.30	0.84	0.47	0.82
Aacu30	(CA) ₁₈	129-303	8	0.24	0.82	0.29	0.79
Mean	-	-	9	0.32	0.79	0.37	0.76

The total volume of each reaction was 20 µL containing 30 ng genomic DNA, 1X PCR buffer, 250 µM dNTPs, 0.15 µM primer SSR, 3 mM MgCl₂ and 1 U Taq polymerase. The thermocycler was programmed for one initial pre-denaturation step of 5 min at 94°C followed by 30 cycles of 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C for extension with a final extension at 72°C for 8 min. The resulting DNA fragments from the amplification were separated by electrophoresis on 6.0% denaturing polyacrylamide (w/v) gels [(19:1) acrylamide/bis-acrylamide (Bio-Rad), 7.5 M urea, 5.0X TBE] using the electrophoresis system "Sequi-Gen GT" (Bio-Rad) in 1X TBE buffer (0.09 M Tris base, 0.09 M boric acid and 2 mM EDTA, pH 8.0). The gels were visualized by silver nitrate staining technique.

Data analysis

The number of alleles, the frequency of the most frequent alleles, the expected and observed heterozygosity and the polymorphism information content (PIC) were calculated for each microsatellite locus using the software PowerMarker version 3.25 (Liu and Muse 2005). The size of the alleles in each locus was determined using AlphaDigiDoc™ 1201 software (version 3.3.0, Alpha Innotech), and a 100 bp DNA ladder was used as the standard (Invitrogen, São Paulo, Brazil).

Population sub-structuring within the accessed genotypes of *A. aculeata* was examined following the Bayesian clustering approach described by Pritchard et al.

(2000) using the software Structure version 2.3.4 (Falush et al. 2003; Pritchard et al. 2000). Fifteen runs were performed by setting the number of clusters (K) from 1 to 15. Each run consisted of a burn-in of 50,000 followed by 100,000 MCMC (Markov Chain Monte Carlo) replications, assuming the admixture model and correlated allelic frequencies. No prior information was used. The choice of the most likely number of clusters (K) was carried out by calculating the statistics ΔK , which is based on the rate of change in the log probability of the data between successive K values, as described by Evanno et al. (2005) using Structure Harvester (Earl and vonHoldt 2012). Among the 15 runs per K, the one with the highest maximum likelihood was used to assign individual genotypes to clusters. Individuals with membership probabilities ≥ 0.70 were assigned to corresponding clusters, and individuals with membership probabilities < 0.70 were assigned to a “mixed” group. Based on the populations identified by Structure, various estimates of population parameters among and within gene pools were analyzed.

An analysis of molecular variance (AMOVA) (Excoffier et al. 1992) and Wright’s F-statistics (F_{st}) across all subpopulations and between pairwise subpopulations was performed using PowerMarker version 3.25 (Liu and Muse 2005) to investigate population differentiations among the subpopulations clustered by Structure. Additionally, Nei’s genetic distances (Nei 1983) among these given subpopulations and individuals were also used to construct a neighbor-joining (NJ) phylogenetic tree with 1000 bootstrapping runs using the software PowerMarker version 3.25 (Liu and Muse 2005). The resulting tree was visualized with Fig Tree version 1.4.0 (Rambaut 2012).

To complement the output of the phylogenetic analyses by NJ, a principal coordinate analysis was conducted (PCoA; Orloci 1978) using the program GenAlEx v.6.4 (Peakall and Smouse 2006), with the former being most sensitive to differences between closely related individuals and the latter being more sensitive to differences among groups (Hauser and Crovello 1982). Two analyses of PCoA were performed, in which the first analysis was with clusters identified by their geographic origin (Brazilian States) and the second based on the population structure analysis by Structure.

The genetic diversity parameters of the populations identified by Structure were characterized by the average number of alleles per locus, the average effective number of alleles per locus, the average number of different alleles with a frequency greater than 0.05, the gene diversity, the observed heterozygosity, the number of rare alleles (frequency of alleles smaller than 0.05) and the percentage of polymorphic loci. The presence of population-specific alleles (hereafter referred to as private alleles, i.e.,

alleles present in only one population and absent in the others) was examined for each population (Maguire et al. 2002). These analyses were all computed using GenAlEx 6.4 (Peakall and Smouse 2006). The geographic distribution of *A. aculeata* accessions was visualized with the DIVA-GIS program (Hijmans et al. 2012).

In fulfillment of data archiving guidelines (Baker 2013), we have deposited the primary data underlying these analyses in Dryad, in accordance with the journal's guidelines.

Results

Allelic diversity of microsatellite loci

In this study, the six microsatellite markers were polymorphic across the 77 *A. aculeata* accessions, amplifying a total of 54 alleles, which ranged in size from 129 to 312 bp. The overall mean of the major allele frequencies was 0.32, and the average number of alleles per microsatellite locus was nine, ranging from six (Aacu26) to ten (Aacu7, Aacu45, Aacu12 and Aacu10). The PIC value of the SSR loci ranged from 0.58 to 0.82 with an average of 0.76. The overall mean of the observed heterozygosity in the macaw palm was 0.37, with the highest value at 0.61 (Aacu7) and the lowest value at 0.22 (Aacu26), while the highest expected heterozygosity was 0.84 (Aacu12 and Aacu10) and the lowest was 0.63 (Aacu45) with an average of 0.79 (Table 2).

Population structure and genetic diversity

Bayesian analysis of population structure, without incorporating prior population information, indicated that the best estimate of K was two populations (Pop1 and Pop2) according to the peak value of ΔK ($\Delta K = 436.1157$) (Figure 2A). For each inferred population, a new analysis of the population structure was conducted by Structure, where a sub-structuring was detected. Pop2 was the only population to have the sub-structuring present, and this population was consistently and clearly divided into three ($K = 3$) subgroups among the 15 iterations for each K from 1 to 10 (Figure 2B). Finally, a total of four clusters was identified. Therefore, the three subgroups originating from Pop2 were labeled as Pop2A, Pop2B and Pop2C. Among these four populations there is evidence of individuals with admixture ancestry. The blend (crossbred) is mainly between Pop2A and Pop2B, Pop2A and Pop2C, and Pop2B and Pop2C. A total of 55 accessions (71%) were assigned to one of the four populations, where $\geq 70\%$ of their inferred ancestry was derived from one of the model-based populations. Among these accessions, 25, 7, 12 and 11 clearly belong to Pop1, Pop2A, Pop2B and Pop2C,

respectively. The remaining 22 accessions (29%) were categorized as having admixed ancestry, including eight with an admixture between populations Pop2A and Pop2B, three with an admixture between populations Pop2A and Pop2C, eight with an admixture between the populations Pop2B and Pop2C, and three that share genomes from populations Pop1 and Pop2 (A, B and C) (Figure 2A, B). Of the four populations identified by Structure, three of these gene pools consisted essentially of *A. aculeata* accessions obtained from the State of Minas Gerais, 100, 95 and 88% composing the populations Pop2A, Pop2B and Pop2C, respectively. The population Pop1 consisted mainly of accessions from the State of Mato Grosso do Sul, which displayed 95% of accessions from this State with more than 90% of their inferred ancestry derived from that population (Figure 2A).

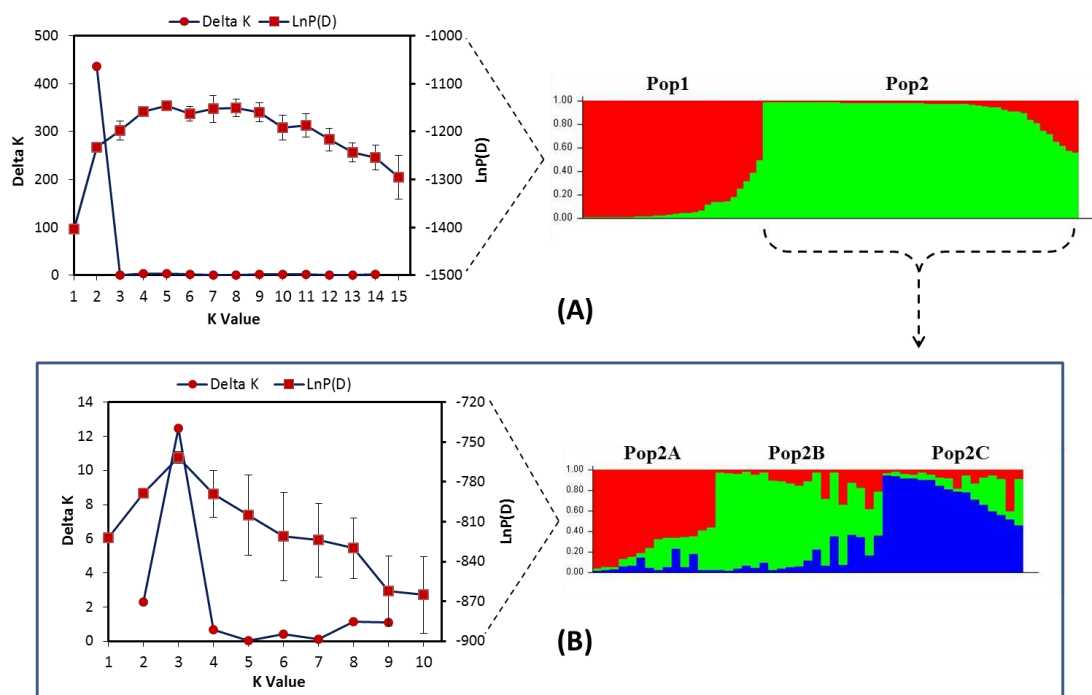


Figure 2. Population structure of *Acrocomia aculeata* from the genebank BAG – Macauba. Mean posterior probabilities LnP(D) and estimated Δk over fifteen repeats of Structure analysis; (A) Optimal population structure of 77 accessions of *A. aculeata* (K=2) and (B) sub-structuring (K=3) of Pop2 assessed by Structure. Each individual is represented by a thin vertical line, which is partitioned into five colored segments that represent the individual estimated membership (EM) to the five clusters.

The neighbor-joining tree constructed by Nei's genetic distance (Figure 3A, B) revealed genetic relationships that were fairly consistent with the Structure-based membership assignment for most of the accessions. Furthermore, a principal coordinate analysis of the 77 accessions was conducted to further assess the population subdivisions identified by both their geographic origin and Structure analysis (Figure 4A, B). The first principal coordinate explained 27% of the genetic variance, and the

second principal coordinate explained 24% of the genetic variance. Plots of the first and second main coordinates shows a clear separation of Pop1 from the remaining populations, explaining 51% of the total molecular marker variance.

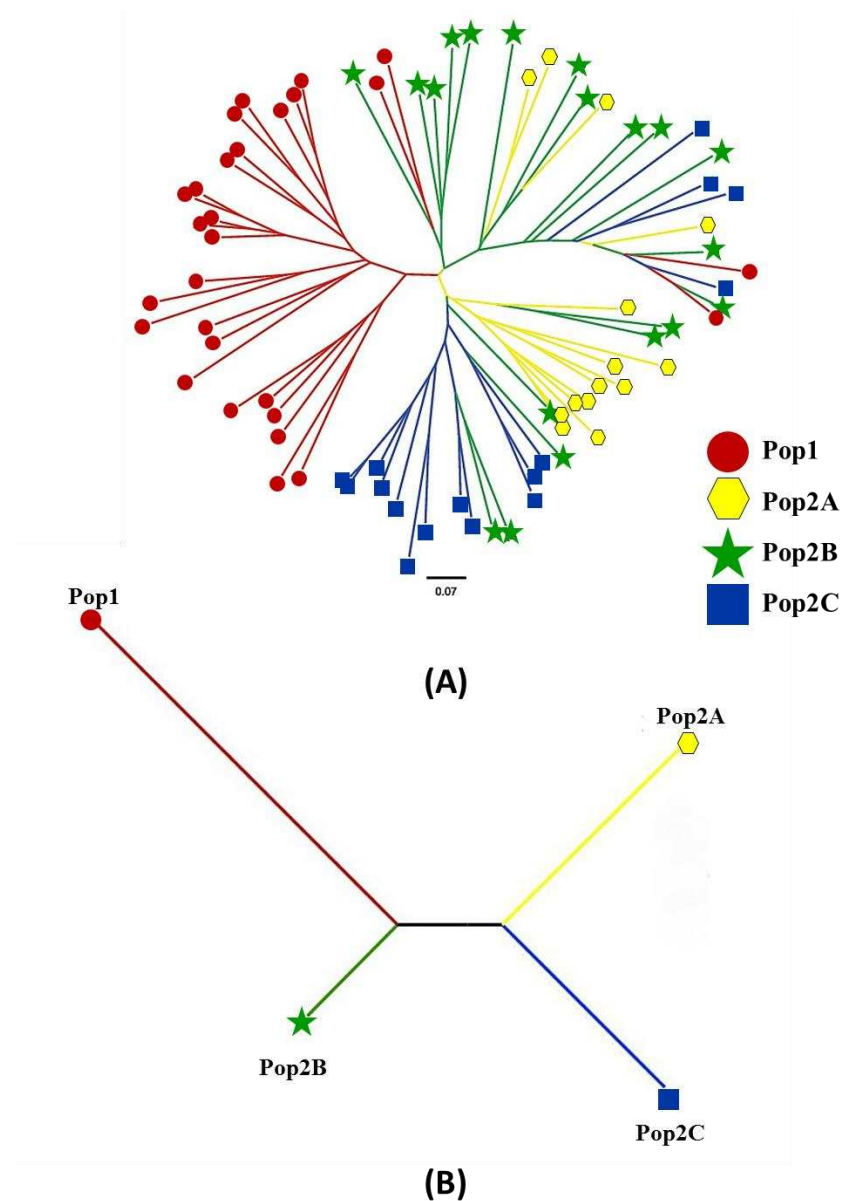
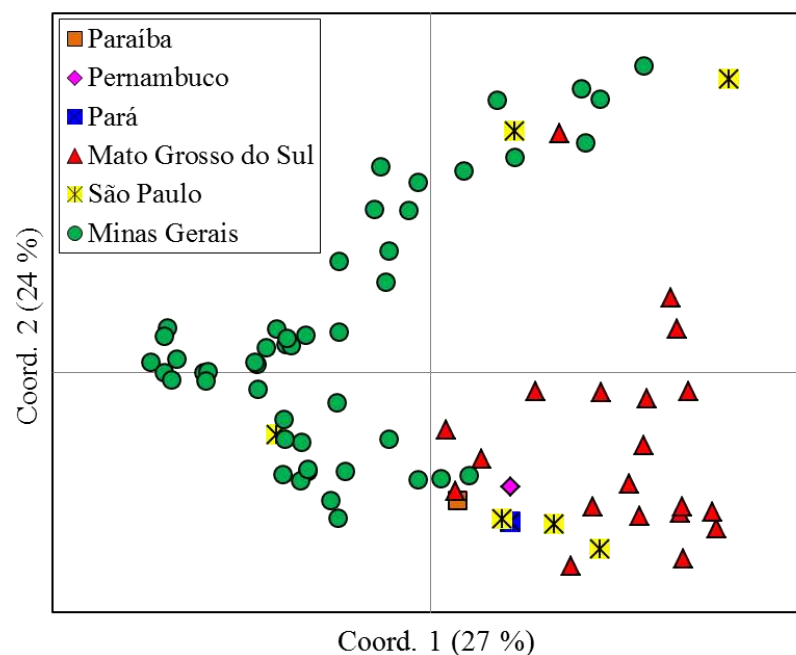
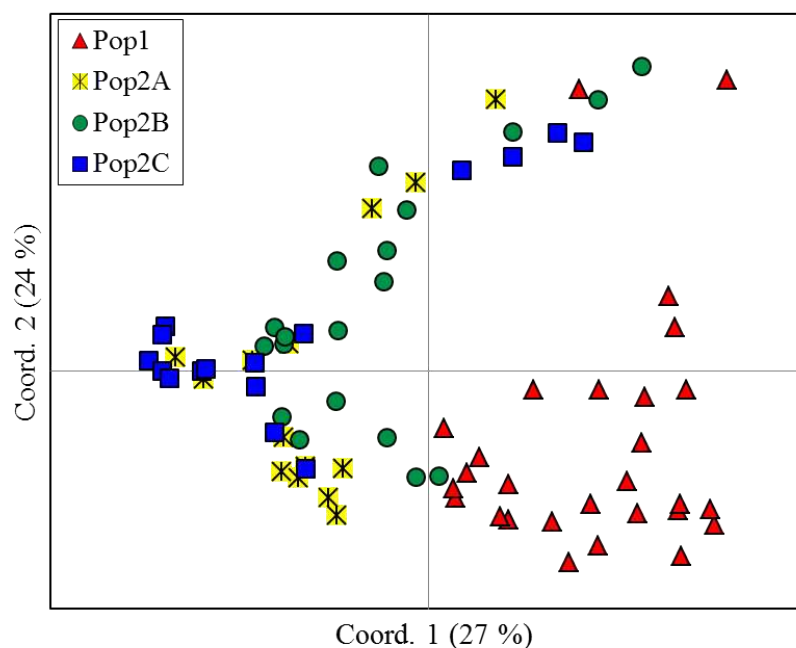


Figure 3. Neighbor-joining (NJ) analysis of 77 *Acrocomia aculeata* accessions based on Nei's genetic distance. (A) Unrooted neighbor-joining tree of 77 accessions; each colored branch represents one accession collected from corresponding subpopulation inferred by Structure. (B) NJ tree of the four inferred subpopulation.



(A)



(B)

Figure 4. Principal coordinate analysis of the 77 *Acrocomia aculeata* accessions. Diversity between the accessions according to the values of the first two principal coordinates with clusters identified by their provenance (Brazilian States) (A) and population structure analysis (B).

The genetic diversity was assessed per locus for each subpopulation (Table 3). The population parameter estimates indicate that Pop1 had the highest number of alleles per locus, the highest population-specific allele number, and the highest number of rare alleles, whereas Pop2B had the highest gene diversity, the highest number of alleles

with frequency greater than 0.05, the highest observed heterozygosity and the highest PIC value. Pop2B and Pop2C have approximately four and three alleles/locus, respectively, with a frequency that is distinctly different from zero. The mean observed heterozygosity levels were consistently lower than the expected panmictic proportions. The mean observed heterozygosity ranged from 0.31 for Pop1 to 0.48 for Pop2B with an average of 0.38.

Table 3. Molecular diversity of model-based subpopulations of *Acrocomia aculeata* as inferred by Structure: sample size (SS), average number of alleles per locus (Na), average effective number of alleles per locus (Ne), average number of different alleles with a frequency greater or equal to 0.05 (N_{Fa}), gene diversity (GD), observed heterozygosity (Ho), number of rare alleles (Nr), number of private alleles per subpopulations (AP) and percentage of polymorphic loci (P).

Population	SS	Na	Ne	N_{Fa}	GD	Ho	P	PIC	AP	Nr
Pop1	28	7.17	3.25	3.83	0.61	0.31	100	0.58	7	20
Pop2A	14	4.33	2.44	2.83	0.55	0.33	100	0.50	1	9
Pop2B	19	7.00	4.03	5.33	0.74	0.48	100	0.71	3	10
Pop2C	16	4.67	2.63	3.50	0.58	0.39	100	0.53	0	7
Mean	19.3	5.79	3.09	3.87	0.62	0.38	100	0.58	2.75	11.5

Relationships among populations and/or subpopulations inferred from Structure were identified by F_{st} analysis (Table 4). Pop1 showed a relatively high genetic differentiation when compared with all the other populations, where the pairwise F_{st} values ranged from 0.173 between population Pop1 and Pop2B to 0.314 between population Pop1 and Pop2C, with an overall average of 0.219. There was less genetic differentiation between the population Pop2B and Pop2C (F_{st} = 0.152).

Table 4. Genetic distances among four subpopulations of *Acrocomia aculeata* identified by Structure.

	Pop1	Pop2A	Pop2B	Pop2C
Pop1	-			
Pop2A	0.297	-		
Pop2B	0.173	0.157	-	
Pop2C	0.314	0.219	0.152	-

Pairwise F_{st} (below diagonal) calculated based on allele frequencies.

The hierarchical partitioning of allele frequency variation using the nested Analysis of Molecular Variance (AMOVA) indicated that 20.58% of total genetic variance was explained among the subpopulations, 23.90% was explained within individuals, and the remaining 55.52% was explained within subpopulations (Table 5).

Table 5. Analysis of molecular variance (AMOVA) of *Acrocomia aculeata* accessions among the four subpopulations identified by Structure.

Source of variation		Sum of squares	Percentage of variation (%)
Within populations	Pop1	124.3299	20.50
	Pop2A	56.7617	9.36
	Pop2B	94.2462	15.54
	Pop2C	61.4176	10.12
Within individuals	Pop1	40.0000	7.25
	Pop2A	25.0000	4.12
	Pop2B	45.0000	7.42
	Pop2C	31.0000	5.11
Among populations		124.8700	20.58
Total		606.6254	100.00

Discussion

Our study represents the first comprehensive assessment of population structure and genetic diversity of the species *A. aculeata* conserved in an ex situ collection. Knowledge of the genetic structure of a genebank offers insight into the amount and patterns of the distribution of the genetic diversity and thus enables efficient use and management of plant genetic resources. This is due to the nature of these allele banks, which retain the relevant allelic diversity that is necessary for plant breeding (Odong et al. 2011).

The collection of *A. aculeata* that was characterized by microsatellite loci in this study represents 30% of the accessions maintained in the Active Germplasm Bank of Macauba. The average number of alleles per locus (9 alleles) was higher than the 5 alleles per SSR locus reported by Nucci et al. (2008). A total of 47 polymorphic loci (average 4.2 fragments) per primer RAPD was found by Oliveira et al. (2012). The abundant allelic variation found indicates a high level of genetic diversity among the macaw palm accessions, which is possibly due to the large representativeness of the collection that is made up of materials from different eco-geographic regions. This

species has a high capacity to withstand heterogeneous environments (Ciconini et al. 2013). Thus, the large number of alleles and PIC reflect adaptation strategies in response to environments with contrasting climate and soil characteristics.

The population structure analysis based on Structure revealed the presence of four clusters among the studied accessions of *A. aculeata*. This observation is consistent with the grouping observed by the NJ hierarchical clustering method (Figure 3A, B) and thus confirms the existence of these four populations in the germplasm bank (BAG – Macauba). Analogously to the results based on the geographical origin and Structure analyses, the PCoA (Figure 4A, B) also showed remarkable subdivision between populations composed of genotypes from Mato Grosso do Sul (Pop1) and Minas Gerais (Pop2A, Pop2B and Pop2C). This discrimination can be explained by a possible specialization process, or even speciation, in progress (lineage splitting event) in the population Pop1 of Mato Grosso do Sul, as reinforced by the high degree of inferred ancestry ($\geq 90\%$) of accessions from this provenance. Moreover, there was a large genetic differentiation of Pop1 when compared with other populations ($F_{st} = 0.173$ to 0.314). This phenomenon is feasible considering that the Pantanal of Mato Grosso do Sul, the location of the accessions' mother trees of the germplasm bank, comprises a sedimentary basin surrounded by a large hydrographic network (Willink et al. 2000) and plateaus with altitudes between 600–700 meters. This topography may, as noted by Eiserhardt et al. (2011), strongly influence the distribution and diversity of plant species.

Added to the terrain barriers, the dispersal mode of both seed and pollen grains of *A. aculeata* are factors that can contribute to the isolation of the Pop1 population. When combined, these abiotic and biotic components can lead to the formation of populations with a high degree of relatedness, as with Pop1. According to Abreu et al. (2012), the relevant indices of inbreeding (endogamy) within *A. aculeata* is primarily the result of a seed dispersal restriction. Because the fruits are considerably heavy, their dispersal is mostly performed by gravity, thus, remaining close to the mother-tree, and mammals, which make the process dependent on a tuned synchronization between the occurrence and abundance of the vertebrates with the fruits' availability (Eiserhardt et al. 2011). In addition, restricted pollen dispersal, due to the limited flight range of pollinators, may have a considerable effect in the formation of an endogamous population.

The main pollinators of *A. aculeata* are the beetles *Andranthobius* sp. (Curculionidae), *Mystrops* cf *Mexicana* (Nitidulidae) and *Cyclocephala forsteri* (Scarabaeidae) (Scariot et al. 1991). The territory covered by these pollinators is small,

most likely due to their gregarious and foraging behavior, which prompt the insects to stay within clusters of individual *A. aculeata* in the same locality (Abreu et al. 2012). Young (1986) reported that *Cyclocephala* sp. fly at an average of 80 m between inflorescences, which is a short distance, to assure effective exchange of genetic material. However, studies on the behavior of pollinators of *A. aculeata* need to be further elucidated. Wind pollination plays a secondary role in the pollination of the macaw palm, as observed in several species of palms, such as the peach palm (*Bactris gasipaes*) (Scariot et al. 1991; Mora-Urpí et al. 1997). Therefore, geographical factors associated with the dispersion of pollen and seed become potential barriers to gene flow in *A. aculeata* accessions of the Pantanal of Mato Grosso do Sul and, therefore, create a favorable scenario for speciation.

In Brazil, no consensus exists on the number of *Acrocomia* species present in the territory. Specifically, the existence of *A. hassleri*, an herbaceous palm endemic to the Brazilian Midwest, and the arboreal *A. aculeata*, which has many synonyms including *A. totai*, *A. intumescens*, *A. esclerocarpa*, *A. media*, etc. is known (<http://www.tropicos.org>; <http://www.palmweb.org>; <http://www.kew.org>; <http://www.nybg.org/science>). Upon analyzing the genetic structure of eleven natural populations of *A. aculeata*, Abreu et al. (2012) reported that the population “Piquerobi” from São Paulo State, which is also located near the border of the Mato Grosso do Sul State, had 95% of their inferred individuals grouped in a single cluster with a very low levels of mixed genomes. According to these authors, who accept the existence *A. totai* as a species that is different from *A. aculeata*, the isolation and the location of this specific population on the geographical border between the two species may have led to these results. However, only works on phylogeny may clarify the assumption of whether they are different species.

Countless studies have made use of molecular tools to infer the likely center of diversity of many cultivated species: soybean (Dong et al. 2004), foxtail millet (Wang et al. 2012), pigeonpea (Kassa et al. 2012), oil palm (Maizura et al. 2006), coconut (Gunn et al. 2011), and oriental mustard (Chen et al. 2013). Among the macaw palm inferred populations, three gene pools were identified in the State of Minas Gerais (Pop2A, Pop2B and Pop2C) and constituted more than 88% of accessions from this State with a high degree of inferred ancestry ($\geq 70\%$) for only one of the populations. Brazil has long been recognized as one of the main centers of origin and genetic diversity of plants (Vavilov 1926; Darlington and Janaki Ammal 1945). The region with the greatest genetic diversity is generally regarded as the core of a species origin (Maxted et al.

1997; Dong et al. 2004). Hence, according to data obtained from accessions of the ex situ collection, our interpretation is that the State of Minas Gerais is a center of genetic diversity for *A. aculeata* (Figure 1B). This explanation is supported not only by the number of gene pools detected but also by the estimated population parameters (Table 3), such as the high values of allele richness, the number of alleles of high frequency, the genetic diversity, the number of rare alleles and the PIC. In addition, the veracity of this claim is reinforced by the Wright's index of genetic differentiation (F_{st}) and mainly by the AMOVA, which indicated a high allelic variation within and between populations (Table 4 and Table 5). The molecular variation of individuals within each of our study populations (> 4%) was superior to the two centers of diversity suggested by Wang et al. (2012) for foxtail millet (<0.5%). Morcote-Rios and Bernal (2001) reported that South America is the probable center of origin for *A. aculeata*. According to the authors, this hypothesis is supported by the fact that the oldest fossils of *A. aculeata* were found in this part of the continent and by the presence of the endemic species *A. hassleri* being restricted to the Cerrado (Brazilian savanna-like ecosystem).

The São Francisco River, which has its source in the central region of the Minas Gerais State, most likely plays a major role in seed dispersal of the macaw palm, as indicated by the remnants of the gene pools (Pop2A, Pop2B and Pop2C) throughout the State (Figure 1B). However, a great variability among the accessions reflects the undomesticated stage in which the species is found. These data suggest that the accessions of the Minas Gerais provenance are a promising sources of useful alleles to start a breeding program. In addition to this broad genetic base, samples harvested from the accessions' matrices showed that bunch production and oil yield were higher than the representatives from Mato Grosso do Sul (Ciconini et al. 2013).

Our results clearly indicate that the high genetic variability observed in the germplasm collection (BAG – Macauba) has a potential relationship with the provenance of the accessions, which is possibly due to the selective pressure of each environment in association with geographic barriers that prevent gene flow between populations. Furthermore, the morphological diversity and the fact that this species is scattered throughout Brazil reveal the high phenotypic plasticity that allowed the species to explore a wide variety of eco-regions. This conjuncture explains why *A. aculeata* is a rustic palm that is resistant to low water availability, variations in temperature and light and adaptable to different soil types (Arkcoll 1990, Moura et al. 2008; Motoike and Kuki 2009; Abreu et al. 2011).

Understanding the diversity center of *A. aculeata* is of high relevance for the management of germplasm and breeding programs. In the search for variability, the center of diversity is, ultimately, the main repository of genes or alleles for crop development, as predicted by Vavilov's theory. However, more research needs to be conducted with a larger number of genotypes and molecular markers and even the use of specific primers for nuclear gene regions and chloroplast DNA (cpDNA). These experiments could provide additional information about the origin, dispersion and evolution of the center of diversity or even lead to the discovery of new gene pools.

Funding

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES-Proex: #32002017004P2); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq: #143353/2011-0) and Petróleo Brasileiro S.A. (PETROBRAS: #0050.0061.571.109).

Acknowledgments

This study is part of the Doctoral Thesis (Genetic and Breeding Program — Universidade Federal de Viçosa — UFV) of the first author. The authors thank PhD. Marcos Deon Vilela de Resende and PhD. Eveline Teixeira Caixeta for the important comments, suggestions and constructive criticism that improved the manuscript. We also thank Dr. F. Andrew Jones and anonymous reviewers for their many helpful comments on this manuscript.

References

- Abdalla AL, Silva Filho JC, Godoy AR, Carmo CA, Eduardo JLP. 2008. Utilização de subprodutos da indústria de biodiesel na alimentação de ruminantes. *R Bras Zootec.* 37:260–258.
- Abreu IS, Carvalho CR, Carvalho GMA, Motoike SY. 2011. 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.* 59:149–155.
- 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* (Arecaceae). *Genet Mol Biol.* 35:119–121.
- Aradhya MK, Stover E, Velasco D, Koehmstedt A. 2010. Genetic structure and differentiation in cultivated fig (*Ficus carica* L.). *Genetica.* 138:681–694.
- Arkcoll D. 1990. New crops from Brazil. In: Janick, J., Simon, E.J. (Eds.), *Advances in New Crops*. Timber Press, Portland, OR, p. 367–371.
- Baker CS. 2013. Journal of heredity adopts joint data archiving policy. *J Hered.* 104:1.
- Belaj A, Dominguez-Garcia MC, Atienza SG, Urdiroz NM, Rosa RD, Satovic Z, Martin A, Kilian A, Trujillo I, Valpuesta V, Rio CD. 2012. Developing a core collection of olive (*Olea europaea* L.) based on molecular markers (DARts, SSRs, SNPs) and agronomic traits. *Tree Genet Genomes.* 8:365–378.
- Bora PS, Rocha RVM. 2004. Macaíba palm: fatty and amino acids composition of fruits. *Cienc Tecnol Aliment.* 4:158–162.
- Bracco M, Lia VV, Gottlieb AM, Hernández JC, Poggio L. 2009. Genetic diversity in maize landraces from indigenous settlements of Northeastern Argentina. *Genetica.* 135:39–49.
- Chen S, Wan Z, Nelson MN, Chauhan JS, Redden R, Burton WA, Lin P, Salisbury PA, Fu T, Cowling WA. 2013. Evidence from Genome-wide Simple Sequence Repeat Markers for a Polyphyletic Origin and Secondary Centers of Genetic Diversity of *Brassica juncea* in China and India. *J Hered.* 104:416–427.
- 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.
- Ciconini G. 2012. Caracterização de frutos e óleo de polpa da macaúba dos biomas Cerrado e Pantanal do estado de Mato Grosso do Sul, Brasil. PhD Thesis, University Católica Dom Bosco, Brazil. 128p.
- Corley RHV, Tinker PB. 2003. The oil palm, 4th edition. Blackwell Science, Oxford, United Kingdom. 562p.
- Darlington CD, Janaki Ammal EK. 1945. Chromosome atlas of cultivated plants. Allen & Unwin, London.
- Dong YS, Zhao LM, Liu B, Wang ZW, Jin ZQ, Sun H. 2004. The genetic diversity of cultivated soybean grown in China. *Theor Appl Genet.* 108:931–936.
- Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour.* 4:359–361.
- Eiserhardt WL, Svenning JC, Kissling WD, Balslev H. 2011. Geographical ecology of the palms (Arecaceae): determinants of diversity and distributions across spatial scales. *Ann Bot.* 108:1391–1416.
- 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.
- Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.* 131:479–491.
- Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics.* 164:1567–1587.
- Gunn BF, Baudouin L, Olsen KM. 2011. Independent Origins of Cultivated Coconut (*Cocos nucifera* L.) in the Old World Tropics. *PLoS ONE.* 6:e21143.
- Harlan JR. 1971. Agricultural origins: centers and noncenters. *Science.* 174:468–474.
- Hauser LA, Crovello TJ. 1982. Numerical analysis of genetic relationships in *Thelypodieae* (Brassicaceae). *Syst Bot.* 7:249–268.

- Henderson A, Galeano G, Bernal R. 1995. 'Field guide to the palms of the Americas.' (Princeton University Press: New Jersey).
- Hiane PA, Baldasso PA, Marangoni S, Macedo MLR. 2006. Chemical and nutritional evaluation of kernels of bocaiúva, *Acrocomia aculeata* (Jacq.) Lodd. *Cienc Tecnol Aliment*. 26:683–689.
- Hiane PA, Ramos Filho MM, Ramos MIL, Macedo MLR. 2005. Bocaiuva, *Acrocomia aculeata* (Jacq.) Lodd., pulp and kernel oils: Characterization and fatty acid composition. *Braz J Food Technol*. 8:256–259.
- Hijmans RJ, Guarino L, Mathur P. 2012. DIVA-GIS Version 7.5. Manual. <<http://www.diva-gis.org>> (accessed 03.05.13).
- Jeffery MI. 2002. Bioprospecting : access to genetic resources and benefit-sharing under the convention on biodiversity and the Bonn Guidelines. *SJICL*. 6:747–808.
- Kassa MT, Penmetsa RV, Carrasquilla-Garcia N, Sarma BK, Datta S, Upadhyaya HD, Varshney RK, von Wettberg EJB, Cook DR. 2012. Genetic patterns of domestication in pigeonpea (*Cajanus cajan* (L.) Millsp.) and wild *Cajanus* relatives. *PLoS One*. 7:e39563.
- Kusdiana D, Saka S. 2001. Methyl esterification of free fatty acids of repeseed oil as trated in supercritical methanol. *J Chem Eng Jpn*. 34:383–387.
- Lanes ECM, Nick C, Kuki KN, Freitas RD, Motoike SY. 2013. Genomic DNA isolation of *Acrocomia aculeata* (Arecaceae) from leaf and stipe tissue samples for PCR analysis. *Gen Mol Res*. 12:3905–3911.
- Liu KJ, Muse SV. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*. 21:2128–2129.
- Lorenzi H, Souza HM, Costa JTM, Cerqueira LSC, Ferreira E. 2004. *Palmeiras brasileiras e exóticas cultivadas*. [Brazilian palm trees and exotic cultivated]. Instituto Plantarum, Nova Odessa.
- Maguire TL, Peakall R, Saenger P. 2002. Comparative analysis of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae) detected by AFLPs and SSRs. *Theor Appl Genet*. 104:388–398.
- Maizura I, Rajanaidu N, Zakri AH, Cheah SC. 2006. Assessment of genetic diversity in oil palm (*Elaeis guineensis* Jacq.) using Restriction Fragment Length Polymorphism (RFLP). *Genet Resour Crop Evol*. 53:187–195.

- Maxted N, Ford-Llyod BV, Hawkes JG. 1997. Plant Genetic Conservation, the in situ Approach. Chapman & Hall, London.
- Mora-Urpí J, Weber JC, Clement CR. 1997. Peach palm. *Bactris gasipaes* Kunth. Promoting the conservation and use of underutilized and neglected crops. 20. Institute of Plant Genetics and Crop Plant Research, Gatersleben/ International Plant Genetic Resources Institute, Rome, Italy.
- Morcote-Ríos G, Bernal R. 2001. Remains of Palms (Palmae) at Archaeological Sites in the New World: A Review. *Bot Rev.*67:309–350.
- Motoike SY, Kuki KN. 2009. The potential of macaw palm (*Acrocomia aculeata*) as source of biodiesel in Brazil. *IRECHE*. 1:632–635.
- Moura EF, Ventrella MC, Motoike SY, Sá Júnior AQ, Carvalho M, Manfio CE. 2008. Histological study of somatic embryogenesis induction on zygotic embryos of macaw palm (*Acrocomia aculeata* (Jacq.) Lodd. Ex Martius). *Plant Cell Tiss Organ Cult.* 95:175–184.
- Moura EF, Ventrella MC, Motoike SY. 2010. Anatomy, histochemistry and ltrastructure of seed and somatic embryo of *Acrocomia aculeata* (Arecaceae). *Sci Agric.* 67:399–407.
- Nei M, Takezaki N. 1983. Estimation of genetic distances and phylogenetic trees from DNA analysis. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, University of Guelph, Guelph, Ontario, Canada 21:405–412.
- Nucci SM, Azevedo-Filho JA, Colombo CA, Priolli RHG, Coelho RM, Mata TL, Zucchi MI. 2008. Development and characterization of microsatellites markers from the macaw. *Mol Ecol Resour.* 8:224–226.
- Odong TL, Heerwaarden J, Jansen J, Hintum TJL, Eeuwijk FA. 2011. Determination of genetic structure of germplasm collections: are traditional hierarchical clustering methods appropriate for molecular marker data? *Theor Appl Genet.* 123:195–205.
- Oliveira DA, Melo Júnior AF, Brandão MM, Rodrigues LA, Menezes EV, Ferreira PRB. 2012. Genetic diversity in populations of *Acrocomia aculeata* (Arecaceae) in the northern region of Minas Gerais, Brazil. *Gen Mol Res.*11:531–538.
- Orloci L. 1978. *Multivariate analysis in vegetation research.*, The Hague: Dr. W. Junk B. V.

- Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel Population genetic software for teaching and research. *Mol Ecol Notes*. 6:288–295.
- Pires TP, Souza ES, Kuki KN, Motoike SY. 2013. Ecophysiological traits of the macaw palm: A contribution towards the domestication of a novel oil crop. *Ind Crop Prod*. 44:200–210.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure from multilocus genotype data. *Genetics*. 155:945–959.
- Rambaut A. 2012. Tree Figure Drawing Tool Version 1.4.0. Institute of Evolutionary Biology. University of Edinburgh, UK.
- Renau-Morata B, Nebauer SG, Sales E, Allaniguillaume J, Caligari P, Segura J. 2005. Genetic diversity and structure of natural and managed populations of *Cedrus atlantica* (Pinaceae) assessed using a random amplified polymorphic DNA. *Am J Bot*. 92:875–884.
- Röser M, Johnson MAT, Hanson L. 1997. Nuclear DNA amounts in palms (Arecaceae). *Bot Acta*. 110:79–89.
- Scariot A, Lleras E, Hay JD. 1991. Reproductive biology of the palm *Acrocomia aculeata* in Central Brazil. *Biotropica*. 23:12–22.
- Scariot A, Lleras E. 1995. Flowering and fruiting phenologies of the palm *Acrocomia aculeata*: patterns and consequences. *Biotropica*. 27:168–173.
- Silva JC, Barrichelo LEG, Brito JO. 1986. Endocarpos de babaçu e macaúba comparados a madeira de *Eucalyptus grandis* para a produção de carvão vegetal. *IPEF*. 34:31–34.
- Stajner N, Satovic Z, Cerenak A, Javornik B. 2008. Genetic structure and differentiation in hop (*Humulus lupulus* L.) as inferred from microsatellites. *Euphytica*. 161:301–311.
- Sumathi S, Chai SP, Mohamed AR. 2008. Utilization of oil palm as a source of renewable energy in Malaysia. *Renew Sust Energ Rev*. 12:2404–2421.
- Toro MA, Caballero A. 2005. Characterization and conservation of genetic diversity in subdivided populations. *Philos Trans R Soc Lond B Biol Sci*. 360:1367–1378.
- Vavilov NI. 1926. Studies on the origin of cultivated plants. *Appl Bot Plant Breed*. 26:1–248.

- Von Lisingen L, Cervi AC. 2009. *Acrocomia aculeata* (Jacq.) Lodd ex Mart., nova ocorrência para a flora do Estado do Paraná. *Acta Biol Parana*. 38:187–192.
- Yang X, Gao S, Xu S, Zhang Z, Prasanna BM, Li L, Li J, Yan J. 2011. Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. *Mol Breeding*. 28:511–526.
- Young HJ. 1986. Beetle Pollination of *Dieffenbachia Longispatha* (Araceae). *Am J Bot* 73:931–944.
- Wang C, Jia G, Zhi H, Niu Z, Chai Y, Li W, Wang Y, Li H, Lu P, Zhao B, Diao X. 2012. Genetic Diversity and Population Structure of Chinese Foxtail Millet [*Setaria italica* (L.) Beauv.] Landraces. *G3-Genes Genom Genet*. 2:769–777.
- Willink PW, Chernoff B, Alonso LE, Montambault JR, Lourival R. 2000. A biological assessment of the aquatic ecosystems of the Pantanal, Mato Grosso do Sul, Brasil. *RAP Bulletin of Biological Assessment* 18. Conservation International, Washington, DC.
- <http://apps.kew.org/wcsp/qsearch.do> (last accessed 20/04/2013).
- http://www.palmweb.org/cdm_dataportal/taxon/d58c0b44-29f8-4071-aa49-32baa1852 (last accessed 20/04/2013).
- <http://sciweb.nybg.org/Science2/hcol/lists/arecaceae.html> (last accessed 20/04/2013).
- <http://www.tropicos.org/Name/2401428?tab=synonyms> (last accessed 20/04/2013).

CAPÍTULO 2

Este manuscrito segue as normas de formatação do periódico Journal of Heredity

Sistema de reprodução e estimativa de depressão por endogamia em macaúba, *Acrocomia aculeata* (Arecaceae) utilizando dados quantitativos e moleculares

Laboratório de Biotecnologia e Melhoramento Vegetal, Universidade Federal de Viçosa, 36570-000, Viçosa, MG, Brasil.

Departamento de Fitotecnia, Universidade Federal de Viçosa, 36570-000, Viçosa, MG, Brasil.

O endereço postal para correspondência a Éder C. M. Lanes encontra-se acima, ou e-mail: edercml@gmail.com

Resumo

A *Acrocomia aculeata* é uma das mais promissoras fontes de óleo vegetal das Américas. A fim de definir a melhor estratégia de melhoramento e conservação genética da espécie, determinou-se o modo como são constituídas as recombinações gênicas a cada evento reprodutivo, e as estimativas de depressão por endogamia (δ) em caracteres vegetativos. O sistema de cruzamento foi analisado com genótipos multilocos de nove marcadores microssatélites determinados em 157 progênies de 19 árvores matrizes. Os níveis de depressão por endogamia foram estimados apenas nas famílias com progênie de autofecundação indicado pela análise de paternidade. Os parâmetros do sistema de cruzamento indicam que *A. aculeata* é predominantemente de cruzamento ($\hat{t}_m = 0,986$), provavelmente, com sistema parcialmente auto-incompatível. Contudo, pode haver variações para auto-incompatibilidade, indicando uma estratégia para sobreviver e persistir em habitats isolados. Detectou-se desvios nos cruzamentos aleatórios ($N_{e(v)} = 1,937$), em decorrência das altas taxas de cruzamentos biparentais e entre indivíduos aparentados, sendo a maior fração das progênies compostas por irmãos-completos (54,8%). Sendo assim, a conservação genética ex situ da *A. aculeata*, deve priorizar o número de árvores matrizes para a coleta de sementes, em pelo menos 77 indivíduos. Quanto à depressão por endogamia, a altura de planta foi significativamente afetada, com níveis percentuais ($\delta > 19\%$). A confirmação do alto desempenho (alto valor per se) ao mérito genético de progênies endogâmicas será de grande utilidade para o programa de melhoramento genético da *A. aculeata*, permitindo a adoção de métodos de melhoramento refinados, como a Seleção Recorrente Recíproca Individual (SRRI).

Introdução

As palmáceas são elementos chave na estrutura e funcionamento dos ecossistemas terrestres tropicais (Zona & Henderson, 1989; Galetti et al., 1999; Salm et al., 2005; Genini, 2009; Barfod et al., 2011). A família Arecaceae possui aproximadamente 2.450 espécies (Barfod et al., 2011), dentre estas, a macaúba (*Acrocomia aculeata* (Jacq.) Lodd. ex Martius), uma planta endêmica das Américas do Sul e Central (Henderson et al., 1995) que vem despertando grande interesse da comunidade científica devido ao seu enorme potencial como fonte de energia renovável.

A *A. aculeata* é um “recurso genético modelo”, uma vez que possibilita uma cadeia produtiva totalmente sustentável, com potencialidades para atender os principais desafios do planeta na atualidade. As estimativas de produtividade da *A. aculeata* (acima de 6200 kg/ha⁻¹ de óleo, Wandeck & Justo, 1988) e sua dispersão por quase todo o território do Brasil, coloca o país em uma posição estratégica e de referência em sustentabilidade no cenário mundial. A extração de óleos da polpa e da amêndoa atende ambas as indústrias de biocombustível e de fármaco-cosmética (Hiane et al., 2005; Bora & Rocha 2004; Coimbra & Jorge, 2012; Ciconini et al., 2013; Berton et al., 2014); mas também produz uma gama de resíduos sólidos que podem ser aproveitados para a geração de co-produtos em outros setores energéticos como carvão vegetal, etanol de segunda geração e bio-querosene para aviação (Silva et al., 1986; Gonçalves et al., 2013, Lanes et al., 2014). A espécie apresenta plasticidade fenotípica que lhe permite adaptar-se a condições edafoclimáticas diversas, inclusive a baixa oferta hídrica (Motoike & Kuki, 2009; Abreu et al., 2011). Desta forma, ela pode ser integrada em diversos sistemas de produção, e dentre estes o agrosilvipastoril (Ciconini et al., 2013). A inserção desta espécie apenas em pastagens já existentes pode ajudar a conter a expansão de novas áreas agrícolas, protegendo as florestas, sequestrando carbono (CO₂) e restaurando a biodiversidade de áreas degradadas (Lanes et al., 2014), exercendo serviços ambientais inestimáveis.

A *A. aculeata* (2n = 30, Abreu et al., 2011) é uma espécie monoica, com numerosas flores masculinas e femininas dispostas em partes distintas da mesma inflorescência (ou seja, andrógenas). A inflorescência apresenta uma marcada protoginia, em que as flores femininas entram em antese antes das flores masculinas (Scariot et al., 1991). Embora seja um mecanismo que favoreça a polinização cruzada, a presença de várias inflorescências em um único indivíduo permite a autofecundação. As plantas hermafroditas podem combinar autofecundações com cruzamentos aleatórios ou

sistemáticos (cruzamentos correlacionados), gerando uma descendência composta por diferentes graus de parentescos. O padrão ou mesmo as variações do sistema de cruzamento dentro de uma espécie é determinada por fatores reprodutivos e ecológicos (Núñez-Avellaneda et al., 2005; Henderson, 1986; Khanduri et al., 2013; Murawski & Hamrick, 1991; Fuchs & Hamrick, 2011), ou por fatores genéticos inerentes, como um sistema de auto-incompatibilidade (Goodwillie et al., 2005).

Várias espécies arbóreas hermafroditas são auto-incompatíveis ou parcialmente auto-incompatíveis (Bawa, 1974; Zapata & Kalin-Arroyo, 1978; Bullock, 1985); sendo que variações no sistema de auto-incompatibilidade são frequentes em muitas espécies de Arecaceae, como *Euterpe edulis* (Gaiotto et al., 2003), *Astrocaryum aculeatum* (Ramos et al., 2011), *Oenocarpus bataua* (Ottewell et al., 2012) e *Butia eriospatha* (Nazareno & Reis, 2012). Porém, a autofecundação pode representar uma boa estratégia para colonizar e persistir em um hábitat, principalmente na ausência de polinizadores ou devido ao isolamento de pequenas populações. Contudo, sucessivas gerações de endogamia podem resultar na redução da diversidade genética e afetar a capacidade destas populações em lidar com as pressões ambientais.

Entretanto, a magnitude depressiva da endogamia atua de forma diferente entre as características, variando conforme o sistema de cruzamento, a fase do ciclo de vida de uma espécie, e particularmente pelas informações implícitas no controle genético do carácter (Lande & Schemske, 1985; Charlesworth et al., 1990; Husband & Schemske, 1996; Thiele et al., 2010). Para fins de melhoramento genético, informações a respeito da base genética da depressão por endogamia de caracteres quantitativos ampliam a eficiência do processo seletivo de genótipos superiores. O uso de medidas vegetativas como critérios de seleção têm sido sugeridos em vários estudos com palmáceas (Uzzo et al., 2002; Luyindula et al., 2005; Agyei-Dwarko et al., 2012; Farias Neto et al., 2013).

O conhecimento sobre o sistema de cruzamento da *A. aculeata* se resume a sua biologia reprodutiva (Scariot et al., 1991) e a marcadores moleculares aplicados em populações naturais (Abreu et al., 2012; Nucci et al., 2008). Estudos de parâmetros de endogamia e estrutura genética, como também, estimativas de depressão endogâmica em caracteres vegetativos para a espécie são inexistentes. Em face ao exposto, as seguintes questões foram levantadas: (i) A espécie possui um sistema misto de reprodução ou é estritamente de polinização cruzada (alogamia)? (ii) Quais são as estimativas dos parâmetros de cruzamento, incluindo cruzamento entre parentes, correlação de paternidade e número efetivo de doadores de pólen? (iii) Quais são as estimativas do parâmetro de coancestria e do tamanho efetivo da variância? (iv) Qual é

o número mínimo necessário de árvores matrizes para a coleta de sementes em um programa de conservação genética? (v) Quais são os níveis de depressão por endogamia sobre caracteres vegetativos na fase juvenil?

Material e Métodos

Banco de germoplasma

O Banco Ativo de Germoplasma de Macaúba (BAG - Macaúba), situado na Fazenda Experimental de Araponga, no município de Araponga, Minas Gerais / Brasil (20°40'1"S "latitude e 42°31'15"W longitude; altitude de aproximadamente 1000 m), clima Cwb pela classificação de Köppen. A BAG - Macaúba é o primeiro repositório oficial registrado pelo Conselho de Gestão do Patrimônio Genético Brasileiro (nº 084/2013-SECEX / CGEN) e destaca-se como um dos maiores bancos de germoplasma de *A. aculeata* da América Latina. Esta área pertence à Universidade Federal de Viçosa, e é mantido pelo Programa de Melhoramento Genético de Macaúba, coordenado pelo Departamento de Fitotecnia da Universidade Federal de Viçosa (UFV, Brasil).

Atualmente, o banco mantém uma coleção com mais de 300 famílias maternas (aproximadamente 1500 acessos), representativa de quase todas as regiões do Brasil. A inserção de novos acessos na coleção é um processo contínuo, onde todas as plantas são obtidas por propagação semínifera. As árvores nativas fornecedoras de sementes (referida como árvore materna ou árvore matriz) são georreferenciadas usando um GPS com precisão ± 7 m (Garmin, Atchison, KS, EUA). Durante a época de frutificação, os frutos maduros são coletados, e as sementes são extraídas e pré-germinadas, aplicando o protocolo desenvolvido especificamente para a espécie (Motoike et al., 2007). Quando as plântulas (referidas como progênies) atingem os 10 meses de idade, estes são transferidas para o campo. O termo "família materna" refere-se aos indivíduos descendentes de uma mesma árvore matriz, e que são conservados na BAG - Macaúba. Para obter as práticas de manejo de campo ver Pires et al. (2013).

Para a genotipagem foram coletados tecidos vegetais das árvores matrizes georreferenciadas (plantas adultas nativas) e de suas respectivas progênies (plantas na fase juvenil mantidas na BAG - Macaúba). A fonte de material para extração de DNA foi composta por: i) amostras de tecido de estipe (caule) de árvores matrizes (n = 27), que foram coletadas e armazenadas em tubos de Eppendorfs, e acondicionados em nitrogênio líquido, de acordo com Lanes et al. (2013); e ii) folíolos de folhas saudáveis não senescentes de 157 progênies provenientes de 19 das 27 matrizes (média de 8,3 progênie por árvore materna), estes foram coletados no banco de germoplasma (Figura

1S). Os folíolos foram excisados das ráquis e imediatamente colocado em nitrogênio líquido, e posteriormente, liofilizados durante um período de 72 h e, em seguida, armazenados a -20°C.

Extração de DNA genômico e análise de microssatélites

As amostras (tecidos de estipes e de folíolos) foram enviadas e processadas no Laboratório de Biotecnologia e Melhoramento Vegetal do Departamento de Fitotecnia - DFT / UFV, em Viçosa, Minas Gerais (MG), Brasil. O DNA genômico foi extraído com brometo de hexadeciltrimetilamônio (CTAB), conforme o protocolo modificado por Lanes et al. (2013). A genotipagem foi realizada com um total de nove locos microssatélites, seis locos desenvolvidos para a *A. aculeata* por Nucci et al. (2008) e três locos de *Elaeis guinensis* (Singh et al., 2007; Billotte et al., 2005).

As amplificações em PCR (Polymerase Chain Reaction) foram conduzidas em termociclador Veriti™ de 96 poços (Applied Biosystems). O volume final de cada reação foi de 20 µL contendo 30 ng de DNA genômico, tampão de reação 1x (20 mM de Tris-HCl, pH 8,4; 50 mM KCl - Invitrogen), 250 µM de cada dNTP (Invitrogen), 0,15 µM de primer SSR (Forward e Reverse), 3 mM de MgCl₂ (Invitrogen) e 1 U de Taq DNA polymerase (Invitrogen). O termociclador foi programado para uma etapa inicial de pré-desnaturação de 94 ° C por 5 min, seguido por 30 ciclos de 94 ° C por 1 min, 55 ° C por 1 min, e 72 ° C por 1 min. Por fim, realizou-se uma extensão final de 72 ° C por 8 min. Os fragmentos de DNA resultantes da amplificação foram separados por eletroforese vertical em gel desnaturante de poliacrilamida (6,0 %, w / v) [acrilamida / bis-acrilamida (19:1) (Bio-Rad), 7,5 M de uréia, TBE 5,0X] usando o sistema eletroforético "Sequi-Gen GT" (Bio-Rad). Os géis foram visualizados após a coloração com nitrato de prata (Brito et al., 2010).

Análise da diversidade genética

A fim de comparar a diversidade genética entre as fases fenológicas, foram estimados para cada loco SSR de árvores adultas (n = 27) e de árvores juvenis (n = 157), o número total e médio de alelos por loco (k), heterozigose observada (Ho) e heterozigosidade esperada (He), utilizando o programa GenAlEx 6.5 (Peakall & Smouse, 2006). Para comparar os valores médios de k, Ho, e He entre árvores adultas e juvenis, o intervalo de confiança de 95% destes parâmetros foram calculados a partir do erro padrão (média ± 1,96 x SE). A endogamia entre árvores adultas e juvenis foi determinada pelo índice de fixação (F) e sua significância estatística (isto é, se F é significativamente diferente de zero). A significância estatística dos valores de F foram

testados por 10000 permutações (alelos entre indivíduos), aplicando uma correção de Bonferroni (95%, $\alpha = 0,05$). Nas progênes (F_{of}), a endogamia foi estimada pelo índice de fixação individual utilizando como frequências alélicas de referência as frequências estimadas nas árvores adultas pelo o programa SPAGeDI versão 1.4 (Hardy & Vekemans, 2002).

Análise do sistema de cruzamento

O sistema de reprodução da *A. aculeata* foi determinado a nível de famílias (cada família materna) e de população (conjunto de famílias maternas), com base nos modelos misto de reprodução (Ritland & Jain, 1981) e de cruzamentos correlacionados (Ritland, 1989) implementado no software MLTR 3.4 (Ritland, 2002). Os parâmetros estimados pelo método de máxima verossimilhança (Expectation-Maximization), admitindo-se a presença de alelos nulos, foram: coeficiente de endogamia nas árvores maternas ($\hat{F}_m$), taxa de cruzamento multiloco ($\hat{t}_m$), taxa de cruzamento uniloco ($\hat{t}_s$), taxa de cruzamento entre indivíduos aparentados ($\hat{t}_m - \hat{t}_s$), correlação de autofecundação ($\hat{f}_s$), correlação multilocos de paternidade ($\hat{r}_{p(m)}$). O intervalo de confiança de 95% foi calculado com base no erro-padrão obtido por 1000 reamostragens de bootstraps.

Com base nos parâmetros do sistema de cruzamento, outros parâmetros demográficos e genéticos foram estimados. O número efetivo de árvores polinizadoras (N_{ep}) foi calculado da correlação multiloco de paternidade ($N_{ep} = 1 / r_{p(m)}$) (Ritland, 1989). A proporção de progênes de irmãos de autofecundação ($\hat{P}_{IA}$), meios-irmãos ($\hat{P}_{MI}$), irmãos-completos ($\hat{P}_{IC}$) e irmãos de autofecundação e de cruzamento ($\hat{P}_{IAC}$) foi calculada, respectivamente, por: $\hat{P}_{IA} = \hat{s}^2$; $\hat{P}_{MI} = \hat{t}_m^2 (1 - \hat{r}_{p(m)})$; $\hat{P}_{IC} = \hat{t}_m^2 \hat{r}_{p(m)}$ e $\hat{P}_{IAC} = 2\hat{s}\hat{t}_m$, onde $\hat{s}$ é a taxa de autofecundação ($\hat{s} = 1 - \hat{t}_m$). O intervalo de confiança para estes parâmetros foram determinados com base nos limites superior e inferior do intervalo de confiança de 95% dos parâmetros do sistema de cruzamento.

O coeficiente médio de coancestria entre plantas dentro de progênes ($\hat{\theta}_{xy}$) foi obtido pelo coeficiente de correlação de parentesco entre plantas dentro de progênes ($\hat{r}_{xy}$), conforme derivado por Ritland (1989):

$$\hat{r}_{xy} = 0,25 (1 + \hat{F}_m) [4\hat{s} + (\hat{t}_m^2 + \hat{t}_m \hat{s} \hat{f}_s)(1 + \hat{r}_{p(m)})]$$

em que, $\hat{F}_m$ é o coeficiente de endogamia na geração parental, e $\hat{s}$ é a estimativa da taxa de autofecundação. Como em espécies diplóides o coeficiente de coancestria corresponde à metade do coeficiente de parentesco, tem-se que, ($\hat{\theta}_{xy} = \hat{r}_{xy} / 2$).

Para conhecer a representatividade genética das progênes, o tamanho efetivo de variância foi estimado com base no coeficiente de coancestria, $\hat{N}_{e(v)} = 0,5 / \hat{\theta}_{xy}$ (Cockerham, 1969). Ressalta-se que o valor mínimo que o coeficiente de coancestria pode atingir dentro de progênes é 0,125, o que corresponde ao parentesco composto puramente de meios-irmãos. Contudo, diferentes níveis de parentesco são esperados dentro de famílias: irmãos de autofecundação, meios-irmãos, irmãos-completos, e irmãos de autofecundação e de cruzamento. Portanto, é esperado que o coeficiente de coancestria apresente valores entre 0,125 e 1, o que corresponde a progênes de polinização aberta, incorporando diferentes tipos de parentesco. Além destes, o coeficiente de parentesco para árvores matrizes foi calculado utilizando o estimador proposto por J. Nason ($\hat{f}_{ij}$) descrito por Loiselle et al. (1995) implementado no software SPAGeDi 1.4 (Hardy & Vekemans, 2002).

O coeficiente de endogamia nas progênes ($\hat{F}_0$) foi inferido por meio do cálculo do índice de fixação. Para esta estimativa, valores de $\hat{F}_0$ negativos foram assumidos como zero (pois, o coeficiente de endogamia varia de 0 a 1). Concomitantemente, determinou-se o índice de fixação para cada família materna ($\hat{f}_x$). Os índices de fixação e seus respectivos intervalos de confiança foram obtidos por 30000 reamostragens de bootstraps, utilizando o programa GDA (Lewis & Zaykin, 2001). A endogamia nas progênes é composta por dois componentes: um devido a autofecundação $\hat{F}_s = 0,5 \hat{S}(1 + \hat{F}_m)$ (Barrett & Kohn, 1991), e um devido a cruzamentos entre aparentados $\hat{F}_{tm} - \hat{t}_s = \hat{F}_0 - \hat{F}_s$, de forma que, o coeficiente de endogamia total é $\hat{F}_0 = \hat{F}_s + \hat{F}_{tm} - \hat{t}_s$. De posse destas expressões, estimou-se a contribuição da autofecundação e a taxa de cruzamento entre parentes para a endogamia total das progênes ($\hat{F}_0$). Adicionalmente, estimou-se o coeficiente de endogamia em equilíbrio de Wright ($\hat{F}_{eq}$) para as progênes, usando como base a estimativa da taxa de cruzamento multiloco, $\hat{F}_{eq} = (1 - t_m)/(1 + t_m)$ (Fyfe & Bailey, 1951).

O número de árvores matrizes ($\hat{m}$) necessárias para a coleta de sementes foi calculado assumindo que o objetivo é reter na amostra total o tamanho efetivo de referência ($N_{e(referência)}$) de 150, pela seguinte expressão $\hat{m} = N_{e(referência)} / \hat{N}_{e(v)}$ (Sebbenn, 2006). Uma importante observação é que esta estimativa é baseada em duas premissas: (i) as árvores matrizes não são parentes e não se acasalam umas com as

outras; e (ii) que as árvores matrizes recebem pólen de diferentes conjuntos gênicos, ou seja, não há sobreposição no conjunto de pólen entre as diferentes árvores.

Para verificar a frequência alélica retida (ω_{af}) na população de referência ($N_{e(referência)}$), um intervalo de confiança construído para a frequência alélica foi determinado pela seguinte expressão proposta por Resende (2002):

$$\omega_{af} = p_0 \pm z \{ [p_0(1 - p_0)] / [2 N_e] \}^{0.5}$$

onde p_0 é a frequência paramétrica do alelo na população original; Z é o valor tabelado da distribuição normal, equivalendo a 1,96 para um intervalo de confiança de 95%; e o N_e é o valor do tamanho efetivo populacional de referência. Se o limite inferior do intervalo de confiança é diferente de zero, pode-se inferir que a frequência alélica (testada arbitrariamente) é capturada na população (Resende, 2002).

Análise de paternidade e depressão por endogamia

A identificação das progênies geradas de autofecundação foi determinada pela análise de paternidade usando software Cervus 3.0.3 (Kalinowski et al., 2007). As análises foram conduzidas a partir dos genótipos de 19 árvores matrizes e de suas progênies ($n=157$). O software Cervus calcula, para cada um dos locos, uma razão entre as probabilidades de duas hipóteses alternativas (likelihood ratio) para cada candidato a pai, de modo que a razão de máxima verossimilhança (likelihood ratio) é $LR = H_{(1)}/H_{(0)}$, onde a hipótese $H_{(1)}$ representa que o candidato é o verdadeiro pai, e $H_{(0)}$ representa que o candidato a pai é de ocorrência aleatória na população. Um LR total para cada candidato é calculado por multiplicação dos LRs de cada loco individual, assumindo-se segregação independente entre locos (locos em equilíbrio de ligação). O logaritmo natural ($\log_e$) desse LR total é denominado LOD score. Valores positivos de LOD score indicam que o candidato possui maior probabilidade de ser o verdadeiro pai do que um outro macho selecionado aleatoriamente. Nos casos em que dois ou mais candidatos a pai apresentam um LOD score positivo, uma estatística Δ é aplicada (Marshall et al., 1998).

A significância dos valores da estatística Δ foi obtida por simulações no software Cervus (Δ crítico), a partir das frequências alélicas observadas na população (baseada nos genótipos de todas as árvores adultas e de progênies) com o objetivo de gerar distribuições do Δ . O parental com o maior LOD score será assinalado como o pai se a diferença entre o seu LOD score e o do segundo candidato a parental for maior do que o Δ crítico. Mas, se o indicado a candidato a pai é a própria árvore materna, a progênie é considerada como de autofecundação. Os valores críticos de Δ foram simulados

considerando uma proporção de árvores masculinas amostradas de 10%, uma taxa de erro de genotipagem de 0,01, níveis de confiança de 80% e 100000 repetições (Marshall et al., 1998). Para determinar o suposto pai das progênie, todas as árvores adultas reprodutivas ($n = 27$) foram testadas como candidatos a pai. Para adicionar um nível extra de rigor para as progênie indicadas como de autofecundação ($n = 31$), as atribuições de paternidade para estas progênie foram verificadas manualmente, de modo que as inadequações em cada loco (a partir de três ou mais erros por loco) entre o pai (árvore materna) e o genótipo da progênie, resultaram no descarte de duas progênie indicadas pelo Cervus. Assim, a taxa de autopolinização (S_p) foi estimada pelo número de progênie identificadas como de autofecundação ($n_{\text{autofecundadas}}$), dividido pelo número total de progênie amostradas (n) em cada família. Para verificar o poder de exclusão para a análise de paternidade, a probabilidade de não-exclusão do segundo parental (P_{exc2}) e a probabilidade de não-exclusão para identidade (P_{ID}) foram estimada pelo software Cervus 3.0.3 (Kalinowski et al., 2007). A P_{exc2} é a probabilidade média de não-exclusão de um candidato a pai erroneamente designado, quando um dos pais é conhecido (neste caso a mãe). Quanto a P_{ID} corresponde à probabilidade média de não exclusão de dois indivíduos aleatórios idênticos geneticamente.

Quanto a depressão por endogamia (δ), esta foi estimada apenas nas famílias maternas com progênie de autofecundação indicado pela a análise de paternidade (Tabela 1S), utilizando a seguinte expressão recomendada por Lande & Schemke (1985):

$$\delta = 1 - W_s / W_o$$

onde, W_s e W_o representam os valores médios fenotípicos das progênie de autofecundação e de fecundação cruzada, respectivamente. O valor da δ foi avaliada para os seguintes caracteres vegetativos: altura de planta (AP), diâmetro de estipe (DE), número de folhas (NF), comprimento de folha (CF), largura do ráquis (LR), e o número de folíolos na folha avaliada (NFA). Os dados morfológicos das progênie de *A. aculeata* foram obtidos a partir do estudo de caracterização fenotípica de Rueda (2014).

Resultados

Diversidade genética

As plantas juvenis exibiram um maior número total de alelos ($k = 55$) do que as árvores adultas ($k = 48$) nos nove locos microssatélites empregados, com média de 6,1 e 5,3 alelos por loco, respectivamente (Tabela 1). O número de alelos por locos variaram

de 2 a 8 entre os dois estágios do ciclo de vida, contudo, a heterozigotidade observada foi significativamente maior em árvores adultas ($H_o = 0,527$) do que nas árvores juvenis ($H_o = 0,441$). Nas árvores adultas, dois locos apresentaram elevados níveis de heterozigotos (mEgCIR3365 e P1T14), enquanto que nas juvenis isso ocorreu somente no loco P1T14. A análise detectou alelos exclusivos ($k = 7$) apenas no conjunto das árvores juvenis. O índice de fixação foi positivo e significativamente diferente de zero para oito dos nove locos (Aacu07, Aacu45, Aacu12, Aacu26, Aacu10, Aacu30, mEgCIR3365 e mEgCIR0840) e para a média geral nas juvenis ($F_{\text{offs}} = 0,433$). Para as arvores adultas o índice de fixação não foi significativamente diferente de zero, com exceção dos locos mEgCIR3365 e P1T14, que foram negativos e significativos. A probabilidade média de não-exclusão de paternidade para um pai candidato dado o genótipo da mãe foi de ($P_{\text{exc2}} = 0,0021$). O poder combinatório dos 9 locos estudados alcançou um valor de probabilidade média de não-exclusão para a identidade de dois indivíduos não relacionados de [$P_{\text{ID}} = 5,6 \times 10^{-9}$] (Tabela 1).

Tabela 1. Estimativas de diversidade genética e do índice de fixação de árvores adultas e juvenis de *Acrocomia aculeata*: k é o número de alelos; H_o e H_e são as heterozigotidades observada e esperada; F e F_{of} são os índices de fixação de adultos e progênes; P_{ID} é a probabilidade de não-exclusão para identidade; P_{exc2} é a probabilidade de não-exclusão do segundo parental.

Locos	Adultos (n = 27)						Juvenis (n = 157)			
	k	H_o	H_e	F	P_{ID}	P_{exc2}	k	H_o	H_e	F_{of}
Aacu07	8	0,577	0,801	0,285	0,074	0,404	8	0,532	0,757	0,327*
Aacu45	4	0,481	0,494	0,026	0,329	0,756	4	0,329	0,557	0,475*
Aacu12	6	0,360	0,725	0,514	0,121	0,510	7	0,278	0,709	0,651*
Aacu26	7	0,407	0,762	0,474	0,098	0,460	8	0,367	0,775	0,558*
Aacu10	7	0,593	0,760	0,224	0,104	0,474	7	0,631	0,808	0,278*
Aacu30	5	0,519	0,744	0,308	0,120	0,512	7	0,306	0,728	0,608*
mEgCIR3365	2	0,667	0,492	-0,362*	0,384	0,817	2	0,399	0,453	0,126*
mEgCIR0840	6	0,211	0,741	0,795	0,121	0,507	7	0,353	0,837	0,751*
P1T14	3	0,926	0,586	-0,590*	0,262	0,708	5	0,771	0,677	-0,049*
Média	5,3	0,527	0,678	0,219	-	-	6,1	0,441	0,700	0,433*
IC _[Class-inf]	4,0	0,389	0,597	0,141	-	-	4,8	0,331	0,619	0,402
IC _[Class-Sup]	6,6	0,664	0,759	0,297	-	-	7,4	0,550	0,781	0,464
Total	48	-	-	-	$5,6 \times 10^{-9}$	0,0021	55	-	-	-

F_{of} : é o índice de fixação individual calculado utilizando como frequências alélicas de referência as frequências estimadas nas árvores adultas.

IC: intervalo de confiança (95%, $\alpha = 0,05$).

* $P < 0,05$ após as correções de Bonferroni.

Taxa de fecundação cruzada e parâmetros do sistema de cruzamento

A taxa de cruzamento multiloco ($\hat{t}_m$) e uniloco ($\hat{t}_s$) a nível populacional para *A. aculeata* é apresentada na Tabela 2. A taxa de cruzamento multiloco foi alta ($\hat{t}_m = 0,986$), contudo a média não diferiu estatisticamente da unidade. Quanto à taxa de cruzamento uniloco, esta foi significativamente $< 1,0$ ($\hat{t}_s = 0,603$, $P < 0,05$). Enquanto que ao nível de famílias foram observados variações das taxas de cruzamentos multilocos de $\hat{t}_m = 0,835$ a $1,000$ (Tabela 3).

Tabela 2. Estimativas de parâmetros do sistema de reprodução, endogamia e parentesco em *Acrocomia aculeata*.

Sistema de reprodução	Estimativa	IC 95%
Taxa de cruzamento multiloco: $\hat{t}_m$	0,986	0,961 – 1,000
Taxa de cruzamento uniloco: $\hat{t}_s$	0,603	0,534 – 0,672
Taxa de autofecundação: $\hat{s} = 1 - \hat{t}_m$	0,014	0,000 – 0,039
Taxa de cruzamento entre parentes: $\hat{t}_m - \hat{t}_s$	0,383	0,312 – 0,454
Correlação de autofecundação: $\hat{r}_s$	0,046	0,011 – 0,081
Correlação multilocos de paternidade: $\hat{r}_{p(m)}$	0,564	0,272 – 0,856
Número efetivo de árvores doadoras de pólen: $\hat{N}_{ep}$	1,773	1,168 – 3,677
Endogamia e estrutura genética		
Coeficiente de endogamia nas árvores maternas: $\hat{F}_m$	0,309	0,240 – 0,378
Coeficiente de endogamia nas progênes: $\hat{F}_o$	0,154	0,000 – 0,298
Endogamia nas progênes por autofecundação: $\hat{F}_s$	0,009	0,000 – 0,024
Endogamia nas progênes por cruzamento entre parentes: $\hat{F}_{t_m - t_s}$	0,145	0,000 – 0,274
Coeficiente de endogamia em equilíbrio de Wright: $\hat{F}_{eq}$	0,007	0,000 – 0,020
Proporção (%) de irmãos de autofecundação: $\hat{P}_{IA}$	0,020	0,000 – 0,156
Proporção (%) de meios-irmãos: $\hat{P}_{MI}$	42,388	14,396 – 67,169
Proporção (%) de irmãos-completos : $\hat{P}_{IC}$	54,832	25,091 – 85,604
Proporção (%) de irmãos de autofecundação e de cruzamento: $\hat{P}_{IAC}$	2,761	0,000 – 7,584
Coancestria média dentro de progênes: $\hat{\theta}_{xy}$	0,256	0,218 – 0,304
Coeficiente de coancestria (estimador de J, Nason): $\hat{f}_{ij}$	0,229	0,136 – 0,324
Tamanho efetivo de variância: $\hat{N}_{e(v)}$	1,937	1,646 – 2,294
Número de árvores matrizes para a coleta de semente: $\hat{m}$	77	65 – 91
Frequência alélica retida: ω_{af}	0,02	0,004 – 0,036
Tamanho amostral		
Número de árvores matrizes	19	
Número médio de progênie por árvore matriz	8,3	

IC: Intervalo de confiança calculado por 1000 bootstraps.

Para 15 das 19 famílias, as estimativas da taxa de cruzamento multiloco foram significativamente $< 1,0$ ($P < 0,05$). A diferença entre a taxa de cruzamento multiloco e

uniloco foi elevada e significativamente diferente de zero ($\hat{t}_m - \hat{t}_s = 0,383$) ao nível de população, como também ao nível de famílias (variando de 0,064 a 0,538), sugerindo uma alta proporção de acasalamento entre aparentados. Apenas a família BGP33, apresentou valor baixo, mas positivo e significativo [$(\hat{t}_m - \hat{t}_s) - (1,96 \times SE) > 0$]. A correlação de autofecundação foi baixa e significativamente diferente de zero ($\hat{r}_s = 0,046$, Tabela 2). Enquanto que, a estimativa da correlação multilocos de paternidade foi significativamente maior que zero ($\hat{r}_{p(m)} = 0,564$ $P < 0,05$) na população, variando de 0,035 a 0,998 nas famílias, indicando que 1,0 – 28,6 doadores de pólen acasalaram com as árvores matrizes (média populacional de $\hat{N}_{ep} = 1,773$).

Tabela 3. Estimativas do sistema de reprodução e parâmetros genéticos de famílias maternas de *Acrocomia aculeata*.

Família / [n]	$\hat{t}_m$	$\hat{t}_m - \hat{t}_s$	$\hat{r}_{p(m)}$	$\hat{N}_{ep}$	$\hat{\Theta}_{xy}$	$\hat{N}_{e(v)}$	$\hat{m}$	$\hat{f}_x$
BGP11 [09]	0,979 (0,009)	0,200 (0,059)	0,086 (0,022)	11,6	0,184	2,71	55	0,067
BGP13 [09]	0,996 (0,000)	0,246 (0,054)	0,690 (0,115)	1,4	0,277	1,80	83	0,095
BGP15 [10]	1,000 (0,000)	0,233 (0,044)	0,880 (0,095)	1,1	0,308	1,63	92	0,068
BGP16 [09]	0,998 (0,002)	0,300 (0,040)	0,227 (0,038)	4,4	0,201	2,48	60	0,271
BGP27 [10]	0,990 (0,004)	0,309 (0,047)	0,071 (0,037)	14,1	0,178	2,80	54	-0,030
BGP29 [08]	0,900 (0,000)	0,162 (0,059)	0,100 (0,001)	10,0	0,213	2,35	64	0,281
BGP31 [09]	1,000 (0,000)	0,538 (0,042)	0,998 (0,012)	1,0	0,327	1,53	98	0,128
BGP33 [08]	0,900 (0,000)	0,064 (0,026)	0,100 (0,001)	10,0	0,213	2,35	64	0,078
BGP36 [10]	0,835 (0,001)	0,198 (0,035)	0,173 (0,095)	5,8	0,243	2,06	73	0,128
BGP37 [10]	1,000 (0,003)	0,331 (0,036)	0,675 (0,158)	1,5	0,274	1,82	82	-0,172
BGP40 [09]	0,999 (0,000)	0,192 (0,038)	0,184 (0,062)	5,4	0,194	2,58	58	-0,157
BGP47 [06]	0,990 (0,004)	0,463 (0,038)	0,705 (0,117)	1,4	0,280	1,78	84	-0,155
BGP48 [04]	0,990 (0,004)	0,246 (0,039)	0,682 (0,121)	1,5	0,277	1,81	83	-0,077
BGP50 [10]	0,990 (0,004)	0,503 (0,041)	0,512 (0,146)	2,0	0,249	2,01	75	0,312
BGP63 [04]	0,990 (0,004)	0,210 (0,050)	0,035 (0,013)	28,6	0,173	2,90	52	0,451
BGP65 [08]	0,953 (0,016)	0,409 (0,059)	0,100 (0,001)	10,0	0,195	2,56	59	0,333
BGP74 [10]	0,953 (0,016)	0,225 (0,038)	0,100 (0,001)	10,0	0,195	2,56	59	0,375
BGP78 [08]	0,979 (0,009)	0,427 (0,032)	0,381 (0,161)	2,6	0,231	2,17	69	0,311
BGP87 [06]	0,979 (0,009)	0,507 (0,031)	0,302 (0,128)	3,3	0,218	2,29	66	0,373

Correlação de autofecundação: $\hat{r}_s$ = variando de 0,029 a 0,117.

Abreviações: n é o número de progênie por família; t_m é a taxa de cruzamento multilocos; $t_m - t_s$ é a taxa de cruzamento entre parentes; $r_{p(m)}$ é a correlação de paternidade multilocos; N_{ep} é o número efetivo de árvores doadores de pólen; Θ_{xy} é o coeficiente de coancestria dentro de família; $N_{e(v)}$ é o tamanho efetivo de variância; m é o número de árvores de sementes necessárias para a coleta de sementes; f_x é o índice de fixação dentro de cada família; () Os números entre parênteses referem-se ao erro padrão.

As estimativas do coeficiente de endogamia para árvores maternas ($\hat{F}_m = 0,309$) e nas progênies ($\hat{F}_o = 0,154$) foram altas e positivas, indicando a presença de endogamia, embora apenas o índice estimado para árvores maternas foi estatisticamente diferente de

zero ($P < 0,05$). A maior fração da endogamia observada nas plantas progênes se deve a cruzamento entre árvores aparentadas ($\hat{F}_{t_m - t_s} = 0,145$, Tabela 2), como indicado também pelo menor valor do coeficiente de endogamia em equilíbrio de Wright ($\hat{F}_{eq}$) em relação ao coeficiente de endogamia das progênes ($\hat{F}_o$). O coeficiente de parentesco pelo estimador de J. Nason para árvores maternas foi de $f_{ij} = 0,229$ e o intervalo de confiança de 95% após 10000 permutações variou 0,136 a 0,324. O coeficiente médio de coancestria dentro de progênie foi $\hat{\Theta}_{xy} = 0,256$, maior do que o esperado para famílias de meios-irmãos (0,125), sugerindo que as progênes são compostas por misturas de diferentes tipos de parentesco. As progênes de *A. aculeata* foram compostas predominantemente de irmãos completos (54,83%) e meios irmãos (42,39%). O tamanho populacional efetivo da variância ($N_{e(v)} = 1,937$, Tabela 2) foi menor do que o esperado para populações panmíticas ($N_{e(v)} = 4$). O número de árvores matrizes (m) para a coleta de sementes, com o objetivo de manter um tamanho efetivo populacional de 150, foi de 77 indivíduos de macaúba (variando de 65 a 91, Tabela 2).

Análise de paternidade e depressão por endogamia

A análise de paternidade indicou a presença de progênes provindas de autofecundação em 13 das 19 famílias estudadas (Tabela 4). A taxa de autofecundação entre estas famílias variou de $S_p = 10$ a 50%, com uma taxa média de $S_p = 18\%$. A depressão por endogamia (δ) foi estimada nas 13 famílias identificadas como progênes de autofecundação (Tabela 4). Em geral, a magnitude do efeito da endogamia variou muito entre as famílias de *A. aculeata*. A ação depressiva da endogamia foi observada em todas as características morfológicas analisadas, onde os percentuais de depressão de cada família ($\delta \%$) variaram de 0,2 a 82,2%. A família BGP27 apresentou os valores mais elevados ($\delta > 49\%$), afetando de forma significativa o desempenho de todos os caracteres vegetativos analisados. Além desta, as famílias BGP37 e BGP40 também tiveram seu desempenho fortemente afetado pela endogamia, com $10 > \delta < 81\%$. Entre as variáveis fenotípicas, a altura de planta foi a mais significativamente afetada ($\delta > 19\%$). Para esta característica foram registrados índices percentuais de depressão extremamente elevados ($\delta > 80\%$) nas famílias BGP27 e BGP37. A depressão por endogamia a nível populacional revelou valores moderadamente elevados, sendo os maiores valores depressivos observados para os caracteres altura de planta ($\delta = 12,8\%$), comprimento de folha ($\delta = 9,7\%$) e diâmetro de caule ($\delta = 7,1\%$).

Tabela 4. Estimativas de depressão por endogamia (δ) para os caracteres vegetativos de famílias maternas de *Acrocomia aculeata* com progênie de autofecundação identificadas pelo teste de análise de paternidade: altura da planta (AP), diâmetro do estipe (DE), número de folhas (NF), comprimento da folha (CF), largura do ráquis (LR) e número de folíolos (NFA).

Famílias	S_p	AP cm		DE cm		NF		CF cm		LR cm		NFA	
		W_o	δ (%)	W_o	δ (%)	W_o	δ (%)	W_o	δ (%)	W_o	δ (%)	W_o	δ (%)
BGP11	0,22	95,6	31,5	42,9	9,4	17,9	-3,6	265,7	4,0	1,2	0,2	333,0	-0,9
BGP27	0,10	140,6	82,2	38,4	49,2	18,2	72,5	274,4	75,2	1,4	73,7	306,8	57,3
BGP29	0,13	65,0	-41,5	42,9	-18,9	14,6	-9,8	245,7	2,3	1,0	-13,3	329,7	10,2
BGP31	0,44	83,4	-7,9	47,6	9,6	16,8	16,7	272,0	-1,1	1,3	-6,2	320,4	3,5
BGP36	0,30	95,1	-32,4	47,7	-1,6	20,6	4,4	292,9	4,4	1,4	-0,7	328,0	-10,0
BGP37	0,10	112,2	80,4	44,7	19,5	15,8	23,9	290,0	13,8	1,5	14,8	285,8	-7,1
BGP40	0,11	108,9	34,8	52,5	34,8	19,0	36,8	268,8	10,7	1,3	14,1	336,4	13,8
BGP47	0,50	126,0	-0,3	29,4	-1,7	21,7	-9,2	186,7	0,0	1,0	-0,3	187,7	2,1
BGP50	0,30	89,3	20,4	40,5	4,2	18,2	-6,1	286,4	9,2	1,3	11,5	318,6	0,6
BGP65	0,38	224,0	-8,6	47,2	-5,9	22,2	-11,1	218,8	-8,0	2,3	-11,7	309,2	-6,1
BGP74	0,10	221,1	-4,0	35,2	-36,3	21,2	-55,5	263,6	-2,4	2,8	-6,3	340,2	-1,1
BGP78	0,38	216,0	19,8	36,0	7,4	21,2	18,2	217,2	3,2	2,7	-7,0	306,4	4,3
BGP87	0,50	156,7	-12,8	36,3	12,8	16,3	-26,5	253,0	4,9	2,7	13,1	325,3	11,5
Population ^a	0,18	133,4	12,8	41,6	7,1	18,7	3,2	256,5	9,7	1,7	4,9	309,8	6,1

S_p : Taxa de autofecundação calculada com base no número de progênie idenficadas como de autopolinização pelo teste de paternidade. W_o = Valores médios de caracteres vegetativos de progênies híbridas.

^aA taxa média de autofecundação a nível populacional é com base na média das 19 famílias estudadas.

Discussão

Este é o mais amplo estudo realizado sobre o sistema de cruzamentos da *A. aculeata*, e o primeiro a apresentar os efeitos depressivos sobre o desempenho de caracteres vegetativos no estágio juvenil desta espécie.

Diversidade genética

Observamos uma ampla variabilidade alélica para os nove locos analisados na espécie *A. aculeata* (média de 5,3 e 6,1 alelos por locos, respectivamente, em plantas adultas e juvenis). A maior parte dos alelos exclusivos (85%) detectado nas plantas juvenis foram de baixa frequências dentro de seus locos ($\leq 0,052$), indicando que novas formas alélicas foram introduzidas por uma vizinhança não aparentada (Lanes et al., 2014).

Os baixos valores de heterozigosidade observada (H_o) e altas estimativas do índice de fixação (F) nas plantas juvenis em relação às árvores adultas sugerem uma maior endogamia nas juvenis, e conseqüentemente, a ocorrência de seleção contra

homozigotos. Este resultado sugere uma alteração nas frequências alélicas e genotípicas entre as fases ontogenéticas da *A. aculeata*. A seleção contra homozigotos também tem sido observado em outras espécies, ao comparar-se a fase inicial do ciclo de vida das plantas (semente, plântula e juvenil) com o estágio da planta adulta (Eguiarte et al., 1992; Gaiotto et al., 2003; Carneiro et al., 2011).

A ampla variabilidade genética esta presente em ambas as fases do ciclo de vida da *A. aculeata*. Contudo, o excesso de homozigose na fase juvenil sugere que a endogamia é naturalmente recorrente; todavia, o fenômeno tem forte efeito negativo sobre a capacidade dos indivíduos em tolerar as pressões ambientais, contribuindo para a eliminação gradual destes ao longo do ciclo de vida.

Sistema de cruzamento da *A. aculeata*

Os parâmetros de sistema de cruzamento determinados com base no modelo de cruzamento misto e correlacionado (Ritland & Jain, 1981; Ritland, 1989) revelaram elevados índices de polinização cruzada ao nível populacional, com taxa de autofecundação de $\hat{s} < 5\%$ considerando 95% de IC. Além disso, estimativas elevadas de polinização cruzada ($\hat{t}_m > 95\%$), também foram observadas em um total de 16 famílias (84%). Estes resultados permitem inferir que a *A. aculeata* é uma espécie predominantemente de fecundação cruzada (alógama), conforme a taxa de autofecundação ($\hat{s} < 20\%$) aceita por inúmeros autores (Schemske & Lande, 1985; Goodwillie et al., 2005; Escobar et al., 2011; Winn et al., 2011). O que contraria vários estudos que afirmavam se tratar de uma espécie com níveis intermediários de autofecundação ($20 < \hat{s} < 80\%$), ou seja, sistema misto de reprodução (Scariot et al. 1991; Nucci et al., 2008; Abreu et al., 2012).

O reduzido número de descendentes por autopolinização ($\hat{s} < 5\%$) observado na *A. aculeata* (Tabela 1), pode ser atribuída à baixa disponibilidade ou deficiência dos polinizadores, como também pela elevada taxa de abortamento precoce de frutos, que tem como uma das causas uma possível carga mutacional excessiva (Husband & Schemske, 1996; Wilcock & Neiland, 2002). A depressão decorrente da endogamia no estágio inicial do ciclo de vida é característico de espécies alógamas, e sua causa está associada a alelos recessivos altamente deletérios (Husband & Schemske, 1996; Mustajärvi et al., 2005; Thiele et al., 2010). Diversas pesquisas oferecem indícios de que os efeitos endogâmicos possam ocorrer prematuramente, ainda na fase gametofítica; afetando a taxa de crescimento de tubo polínico, viabilidade e quantidade de pólen e óvulo produzidos (Carr & Dudash, 1995; Jóhannsson et al., 1998; Melser et al., 1999).

Contudo, considerando a base genética da depressão endogâmica, é pouco plausível que esta se estenda para a geração haplóide (Hayes et al., 2005).

Além dos fatores depressivos pós-zigóticos comuns (aborto de frutos, viabilidade das sementes e mortalidade de plântulas), é provável que *A. aculeata* tenha desenvolvido algum outro tipo de sistema de autoincompatibilidade (Abreu et al., 2012). Esta hipótese é ainda mais fortalecida se aplicarmos o “índice de auto-incompatibilidade” (ISI) de Zapata & Arroyo (1978) e Bullock (1985) sobre os percentuais de frutificação da *A. aculeata*, obtidos de polinizações controladas por Scariot et al. (1991). Neste contexto o valor obtido ($ISI = 0,26$) sugere que a espécie é parcialmente auto-incompatível (Zapata & Arroyo, 1978). Todavia, existe a possibilidade de algumas famílias deixarem mais descendentes por autofecundação do que outras, conforme indicado pela correlação de autofecundação entre famílias ($0,029 \leq \hat{f}_s \leq 0,117$), possivelmente devido à variações no sistema de auto-incompatibilidade entre árvores dentro de populações ou certos agrupamentos. Segundo del Castillo & Trujillo (2008), a verdadeira amplitude de variação do sistema de cruzamento dentro de uma mesma espécie é desconhecida.

Cruzamento correlacionados e entre indivíduos aparentados

Detectou-se níveis elevados de cruzamento entre indivíduos aparentados de *A. aculeata* ($\hat{f}_m - \hat{f}_s = 0,383$). Além deste, a correlação de paternidade indica que uma grande proporção das progênes foram geradas pelo mesmo parental materno e paterno, sugerindo que a polinização cruzada é não aleatória. Cruzamentos correlacionados implicam em um número reduzido de árvores contribuindo para a efetiva polinização. Esta forte restrição na contribuição paterna foi confirmada pelo número extremamente baixo de árvores efetivamente polinizadoras ($\hat{N}_{ep}$), com média de dois indivíduos por árvore (Tabela 2). Tal fato também foi observado na palmeira *Butia eriospatha* por Nazareno & Reis (2012). Estes resultados estão, possivelmente, associados à existência de estrutura espacial intrapopulacional, de forma que os indivíduos que compõem uma população natural ou “maciço” possuem um alto grau de parentesco (Smouse & Sork, 2004). Por sua vez, a estrutura genética espacial favorece a endogamia biparental na descendência, o qual foi confirmado, no presente estudo, pelas estimativas de endogamia na descendência de acasalamento entre parentes ($\hat{F}_{t_m - t_s}$).

A elevada endogamia nas progênes de *A. aculeata* pode ser atribuída à restrição no modo de dispersão dos frutos da espécie, primordialmente barocórica, de forma que, estes são dispersos próximos à árvore mãe (Abreu et al., 2012; Lanes et al., 2015),

favorecendo a formação de estrutura espacial intrapopulacional. Este tipo de limitação também é observado em outras palmáceas, como *Attalea phalerata* e *Astrocaryum mexicanum* a maioria das sementes se encontram dispersas a uma distância menor que 30m da árvore materna (Eguiarte et al., 1993; Choo et al., 2012). A dispersão zoocórica também está presente, entretanto, o processo depende de um sistema sincronizado entre a ocorrência de vertebrados com a disponibilidade dos frutos (Eiserhardt et al, 2011).

Outra causa da endogamia observada seria a associação entre a inflorescência andrógina da palmeira e o comportamento dos polinizadores, visitando de forma sistemática uma mesma inflorescência ou inflorescência vizinhas. Os coleópteros são os principais polinizadores da *A. aculeata* (Scariot et al., 1991), e sua limitada autonomia de vôo restringiria o fluxo gênico (Lanes et al., 2015). Porém, pouco se sabe sobre o comportamento destes polinizadores na espécie. A distância média de dispersão do pólen (carreado pelo vento e/ou inseto) em algumas espécies de palmáceas tem sido relativamente baixa; como cerca de 303m para *Oenocarpus bataua* (Ottewell et al., 2012), 71m para *Phoenix canariensis* (Saro et al., 2014), 200m para *Iriartea deltoidea* (Sezen et al., 2007) e 20m para a *Astrocaryum mexicanum* (Eguiarte et al., 1993). Contudo, Barfod et al. (2011) alertaram para o fato de que apenas cerca de ~3% de todas as espécies de palmáceas tiveram o seu mecanismo de polinização estudado em detalhes, e portanto, cuidados devem ser tomados ao se fazer generalizações.

Neste estudo, três famílias de *A. aculeata* (BGP29, BGP33 e BGP36), em que as árvores matrizes se encontram em locais restritos ou fragmentos isolados, apresentaram elevadas taxas de autofecundação ($\hat{s} = 10 - 16,5 \%$). Provavelmente este resultado seja consequência da pressão de antropização nos habitats originais, uma vez que estes localizam-se no entorno de regiões urbanas, densamente povoadas, com cerca de 6,2 milhões de habitantes (IBGE, 2010). Um ambiente sobre forte fragmentação ou mesmo o isolamento de populações propiciam a endogamia (Dudash & Fenster, 2000). Ademais, os dispersores de sementes, como primatas, pequenos marsupiais e aves, são sensíveis à fragmentação, e como consequência suas respectivas populações descresem sensivelmente (Laurance et al., 2002). Segundo Portela & Santos (2014), a fragmentação tem levado a uma ausência de dispersores de sementes da palma *Astrocaryum aculeatissimum*, o que tem contribuído para a redução das populações desta espécie no sudeste do Brasil. A fragmentação de florestas tropicais tem conduzido uma diminuição no número de polinizadores e de visitantes florais (Nayak & Davidar, 2010; Bergsdorf, 2006), ocasionado uma redução do fluxo gênico e perda de diversidade genética, enquanto favorece cruzamentos endogâmicos, aumento da taxa de

inflorescências abortivas e de sementes inviáveis em várias espécies de palmáceas, incluindo *Astrocaryum mexicanum*, *Oenocarpus bacaba*, *Mauritia flexuosa*, e *Butia eriospatha* (Aguirre & Dirzo, 2008; Lepsch-Cunha et al., 2003; Federman et al., 2014; Nazareno & Reis, 2014).

A estimativa da coancestria dentro das progênes de *A. aculeata* (Tabela 2) indica que existe uma probabilidade de 25,6% de que 2 alelos amostrados em 2 plantas da mesma progênie sejam idênticos por descendência. Esse resultado sugere que as progênes analisadas são compostas por uma mistura de parentescos. De fato, essa mistura de parentesco entre as progênes foi reafirmada pelo tamanho efetivo de variância ($\hat{N}_{e(v)}=1,937$), indicando que a maior fração das progênes geradas são aparentadas no grau de irmãos-completos. A representatividade genética na descendência de *A. aculeata* está 52% abaixo da esperada em progênes de cruzamentos aleatórios (população idealizada). Em uma população idealizada, ou seja, perfeitamente panmíticas (tamanho infinito, praticando cruzamentos aleatórios, ausência de seleção, mutação, migração e sobreposição de gerações) o tamanho efetivo de variância equivale a 4,0 ($\hat{N}_{e(v)} = 0,5 / \hat{\theta}_{xy} = 0,5 / 0,125$), o qual corresponde a progênes de meios-irmãos. Sendo assim, em *A. aculeata* os cruzamentos biparentais e endogâmicos foram responsáveis pelos desvios das pressuposições que caracterizam uma população idealizada, promovendo uma redução no tamanho efetivo de variância. Isso implica na necessidade de coleta de maiores tamanhos amostrais para que se garanta a manutenção da variabilidade genética em programas conservação genética ex situ e in situ. Segundo Namoff et al. (2010), a conservação genética ex situ de um total de 15 indivíduos coletados de três árvores matrizes (famílias de meios-irmãos) da palmeira *Leucothrinax morrisii* assegura a captura de mais de 80% da diversidade alélica da população original. Contudo, em famílias com muitos irmãos-completos deve-se priorizar o número de árvores maternas.

Depressão endogâmica entre e dentro de famílias

A magnitude da depressão por endogamia variou entre as famílias maternas da *A. aculeata*, dentre as quais, a família BGP27 foi a mais significativamente afetada (Tabela 4). Segundo Fox (2005), a depressão por endogamia é consequência da redução do número de locos heterozigotos e da restrição de novas formas alélicas presente na progênie endogâmica. Assumindo que as características avaliadas são controladas em sua maior fração por fatores genéticos, e que os locos destes caracteres estão sob uma alta concentração de heterozigose na família BGP27, como demonstrado pelo seu

respectivo índice de fixação ($\hat{f}_x = -0,030$). É provável que a autofecundação imposta a esta família (Tabela 3), conduziu a um acúmulo de locos com mutações deletérias recessivas de pequenos efeitos no estado homozigoto. Isto, somado ao custo da perda da ação da dominância ou sobredominância de alguns locos heterozigotos (Charlesworth & Willis, 2009), tiveram grande impacto sobre o desempenho dos caracteres vegetativos para a família BGP27. Um caráter controlado sob qualquer grau de dominância ou sobredominância, terá sua aptidão reduzida na proporção do aumento do coeficiente de endogamia (Keller & Win, 2002; Charlesworth & Willis, 2009). O que de fato foi confirmado pelo excesso de heterozigotos detectado pelo índice de fixação e a elevada magnitude depressiva das famílias BGP27, BGP37 e BGP40. Entretanto, somado aos fatores genéticos, os níveis de expressão da endogamia entre famílias podem variar sob condições de estresse ambiental (Hayes et al., 2005; Paschke et al., 2005; Waller et al., 2008; Fox & Reed, 2010).

Os níveis de depressão por endogamia foram elevados e afetaram de forma significativa o desempenho de todos os caracteres vegetativos da *A. aculeata*, em especial a altura de planta. Este resultado era esperado, por se tratar de um caráter vegetativo de natureza genética (herdabilidade = 87%, Manfio et al. 2012), provavelmente sob controle genético da dominância e/ou epistasia. Níveis expressivos de depressão para a altura de planta também foram descrito em *Elaeis guineense*, o qual significativamente correlacionou-se ($r = -0,710$) com o coeficiente de endogamia (Corley & Tinker, 2003; Luyindula et al., 2005), e em *Cocos nucifera* L., onde os níveis depressivos para altura de planta foram maiores a cada avanço de geração de autofecundação (Pandi, 2009).

Em contrapartida, valores negativos de depressão para todas as variáveis morfológicas foram detectados nas famílias maternas BGP65 e BGP74, indicando superioridade dos caracteres vegetativos das progênies autofecundadas em relação às progênies de cruzamento (Tabela 4). Há uma série de fatores que podem contribuir para estes resultados: (i) Ocorrência de um possível viés de plantas autofecundadas identificadas pelo teste de paternidade, uma vez que a aplicação de um pequeno número de marcadores polimórfico tem pouco poder em distinguir plantas autofecundadas de cruzamentos endogâmicos (Weller et al., 2005; Winn et al., 2011). Ainda que a probabilidade de exclusão de paternidade (P_{exc2}), para os 9 locos combinados, indique que mais de 99,8% dos pais candidatos poderiam ser excluídos da paternidade, assumido como conhecido o genótipo do genitor materno (Tabela 1); (ii) A

superioridade das progênes endogâmicas é devido ao alto mérito genético das famílias (BGP65 e BGP74). As progênes híbridas da *A. aculeata* destas famílias maternas, apresentaram valores médios fenotípicos (W_0) superiores as médias populacionais em quase todos os caracteres morfológicos analisados (Tabela 4). Na palma *Elaeis guineense*, a descendência dos cruzamentos envolvendo genótipos superiores com genótipos de baixo mérito genético, tiveram uma maior redução do desempenho de parâmetro de crescimento e produtividade do que os cruzamentos entre irmãos (Luyindula et al., 2005); (iii) O declínio da performance dos caracteres vegetativos das progênes híbridas das famílias BGP65 e BGP74 é resultado da depressão exogâmica, ou seja, cruzamento entre famílias excessivamente divergentes em seus genomas (Waser et al., 2000; Escobar et al., 2008; Rousselle et al., 2011).

Os mecanismos de depressão exogâmica que poderiam explicar a perda de desempenho das progênes híbridas da *A. aculeata* podem ser atribuídas ao “mecanismo ecológico” ou o “efeito de diluição” dos genes associados à adaptação das matrizes aos seus respectivos ambientes de origem. Dessa forma, a descendência híbrida introduzida a um novo ambiente demonstraria menor desempenho, por transportar genes mal adaptados a esta nova condição ambiental (Rogers & Montalvo, 2004; Rousselle et al., 2011). Para montagem do BAG-Macaúba, sementes de matrizes nativas foram coletadas de diferentes regiões microclimáticas e utilizadas para implantação da coleção viva de germoplasma ex situ, ou seja, a descendência foi inserida em uma nova condição biótica e abiótica. Outra explicação seria a divergência a nível cromossômico entre populações, conduzindo a um pareamento anormal dos cromossomos, e como resultado a má formação de indivíduos ou até a inviabilidade destes (Rogers & Montalvo, 2004). Abreu et al. (2011) e Röser et al. (1997) encontraram diferentes conteúdo do DNA (2C-value) na *A. aculeata*, possivelmente em decorrência da grande variabilidade genética presente nas diversas regiões eco-geográficas.

Implicações para o melhoramento e a conservação genética

Muitos programas de melhoramento assumem que espécies alógamas como a *A. aculeata* são originadas de cruzamentos aleatórios. No entanto, a análise de sistema de cruzamento revelou desvios da pressuposição de panmixia, indicando que a maior parte das progênes são irmãos-completos. Esta informação tem implicações importantes no melhoramento genético, uma vez que, assumir que progênes de polinização aberta de *A. aculeata* são meios-irmãos gera superestimativas na variância genética aditiva, herdabilidade e ganhos na seleção. Pois, se as progênes de polinização aberta são

coletadas de populações sem endogamia e parentesco, a correlação genética entre plantas dentro de progênies, ou coeficiente de parentesco ($\hat{r}_{xy}$) estima 1/4 (0,25) de variância genética aditiva. Sendo o coeficiente de parentesco entre plantas dentro de progênies correspondente ao dobro do coeficiente de coancestria ($\hat{r}_{xy} = 2\hat{\theta}_{xy}$) em espécies diploides (Lynch & Walsh, 1998), a verdadeira estimativa do coeficiente de parentesco da *A. aculeata* equivale a 0,512 de variância genética aditiva ($0,512 \hat{\sigma}_A^2$). Os desvios das pressuposições que caracterizam uma população idealizada também têm implicação na conservação genética ex situ e coleta de sementes de *A. aculeata*. Para reter o tamanho efetivo populacional de 150, será necessária a coleta de sementes de 77 árvores matrizes. Esse tamanho efetivo populacional garante a retenção de alelos de frequência $\omega_{af} \geq 2\%$ na população original, valor abaixo do mínimo exigido ($\omega_{af} \geq 4\%$) para programas de melhoramento, de acordo com Resende (2002). Mas, segundo este mesmo autor, caso o objetivo seja o uso da espécie para conservação in situ e recuperação de áreas degradadas, será necessária a coleta de sementes de 258 árvores matrizes para manter o tamanho efetivo populacional de 500 e retenção de alelos ($\omega_{af} \geq 0,4\%$) na população original.

Dentre as hipóteses levantadas sobre a superioridade das progênies autofecundadas, se confirmada à atribuição do alto desempenho (alto valor per se) ao mérito genético destas famílias, esta seria de grande relevância para o programa de melhoramento genético da *A. aculeata*. A identificação de genitores superiores e com alta capacidade específica de combinação permitiria a adoção de métodos de melhoramento refinados como a Seleção Recorrente Recíproca Individual (SRRI). Nesta estratégia de seleção, os genitores são autofecundados para eliminar a carga genética (alelos deletérios), gerando duas populações S_1 que se inter cruzam para produzir híbridos superiores (Resende & Barbosa, 2005). Estratégia semelhante de seleção recorrente é empregada em programas de melhoramento de dendê em países como o Congo, Costa do Marfim e Nigéria (Baudouin et al., 1997; Corley & Tinker, 2003). Além deste, os ganhos em componentes de produtividade podem ser maximizados, caso a seleção dentro das progênies endogâmicas, seja praticada em caracteres vegetativos correlacionados e moderadamente ou não afetados pela depressão (caráter sob controle de genes cuja interação alélica predominante seja aditiva ou com dominância parcial).

As altas taxas de cruzamentos endogâmicos e biparentais são consequências da estrutura espacial intrapopulacional e do isolamento de muitas populações devido a

fragmentação do habitat natural (Dudash & Fenster, 2000). Além deste, alterações a nível genômico poderiam ser promovidas em pequenas populações isoladas devido à deriva genética, bem como, em decorrência da adaptação local dos genótipos às condições ambientais. Portanto, sobre a ótica dos efeitos da ação depressiva por endogamia e exogamia (Shapcott et al., 2009; Rogers & Montalvo, 2004), a introdução de um novo material, seja para o melhoramento da população ou para assegurar a variabilidade genética de espécies ameaçadas, deve ser praticada de forma consciente, levando em consideração os impactos genéticos.

Agradecimentos

Este estudo é parte da Tese de Doutorado do primeiro autor (Programa de Genética e Melhoramento — Universidade Federal de Viçosa). O autor agradece a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES-Proex: nº 32002017004P2), e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq: nº 143353/2011-0) pela concessão da bolsa de doutorado; e a Petróleo Brasileiro S.A. (PETROBRAS: nº 0050.0061.571.109) pelo financiamento do projeto.

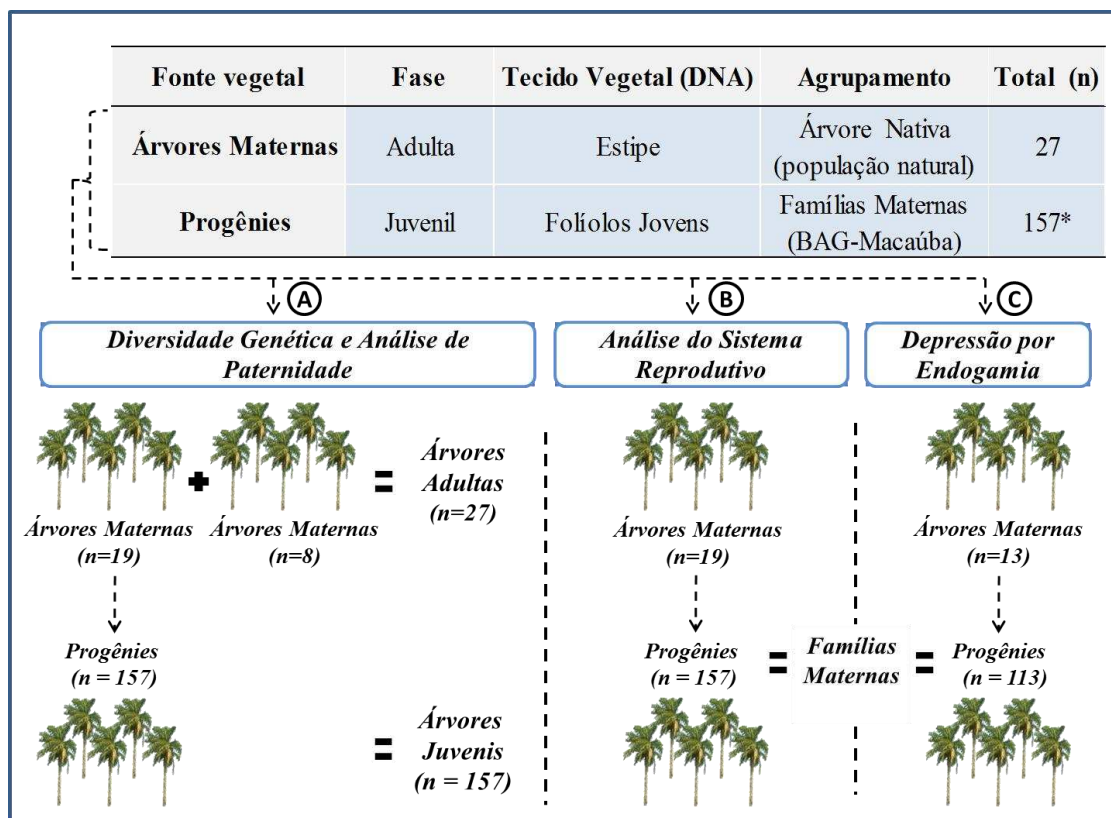


Figura 1S. Esquema do estudo da *Acrocomia aculeata*. O diagrama mostra as fontes de plantas e o tipo de tecido vegetal usado para a extração de DNA. Árvores maternas ou Matrizes: São as árvores adultas nativas, doadoras de semente que deram origem as coleção viva da BAG-Macaúba; Progênie: Plantas jovens que compõem o Banco Ativo de Germoplasma da Macaúba. O termo "família materna" refere-se aos indivíduos descendentes de uma mesma árvore materna. **(A)** Para a pesquisa de diversidade genética e teste de paternidade foram utilizadas 27 árvores adultas nativas e 157 progênie (obtidas de 19 árvores matrizes); **(B)** O sistema de cruzamento foi analisado utilizando 157 progênie juntamente com as respectivas informações genotípicas de suas árvores matrizes (n=19); **(C)** A depressão por endogamia foi estimada nas progênie de 13 famílias maternas, uma vez que, nestas famílias foram identificadas progênie de autofecundação pelo teste de paternidade.

Tabela 1S. Valores fenotípicos dos caracteres vegetativos de cada progênie híbrida e de autofecundação de *Acrocomia aculeata*: altura de planta (AP), diâmetro de estipe (DE), número de folhas (NF) comprimento de folha (CF), largura do ráquis (LR), e número de folíolos na folha avaliada (NFA).

Famílias	AF ¹	Valores Fenotípicos						Famílias	AF	Valores Fenotípicos					
		AP cm	DE cm	NF	CF cm	LR cm	NFA			AP cm	DE cm	NF	CF cm	LR cm	NFA
BGP11-1	*	33	24,7	18	250	1,22	336	BGP40-3		125	51,2	21	300	1,59	376
BGP11-2		113	50,0	12	220	0,85	294	BGP40-4		149	55,0	16	260	1,43	380
BGP11-3		78	32,0	14	250	1,47	352	BGP40-5		140	61,0	22	280	1,20	288
BGP11-4		64	46,0	20	290	1,35	362	BGP40-6		68	36,5	12	250	1,14	329
BGP11-5	*	98	53,0	19	260	1,13	336	BGP40-7		168	67,0	21	350	1,70	305
BGP11-6		94	38,7	18	270	0,91	338	BGP40-8		89	65,0	21	260	1,12	382
BGP11-7		93	49,2	23	290	1,30	324	BGP40-9		97	50,0	23	250	1,17	356
BGP11-8		144	44,0	22	260	1,28	321	BGP47-1	*	171	30,0	29	230	1,04	171
BGP11-9		83	40,2	16	280	1,08	340	BGP47-2		66	28,6	13	180	1,05	227
BGP27-1		261	20,7	26	458	2,39	323	BGP47-3	*	82	23,8	14	160	1,15	224
BGP27-2		172	19,7	25	300	1,32	334	BGP47-4		129	26,5	24	190	1,07	177
BGP27-3		166	20,5	18	300	1,37	318	BGP47-5	*	126	35,8	28	170	0,86	156
BGP27-4		100	40,3	17	280	1,43	332	BGP47-6		183	33,0	28	190	0,92	159
BGP27-5	*	25	19,5	5	68	0,36	131	BGP50-1	*	60	47,3	19	290	1,13	312
BGP27-6		144	38,0	17	270	1,11	334	BGP50-2		76	38,2	18	320	1,13	378
BGP27-7		106	40,5	15	300	1,20	311	BGP50-3		170	36,8	29	490	2,37	306
BGP27-8		121	96,2	13	310	1,47	301	BGP50-4		54	45,5	18	250	1,09	345
BGP27-9		215	47,3	25	290	1,35	292	BGP50-5	*	83	33,1	22	230	1,07	312
BGP27-10		25	18,0	8	130	0,86	190	BGP50-6		59	38,0	15	360	1,62	310
BGP29-1		46	43,0	17	240	0,84	322	BGP50-7		89	48,0	23	280	1,22	440
BGP29-2		74	54,0	18	240	0,93	314	BGP50-8	*	70	36,0	17	260	1,24	326
BGP29-3	*	92	51,0	16	240	1,17	296	BGP50-9		39	23,3	12	220	1,05	235
BGP29-4		78	39,2	13	260	1,11	314	BGP50-10		25	31,5	14	280	0,96	338
BGP29-5		86	33,0	20	230	0,88	385	BGP65-1		230	47,0	24	244	2,18	368
BGP29-6		41	37,5	11	230	1,06	305	BGP65-2	*	220	46,0	19	234	2,72	372
BGP29-7		66	63,0	10	240	1,12	305	BGP65-3	*	270	52,0	32	239	2,72	288
BGP29-8		64	30,5	13	280	1,29	363	BGP65-4		220	48,0	22	205	2,18	264
BGP31-1	*	73	30,5	10	280	1,30	279	BGP65-5		280	57,0	29	231	2,59	322
BGP31-2	*	78	48,4	9	280	1,40	280	BGP65-6	*	240	52,0	23	236	2,28	324

Tabela 1S. Continuação

Famílias	AF	Valores Fenotípicos						Famílias	AF	Valores Fenotípicos					
		AP cm	DE cm	NF	CF cm	LR cm	NFA			AP cm	DE cm	NF	CF cm	LR cm	NFA
BGP31-3		71	54,3	12	280	1,30	315	BGP65-7		180	42,0	15	194	2,08	256
BGP31-4		83	46,2	16	280	1,33	387	BGP65-8		210	42,0	21	220	2,49	336
BGP31-5	*	111	44,2	18	280	1,36	393	BGP74-1		390	36,0	36	254	2,51	356
BGP31-6		103	49,2	22	280	1,30	334	BGP74-2		120	18,0	13	232	2,79	288
BGP31-7		110	47,0	20	260	1,28	290	BGP74-3		200	31,0	17	284	2,92	332
BGP31-8	*	98	49,2	19	260	1,30	285	BGP74-4		130	17,0	12	260	3,12	344
BGP31-9		50	41,5	14	260	1,10	276	BGP74-5	*	230	48,0	33	270	2,93	344
BGP36-1		120	52,7	23	330	1,63	414	BGP74-6		270	41,0	23	328	2,63	364
BGP36-2		94	51,6	24	260	1,16	308	BGP74-7		200	39,0	18	248	2,49	364
BGP36-3	*	141	46,0	22	290	1,42	412	BGP74-8		260	47,0	29	237	2,71	316
BGP36-4	*	100	53,4	22	280	1,34	338	BGP74-9		260	49,0	23	267	2,82	382
BGP36-5	*	137	46,0	15	270	1,32	332	BGP74-10		160	39,0	20	262	2,81	316
BGP36-6		100	45,3	21	290	1,17	318	BGP78-1		310	41,0	28	210	2,52	344
BGP36-7		62	48,1	15	300	1,53	332	BGP78-2		260	41,0	27	255	2,79	288
BGP36-8		100	44,7	23	300	1,43	280	BGP78-3		130	29,0	17	212	2,92	312
BGP36-9		106	46,6	23	280	1,32	280	BGP78-4	*	120	24,0	10	166	2,83	232
BGP36-10		84	45,0	15	290	1,21	364	BGP78-5		190	36,0	19	208	2,59	292
BGP37-1		106	34,0	12	220	1,59	232	BGP78-6	*	170	37,0	23	212	2,93	324
BGP37-2		163	44,0	17	290	1,34	293	BGP78-7	*	230	39,0	19	253	2,91	324
BGP37-3		111	57,0	22	270	1,27	313	BGP78-8		190	33,0	15	201	2,69	296
BGP37-4	*	22	36,0	12	250	1,27	306	BGP87-1	*	230	42,0	25	246	2,73	316
BGP37-5		96	36,0	16	290	1,43	296	BGP87-2	*	190	40,0	23	272	2,43	276
BGP37-6		70	35,5	14	290	1,34	303	BGP87-3	*	110	13,0	14	204	1,93	272
BGP37-7		147	40,0	15	270	1,48	313	BGP87-4		160	33,0	17	248	2,72	324
BGP37-8		134	46,2	19	280	1,56	261	BGP87-5		160	35,0	16	263	2,72	352
BGP37-9		119	54,0	15	430	2,11	283	BGP87-6		150	41,0	16	248	2,72	300
BGP37-10		64	56,0	12	270	1,30	278								
BGP40-1	*	71	34,2	12	240	1,12	290	Média		129	41,3	19	260	1,65	311
BGP40-2		35	34,0	16	200	1,08	275								

¹AF = Progenies de autofecundação: O (*) indica as progenies de autofecundação detectadas pelo teste de paternidade realizado no software Cervus.

Referências

- Abreu IS, Carvalho CR, Carvalho GMA, Motoike SY. 2011. First karyotype, DNA C-value and AT/GC base composition of macaw palm (*Acrocomia aculeata*, Arecaceae) – a promising plant for biodiesel production. *Australian Journal of Botany* 59:149–155.
- Abreu AG, Priolli RHG, Azevedo-Filho JA, et al. 2012. The genetic structure and mating system of *Acrocomia aculeata* (Arecaceae). *Genetics and Molecular Biology* 35:119–121.
- Agwei-Dwarko D, Ofori K, Kaledzi PD. 2012. Variation and correlation analysis of growth parameters in D x P oil palm (*Elaeis guineensis* J.) seedlings. *Elixir Agriculture* 47:8946–8949.
- Ågren J, and Schemske DW. 1993. Outcrossing rate and inbreeding depression in two annual monoecious herbs, *Begonia hirsuta* and *B. semiovata*. *Evolution* 47:125–135.
- Aguirre A, Dirzo R. 2008. Effects of fragmentation on pollinator abundance and fruit set of an abundant understory palm in a Mexican tropical forest. *Biological Conservation* 141:375–384.
- Ashburner GR, Faure MG, James EA, Thompson WK, Halloran GM. 2001. Pollination and breeding system of a population of Tall Coconut Palm *Cocos nucifera* L. (Arecaceae) on the Gazelle Peninsula of Papua New Guinea. *Pacific Conserv Biol.* 6:344–351.
- Barfod AS, Burholt T, Borchsenius F. 2003. Contrasting pollination modes in three species of *Licuala* (Arecaceae: Coryphoideae). *Telopea* 10:207–223.
- Barfod AS, Hagen M, Borchsenius F. 2011. Twenty-five years of progress in understanding pollination mechanisms in palms (Arecaceae). *Annals of Botany* 108:1503–1516.
- Barrett SCH, Kohn J. 1991. The genetic and evolutionary consequences of small population size in plant: implications for conservation. In: Falk D, Holsinger KE, editors. *Genetics and conservation of rare plants*. New York: Oxford University Press. p. 3–30.
- Baudouin L, Baril C, Clement-Demange A, Leroy T, Paulin D. 1997. Recurrent selection of tropical tree crops. *Euphytica* 96:101–114.

- Bawa KS. 1974. Breeding systems of tree species of a lowland tropical community. *Evolution* 28:85–92.
- Bergsdorf T. 2006. Forest Fragmentation and Plant-Pollinator Interactions in Western Kenya. Dissertation Doktorgrades, Rheinische Friedrich-Wilhelms-Universität. Bonn, Germany.
- Berton LHC, Azevedo Filho JA, Siqueira WJ, Colombo CA. 2014 Seed germination and estimates of genetic parameters of promising macaw palm (*Acrocomia aculeata*) progenies for biofuel production. *Industrial Crops and Products* 51:258–266.
- Billotte N, Marseillac N, Risterucci A-M, Adon B, Brottier P, Baurens F-C, Singh R, Herrán A, Asmady H, Billot C, Amblard P, Durand-Gasselin T, Courtois B, Asmono D, Cheah SC, Rohde W, Ritter E, Charrier A. 2005. Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis* Jacq.). *Theoretical and Applied Genetics* 110:754–765.
- Bora PS, Rocha RVM. 2004. Macaíba palm: fatty and amino acids composition of fruits. *Ciencia y Tecnología Alimentaria* 4:158–162.
- Brito GG, Caixeta ET, Gallina AP, Maciel-Zambolim E, Zambolim L, Diola V, Loureiro ME. 2010. Inheritance of coffee leaf rust resistancion and identification of AFLP markers linked to the resistance gene. *Euphytica* 173:255–264.
- Bullock SH. 1985. Breeding systems in the flora of a tropical deciduous forest in Mexico. *Biotropica* 17:287–301.
- Burquez A, Sarukhan J, Pedroza AL. 1987. Floral biology of a primary forest palm, *Astrocaryum mexicanum* Liebm. *Bot. J. Linn. Soc.*, 94:407–419.
- Carneiro FS, Lacerda AEB, Leme MR, Gribel R, Kanashiro M, Wadt LHO, Sebbenn AM. 2011. Effects of selective logging on the mating system and pollen dispersal of *Hymenaea courbaril* L. (leguminosae) in the Eastern Amazon as revealed by microsatellite analysis. *Forest Ecology and Management* 262:1758–1765.
- Charlesworth D, Morgan MT, Charlesworth B. 1990. Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. *Evolution* 44:1469–1489.
- Charlesworth D, Willis JH. The genetics of inbreeding depression. 2009. *Nature Reviews Genetics* 10:783–796.

- Choo J, Juenger TE, Simpson BB. 2012. Consequences of frugivore-mediated seed dispersal for the spatial and genetic structures of a neotropical palm. *Molecular Ecology* 21:1019–1031.
- 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. *Industrial Crops and Products* 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.
- Corley RHV, Tinker PB. 2003. *The oil palm*, 4th edn. Blackwell, Oxford 562p.
- Cockerham CC, 1969. Variance of gene frequencies. *Evolution* 23:72–84.
- Del Castillo RF, Trujillo S. 2008. Effect of inbreeding depression on outcrossing rates among populations of a tropical pine. *New Phytologist* 177:517–524.
- Dudash MR, Fenster CB. 2000. Inbreeding and outbreeding depression in fragmented populations. In: Young AC and Clarke GM (eds) *Genetics, Demography and Viability of Fragmented Populations*, Cambridge University Press: Cambridge, UK. p. 35–54.
- Dudash MR, Fenster CB. 2001. The role of breeding system and inbreeding depression in the maintenance of an outcrossing mating strategy in *Silene virginica* (caryophyllaceae). *American Journal of Botany* 88:1953–1959.
- 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* 69:217–228.
- Eguiarte LE, Burquez A, Rodriguez J, Martinez-Ramos M, Sarukhan J, Pinero D. 1993. Direct and indirect estimates of neighborhood and effective population size in a tropical palm, *Astrocaryum mexicanum*. *Evolution* 47:75–87.
- Eiserhardt WL, Svenning JC, Kissling WD, Balslev H. 2011. Geographical ecology of the palms (Arecaceae): determinants of diversity and distributions across spatial scales. *Annals of Botany* 108:1391–1416.

- Escobar JS, Nicot A, David P. 2008. The different sources of variation in inbreeding depression, heterosis and outbreeding depression in a metapopulation of *Physa acuta*. *Genetics* 180:1593–1608.
- Escobar, J. S., J. R. Auld, A. C. Correa, J. M. Alonso, Y. K. Bony, M.-A. Coutellec, J. M. Koene, J.-P. Pointier, P. Jarne, and P. David. 2011. Patterns of mating-system evolution in hermaphroditic animals: correlations among selfing rate, inbreeding depression, and the timing of reproduction. *Evolution* 65:1233–1253.
- Facelli JM, Pickett STA. 1991. Plant litter: its dynamics and effects on plant community structure. *The Botanical Review* 57:1–32.
- Farias Neto JT, Clement CR, Resende MDV. 2013. Estimativas de parâmetros genéticos e ganho de seleção para produção de frutos em progênies de polinização aberta de pupunheira no estado do Pará, Brasil. *Bragantia* 72:122–126.
- Federman S, Hyseni C, Clement W, Oatham MP, Caccone A. 2014. Habitat fragmentation and the genetic structure of the Amazonian palm *Mauritia flexuosa* L.f. (Arecaceae) on the island of Trinidad. *Conserv Genet* 15:355–362.
- Fenster CB, Galloway LF. (2000). Inbreeding and outbreeding depression in natural populations of *Chamaecrista fasciculata* (Fabaceae). *Conserv Biol* 14:1406–1412.
- Fox, C.W. 2005. Problems in measuring among-family variation in inbreeding depression. *American Journal of Botany* 92:1929–1932.
- Fox CW, Reed DH. 2010. Inbreeding depression increases with environmental stress: an experimental Study and meta-analysis. *Evolution* 65:246–258.
- Fuchs E, Hamrick JL. 2011. Mating system and pollen flow between remnant populations of the endangered tropical tree, *Guaiacum sanctum* (Zygophyllaceae). *Conservation Genetics* 12:175–185.
- Fyfe JL, Bailey NTJ. 1951. Plant breeding studies in leguminous forage crops. 1. Natural cross breeding in winter beans. *The Journal of Agricultural Science* 41:371–378.
- Gaiotto FA, 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.
- Galetti M, Zipparro VB, Morellato LPC. 1999. Fruiting phenology and frugivory on the palm *Euterpe edulis* in a lowland Atlantic forest of Brazil. *Ecotropica* 5:115–122.

- Geninia J, Galettia M, Morellato LPC. 2009. Fruiting phenology of palms and trees in an Atlantic rainforest land-bridge island. *Flora* 204:131–145.
- Gonçalves DB, Batista AF, Rodrigues MQ, Nogueira KM, Santos VL. 2013. Ethanol production from macaúba (*Acrocomia aculeata*) presscake hemicellulosic hydrolysate by *Candida boidinii* UFMG14. *Bioresource Technology* 146:261–266.
- Goodwillie, C., S. Kalisz, and C. G. Eckert. 2005. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics* 36:47–79.
- Hardner C, Tibbits W. 1998. Inbreeding depression for growth and fecundity traits in *Eucalyptus nitens*. *Forest Genet* 5:11–20
- Hardy OJ, Vekemans X. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2:618–620.
- Hayes, C. N., J. A. Winsor, and A. G. Stephenson. 2005. Environmental variation influences the magnitude of inbreeding depression in *Cucurbita pepo* ssp. *Texana* (Cucuritaceae). *J. Evol. Biol.* 18:147–155.
- Henderson A. 1986. A review of pollination studies in the Palmae. *The Botanical Review* 52:221–259.
- Henderson A, Galeano G, Bernal R. 1995 ‘Field guide to the palms of the Americas.’ (Princeton University Press: New Jersey)
- Hiane PA, Ramos Filho MM, Ramos MIL, Macedo MLR. 2005. Bocaiuva, *Acrocomia aculeata* (Jacq.) Lodd., pulp and kernel oils: Characterization and fatty acid composition. *Brazilian Journal of Food Technology* 8:256–259.
- Husband B, Schemske D. 1996. Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution* 50:50–74.
- Instituto Brasileiro de Geografia e Estatística – IBGE. Censo Demográfico 2010. Disponível em: <http://www.censo2010.ibge.gov.br>
- Lande R, Schemske DW. 1985. The evolution of self-fertilization and inbreeding depression in plants I: genetic models. *Evolution* 39:24–40.

- Lanes ECM, Nick C, Kuki KN, Freitas RD, Motoike SY. 2013. Genomic DNA isolation of *Acrocomia aculeata* (Arecaceae) from leaf and stipe tissue samples for PCR analysis. *Genetics and Molecular Research* 12:3905–3911.
- Lanes, ECM, Costa PMA, Motoike SY. 2014. Alternative fuels: Brazil promotes aviation biofuels. *Nature* 511:31–31.
- Lanes ECM, Motoike SY, Kuki KN, Nick C, Freitas RD. 2015. Molecular characterization and population structure of *Acrocomia aculeata* (Arecaceae) ex situ germplasm collection using microsatellites markers. *Heredity* 106:102–112.
- Laurance WF, Lovejoy TE, Vasconcelos HL, Bruna EM, Didham RK, Stouffer PC, Gascon C, Bierregaard RO, Laurance SG, Sampaio E. 2002. Ecosystem decay of Amazonian forest fragments: a 22-year investigation. *Conservation Biology* 16:605–618.
- Laurance WF, Merona JMR, Andrade A, Laurance SG, D'Angelo S, Lovejoy TE, Vasconcelos HL. 2003. Rain-forest fragmentation and the phenology of Amazonian tree communities. *Journal of Tropical Ecology* 19:343–347.
- Lepsch-Cunha N, Lund CA, Hamilton MB. 2003. Isolation and characterization of nuclear microsatellite loci in the tropical arboreal palm *Oenocarpus bacaba* (Arecaceae). *Molecular Ecology Notes* 3:435–437.
- Lewis PO, Zaykin D. 2001. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.1 (d16c). Free program distributed by the authors over the internet from <http://lewis.eeb.uconn.edu/lewishome/software.html>.
- Loiselle BA, Sork VL, Nason J, Graham C. 1995. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* 82:1420–1425.
- Luyindula N, Mantantu N, Dumortier F, Corley RHV. 2005. Effects of inbreeding on growth and yield of oil palm. *Euphytica* 143:9–17.
- Lynch M, Walsh B. 1998. Genetics and analysis of quantitative traits. Sunderland (MA): Sinauer Associates.
- Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how computer program Cervus accommodates genotyping error increase success in paternity assignment. *Mol. Ecol.* 16:1099–1106.

- Keller L, Waller D. 2002. Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17:230–241.
- Khanduri VP, Sharma CM, Kumar KS, Ghildiyal SK. 2013. Annual Variation in Flowering Phenology, Pollination, Mating System, and Pollen Yield in Two Natural Populations of *Schima wallichii* (DC.) Korth. *The Scientific World Journal* 2013: 350157.
- Mantovani A, Morellato LPC. 2000. Fenologia da floração, frutificação, mudança foliar e aspectos da biologia floral. *Sellowia* 49–52:23–38.
- Marshall TC, Slate J, Kruuk LEB, Pemberton JM. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 7:639–655.
- Marten-Rodríguez S, Quesada M, Momen B. 2012. Effects of Local Plant Density and Incomplete Dichogamy on the Reproductive Success of the Rare Neotropical Palm *Geonoma epetiolata*. *Biotropica* 44:680–688.
- Melser C, Bijleveld A, Klinkhmer PGL. 1999. Late-acting inbreeding depression in both male and female function of *Echium vulgare* (Boraginaceae). *Heredity* 83:162–170.
- Mora-Urpí J, Weber JC, Clement CR. 1997. Peach palm. *Bactris gasipaes* Kunth. Promoting the conservation and use of underutilized and neglected crops. 20. Institute of Plant Genetics and Crop Plant Research, Gatersleben/ International Plant Genetic Resources Institute, Rome, Italy.
- Motoike SY, Lopes FA, Sá Júnior AQ, Carvalho M, Oliveira MAR. 2007. Processo de germinação e produção de sementes pré-germinadas de palmeiras do gênero *Acrocomia*. (Protocolo INPI 014070005335).
- Motoike SY, Kuki KN. 2009. The potential of macaw palm (*Acrocomia aculeata*) as source of biodiesel in Brazil. *International Review of Chemical Engineering* 1:632–635.
- Murawski DA, Hamrick JL. 1991. The effect of the density of flowering individuals on the mating systems of nine tropical tree species. *Heredity* 67:167–174.
- Mustajärvi K, Siikamäki P, Åkerberg A. 2005. Inbreeding depression in perennial *Lychnis viscaria* (Caryophyllaceae): Effects of population mating history and nutrient availability. *Am J Bot* 92:1853–1861.

- Namoff S, Husby CE, Francisco-Ortega J, Noblick LR, Lewis CE, Griffith MP. 2010. How well does a botanical garden collection of a rare palm capture the genetic variation in a wild population? *Biological Conservation* 143:1110–1117.
- Nayak, K. G. & P. Davidar. 2010. Pollinator limitation and the effect of breeding systems on plant reproduction in forest fragments. *Acta Oecologica* 36:191–196.
- Nazareno AG, Reis MS. 2012. Linking Phenology to Mating System: Exploring the Reproductive Biology of the Threatened Palm Species *Butia eriospatha*. *Journal of Heredity* 103:842–852.
- Nazareno AG, Reis MS. 2014. At Risk of Population Decline? Na Ecological and Genetic Approach to the Threatened Palm Species *Butia eriospatha* (Arecaceae) of Southern Brazil. *Journal of Heredity* 105:120–129.
- Nucci SM, Azevedo-Filho JA, Colombo CA, Priolli RHG, Coelho RM, Mata TL, Zucchi MI. 2008. Development and characterization of microsatellites markers from the macaw. *Molecular Ecology Resources* 8:224–226.
- Núñez-Avellaneda LR, Bernal R, Knudsen JT. 2005. Diurnal palm pollination by mytropical beetles: is it weather-related? *Plant Systematics and Evolution* 254:149–171.
- Ottewell K, Grey E, Castillo F, Karubian J. 2012. The pollen dispersal kernel and mating system of an insect-pollinated tropical palm, *Oenocarpus bataua*. *Heredity* 109: 332–339.
- Pandin DS. 2009. Inbreeding depression analysis based on morphological characters in four generations of selfed mapanget tall coconut N°32 (*Cocos nucifera* L.). *Indonesian Journal of Agriculture* 2:110–114.
- Paschke M, Bernasconi G, Schmid B. 2005. Effects of inbreeding and pollen donor provenance and diversity on offspring performance under environmental stress in the rare plant *Cochlearia bavarica*. *Basic Appl Ecol* 6:325–338.
- Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288–295.
- Pires TP, Souza ES, Kuki KN, Motoike SY. 2013. Ecophysiological traits of the macaw palm: A contribution towards the domestication of a novel oil crop. *Industrial Crops and Products* 44:200–210.

- Portela RCQ, Santos FAM. 2014. Impact of forest fragment size on the population structure of three palm species (Arecaceae) in the Brazilian Atlantic Rainforest. *Revista de Biología Tropical* 64:433–442.
- Ramos SLF, Lopes MTG, Lopes R, Cunha RNV, Macêdo JLV, Contim LAS, Clement CR, Rodrigues DP, Bernardes LG. 2011. Determination of the mating system of Tucumã palm using microsatellite markers. *Crop Breeding and Applied Biotechnology*. 11:181–185.
- Resende MDV. 2002. *Genética Biométrica e Estatística no Melhoramento de Plantas Perenes*. 1th edn. Embrapa Informação Tecnológica, 975p.
- Resende MDV, Barbosa MHP. 2005. *Melhoramento genético de plantas de propagação assexuada*. 1th edn. Colombo: Embrapa Florestas. 130p.
- Ritland K, Jain S. 1981. A model for the estimation of outcrossing rate and gene frequencies using independent loci. *Heredity* 47:35–52.
- Ritland K. 1989. Correlated matings in the partial selfer *Mimulus guttatus*. *Evolution*, 43:848–859.
- Ritland K. 2002. Extensions of models for the estimation of mating systems using independent loci. *Heredity* 88:221–228.
- Rogers DL and Montalvo AM. Genetically appropriate choices for plant materials to maintain biological diversity. University of California, 2004. Report to the USDA Forest Service, Rocky Mountain Region, Lakewood, CO. http://www.fs.usda.gov/Internet/FSE_DOCUMENTS/fsbdev3_039080.pdf (acessado em 11/08/2014).
- Röser M, Johnson MAT, Hanson L. 1997. Nuclear DNA amounts in palms (Arecaceae). *Botanica Acta* 110:79–89.
- Rousselle Y, Thomas M, Galic N, Bonnin I, Goldringer I. 2011. Inbreeding depression and low between-population heterosis in recently diverged experimental populations of a selfing species. *Heredity* 106:289–299.
- Rueda RAP. 2014. *Avaliação de germoplasma para o melhoramento e a conservação da macaúba*. Tese (Doutorado em Fitotecnia) – Universidade Federal de Viçosa, Viçosa – MG.

- Salm R, Jalles-Filho E, Schuck-Paim C. 2005. A model for the importance of large arborescent palms in the dynamics of seasonally-dry Amazon forest. *Biota Neotropica* 5:1–6.
- Saro I, Robledo-Arnuncio JJ, González-Pérez MA and Sosa PA. 2014. Patterns of pollen dispersal in a small population of the Canarian endemic palm (*Phoenix canariensis*). *Heredity* 113:215–223.
- Sebbenn AM. 2006. Sistema de reprodução em espécies arbóreas tropicais e suas implicações para a seleção de árvores matrizes para reflorestamentos ambientais. In: Higa AR, Silva LD, editors. *Pomares de sementes de espécies florestais nativas*. Curitiba: FUPEF, p. 193–138.
- Shapcott A, Dowe JL, Ford H. 2009. Low genetic diversity and recovery implications of the vulnerable Bankoualé Palm *Livistona carinensis* (Arecaceae), from north-eastern Africa and the southern Arabian Peninsula. *Conservation Genetics* 10: 317–327
- Scariot A., Lleras E., Hay JD. 1991. Reproductive biology of the palm *Acrocomia aculeata* in Central Brazil. *Biotropica* 23:12–22
- Shapcott, A., Dowe, J.L., Ford, H., 2009. Low genetic diversity and recovery implications of the vulnerable Bankoualé Palm *Livistona carinensis* (Arecaceae), from north-eastern Africa and the southern Arabian Peninsula. *Conservation Genetics* 10:317–327.
- Schemske, D. W., and R. Lande. 1985. The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. *Evolution* 39:41–52.
- Sezen UU, Chazdon RL, Holsinger KE. 2007. Multigenerational genetic analysis of tropical secondary regeneration in a canopy palm. *Ecology* 88:3065–3075.
- Silberbauer-Gottsberger I. 1990. Pollination and evolution in palms. *Phyton* 30:231–233.
- Silva JC, Barrichelo LEG, Brito JO. 1986. Endocarpos de babaçu e macaúba comparados a madeira de *Eucalyptus grandis* para a produção de carvão vegetal. *IPEF* 34:31–34.
- Singh R, Nagappan J, Tan SG, Panandam JM, Cheah SC. 2007. Development of simple sequence repeat (SSR) markers for oil palm and their application in genetic

- mapping and fingerprinting of tissue culture clones. *Asia Pacific Journal of Molecular Biology and Biotechnology* 15:121–131.
- Smouse PE, Sork VL. 2004. Measuring pollen flow in forest trees: an exposition of alternative approaches. *Forest Ecology and Management* 197:21–38.
- Tabarelli M, Mantovani W, Peres CA. 1999. Effects of habitat fragmentation on plant guild structure in the ontane Atlantic Forest of southeastern Brazil. *Biological Conservation* 91:119–127.
- Thiele J, Hansen T, Siegismund HR, Hauser TP. 2010. Genetic variation of inbreeding depression among floral and fitness traits in *Silene nutans*. *Heredity* 104:52–60.
- Uzzo RP, Bovi MLA, Spiering SH, Saes LA. 2002. Correlações fenotípicas entre caracteres vegetativos e de produção de palmito da palmeira real australiana. *Scientia Agricola*, 59:505–511.
- Waller D M, Dole J, Bertsch AJ. 2008. Effects of stress and phenotypic variation on inbreeding depression in *Brassica rapa*. *Evolution* 62:917–931.
- Wandeck FA, Justo PG. 1988. A macaúba, fonte energética e insumo industrial: sua significação econômica no Brasil. In: SIMPOSIO SOBRE O CERRADO, SAVANA, Brasília. EMBRAPA, CPAC, p. 541–577.
- Waser NM, Price MV, Shaw RG. 2000. Outbreeding depression varies among cohorts of *Ipomopsis aggregata* planted in nature. *Evolution* 54:485–491.
- Weller SG, Sakai A K, Thai DA, Tom J, Rankin AE. 2005. Inbreeding depression and heterosis in populations of *Schiedea viscosa*, a highly selfing species. *Journal of Evolutionary Biology* 18:1434–1444.
- Wilcock C, Neiland R. 2002. Pollination failure in plants: why it happens and when it matters. *Trends in Plant Science* 7:270–277.
- Winn AA, Elle E, Kalisz S, Cheptou P-O, Eckert CG, Goodwillie C, Johnston MO, Moeller DA, Ree RH, Sargent RD, Mario V-M. 2011. Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution* 65:3339–3359.
- Zangerl AR, Berenbaum MR, e Nitao JK. 1991. Parthenocarpic fruits in wild parsnip: decoy defence against a specialist herbivore. *Evolutionary Ecology* 5:136–145.

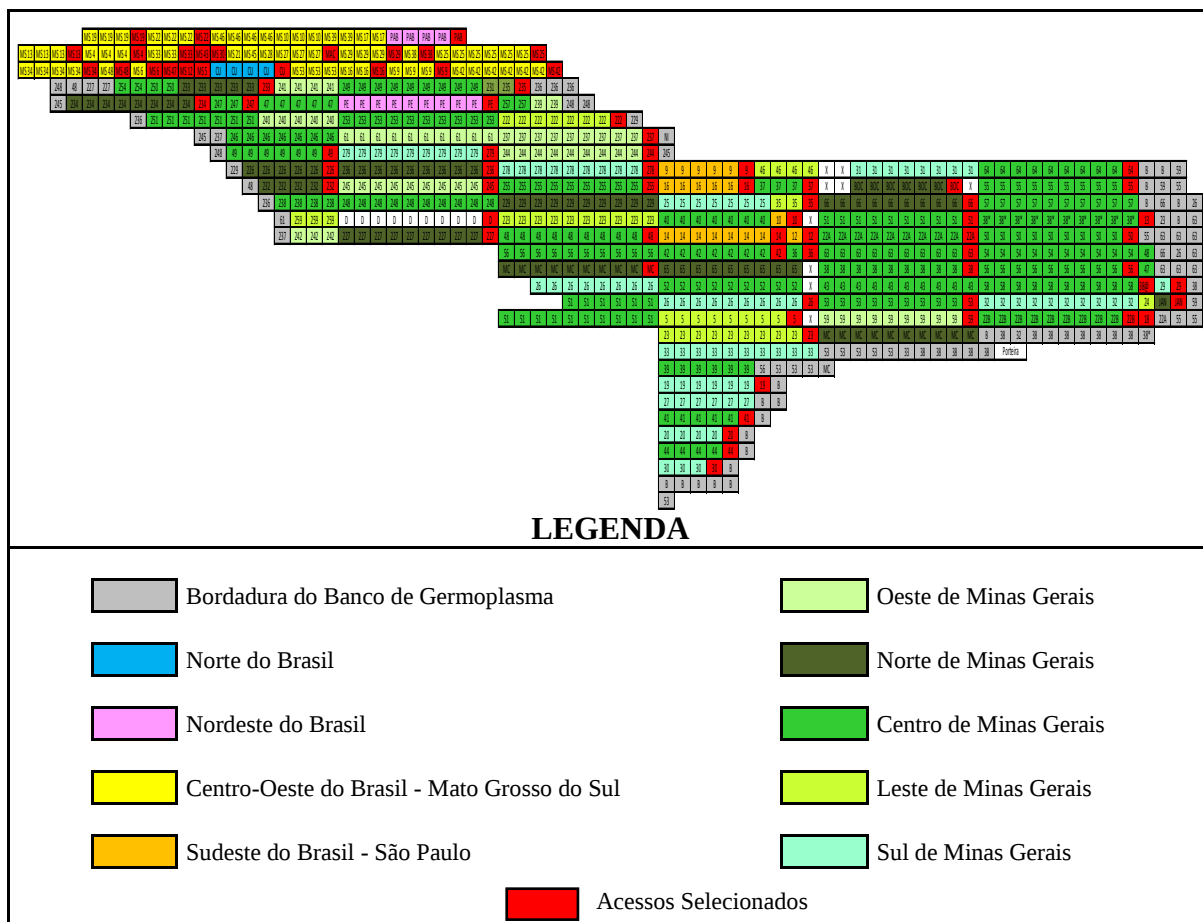
- Zapata TR, Arroyo MTK. 1978. Plant reproductive ecology of a secondary deciduous tropical forest in Venezuela. *Biotropica* 10:221–23.
- Zona S, Henderson A.1989. A review of mediated seed dispersal of palms. *Selbyana* 11:6–21.

Conclusões Gerais

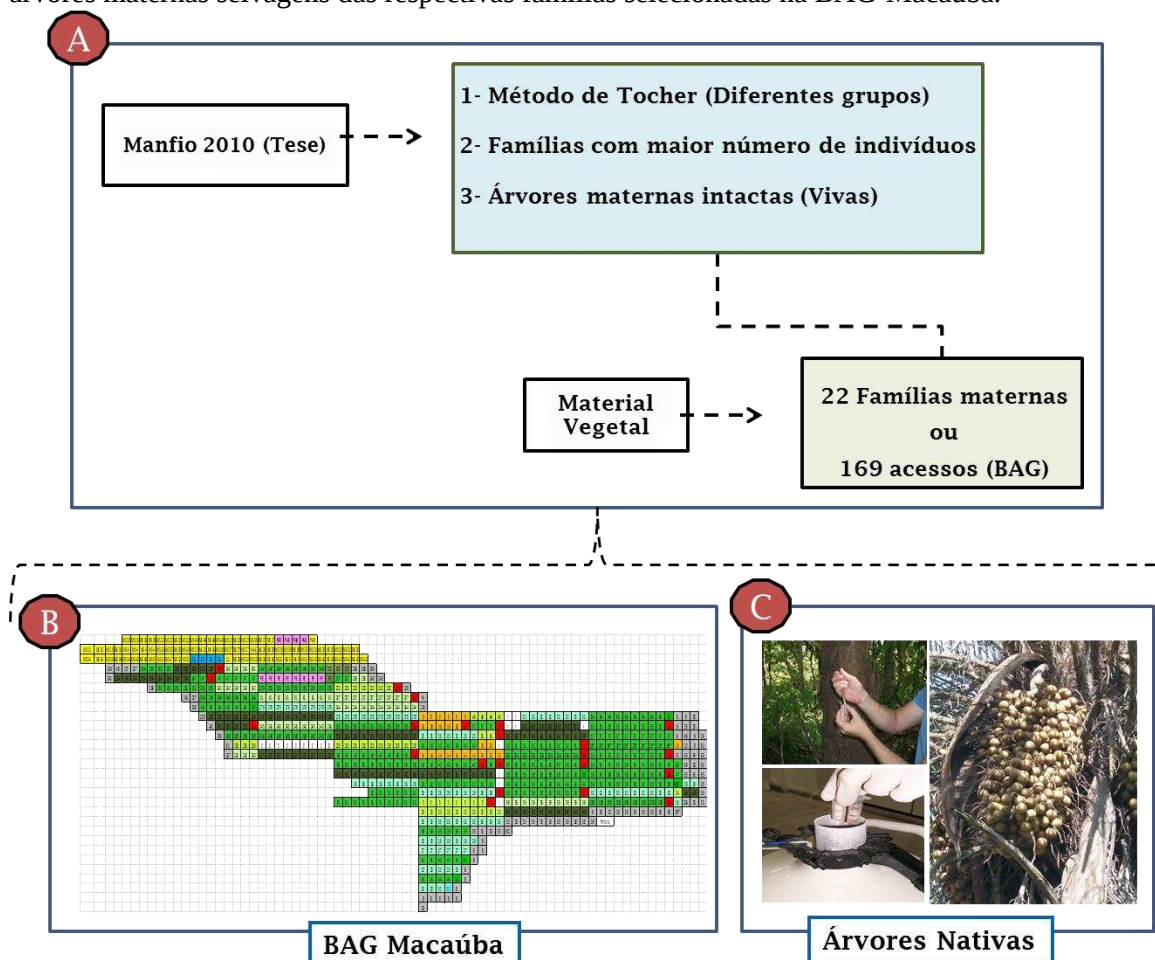
As informações moleculares disponibilizadas neste estudo contribuíram para o preenchimento da lacuna científica existente. Conhecimento que era necessário para a adoção de estratégias mais eficientes para a conservação e o melhoramento genético da espécie *A. aculeata*, sumarizados nos Apêndices A — F. Neste âmbito, os resultados genotípicos indicaram uma abundante variabilidade alélica presente nos acessos mantidos na coleção *ex situ* (BAG-Macaúba), tal como a grande variabilidade morfoagronômica previamente constatada no banco. Os quatro pools gênicos detectados dentro da coleção *ex situ* foram correspondentes com as origens geográficas dos acessos, além de apresentarem uma ampla variabilidade genética tanto dentro, como entre pools genéticos. Destes, três pools gênicos de ampla variabilidade alélica estão no estado de Minas Gerais, sugerindo a presença de um possível centro de diversidade nesta região. Os resultados moleculares indicaram que o sistema de reprodução da *A. aculeata* é predominantemente por fecundação cruzada, ou seja, trata-se de uma espécie alógama com sistema de auto-incompatibilidade parcial. Os cruzamentos não são aleatórios, sendo a maior fração das progênies compostas por irmãos-completos. Isto se deve as altas taxas de cruzamentos biparentais e entre indivíduos aparentados, possivelmente decorrentes de fatores reprodutivos e ecológicos intrínsecos da espécie *A. aculeata*, notadamente o vetor polinizador e dispersor de semente. O coeficiente de parentesco determinado neste estudo foi 105% maior que a esperada em uma população idealizada, portanto assumir que progênies de polinização aberta de *A. aculeata* são meios-irmãos gera superestimativas na variância genética aditiva, herdabilidade, valores genéticos e ganhos na seleção. Ainda que a magnitude da depressão por endogamia tenha variado entre as famílias maternas de *A. aculeata*, níveis elevados de depressão foram detectados em todos os caracteres vegetativos estimados. A superioridade fenotípica detectada em algumas progênies endogâmicas em relação às progênies híbridas pode estar relacionada com o alto mérito genético das famílias. A confirmação desta superioridade genética, bem como a capacidade específica de combinação destas famílias, possibilitaria a implementação da Seleção Recorrente Recíproca Individual no programa de melhoramento, conforme o Apêndice E. Quanto aos programas de conservação genética *ex situ* e *in situ* de *A. aculeata*, estes devem priorizar o número de árvores matrizes para a coleta de sementes em uma determinada população natural.

APÊNDICES

Apêndice A – Croqui do Banco de Germoplasma de Macaúba (BAG-Macaúba). São indicados na cor vermelha os acessos (um indivíduo de cada família materna) de *Acrocomia aculeata* selecionados para o capítulo 1 da tese. A seleção do material genético foi baseada em dois critérios: (1) representatividade da coleção viva mantida na BAG-Macaúba e (2) os estados e regiões do Brasil onde foram realizadas as coletas.

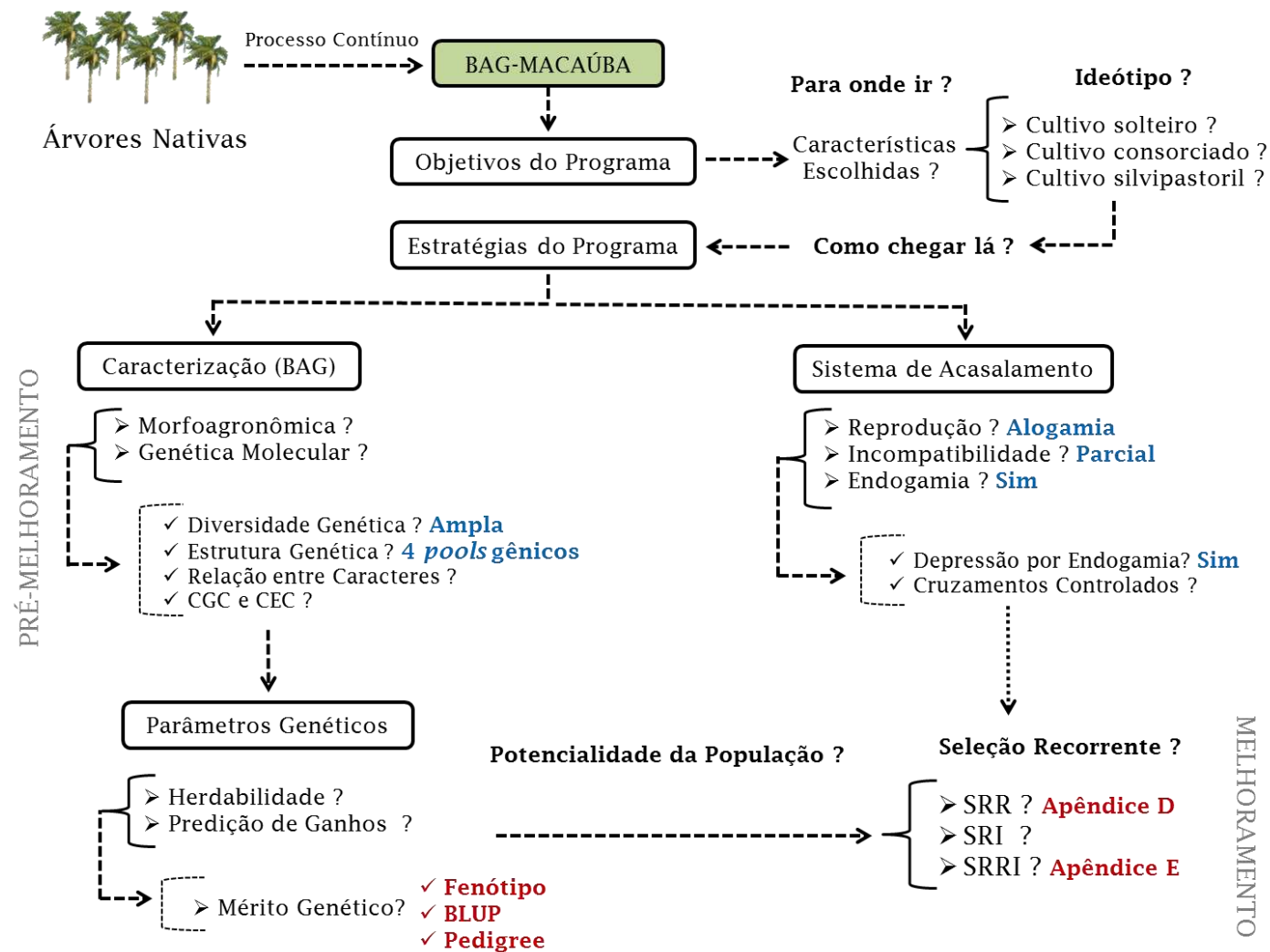


Apêndice B – Esquematização do critério adotado para a seleção das famílias maternas de *Acrocomia aculeata* do capítulo 2 da tese. (A) Seleção do material genético com base nos resultados de diversidade genética de Manfio (2010); (B) Croqui do Banco de Germoplasma de Macaúba indicando na cor vermelha as famílias selecionadas; (C) Coleta do tecido vegetal das árvores maternas selvagens das respectivas famílias selecionadas na BAG-Macaúba.

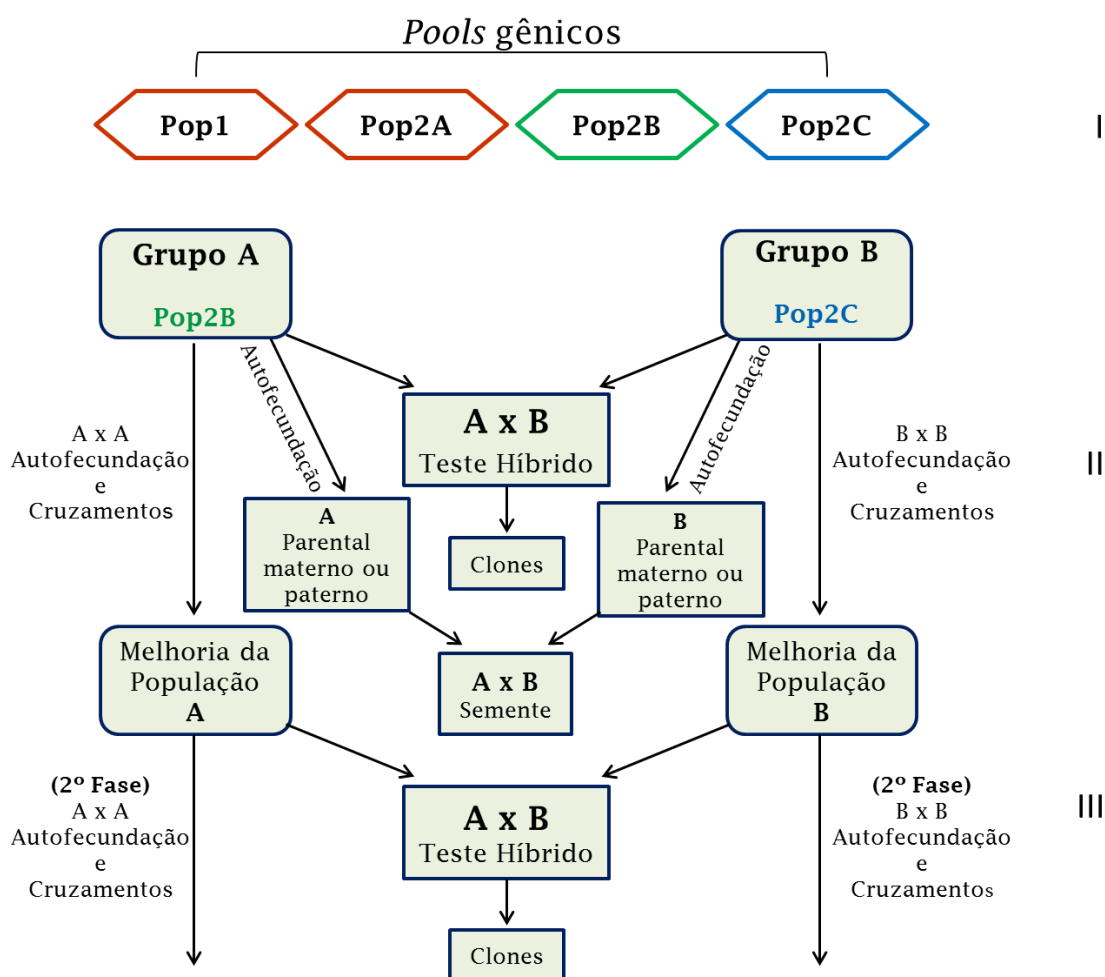


A seleção das famílias maternas teve como critério os resultados de Manfio (2010) sobre diversidade genética dos acessos de *A. aculeata* a partir de características biométricas do fruto de matrizes selvagens. O critério de seleção das famílias maternas deu-se da seguinte forma: (i) Famílias pertencentes a diferentes grupos estabelecidos pelo método de otimização de Tocher; (ii) Preferência das famílias com um maior número de acessos dentro da BAG-Macaúba. (iii) Famílias que possuem árvores maternas devidamente geoferrerienciadas e intactas (vivas) no campo. Das 22 famílias maternas escolhidas, 21 pertenceram ao estado de Minas Gerais devido aos aspectos de produtividade destas plantas (Manfio, 2010), e uma única família do estado de São Paulo. As árvores maternas das respectivas famílias mantidas na BAG também foram inseridas no estudo. Ao final, a análise do sistema de cruzamento foi determinada apenas nas 19 famílias maternas com número igual ou superior a quatro indivíduos dentro de cada família.

Apêndice C – Racionalização para o melhoramento da *Acrocomia aculeata*. As informações geradas por este estudo permitiu a resposta para algumas etapas do programa e são indicadas em negrito na cor azul. Respostas em negrito na cor vermelha são possíveis procedimentos que podem ser incorporados no programa de melhoramento. CGC: Capacidade Geral de Combinação; CEC: Capacidade Específica de Combinação; SRR: Seleção Recorrente Recíproca; SRI: Seleção Recorrente Intrapopulacional; SRRI: Seleção Recorrente Recíproca Individual.

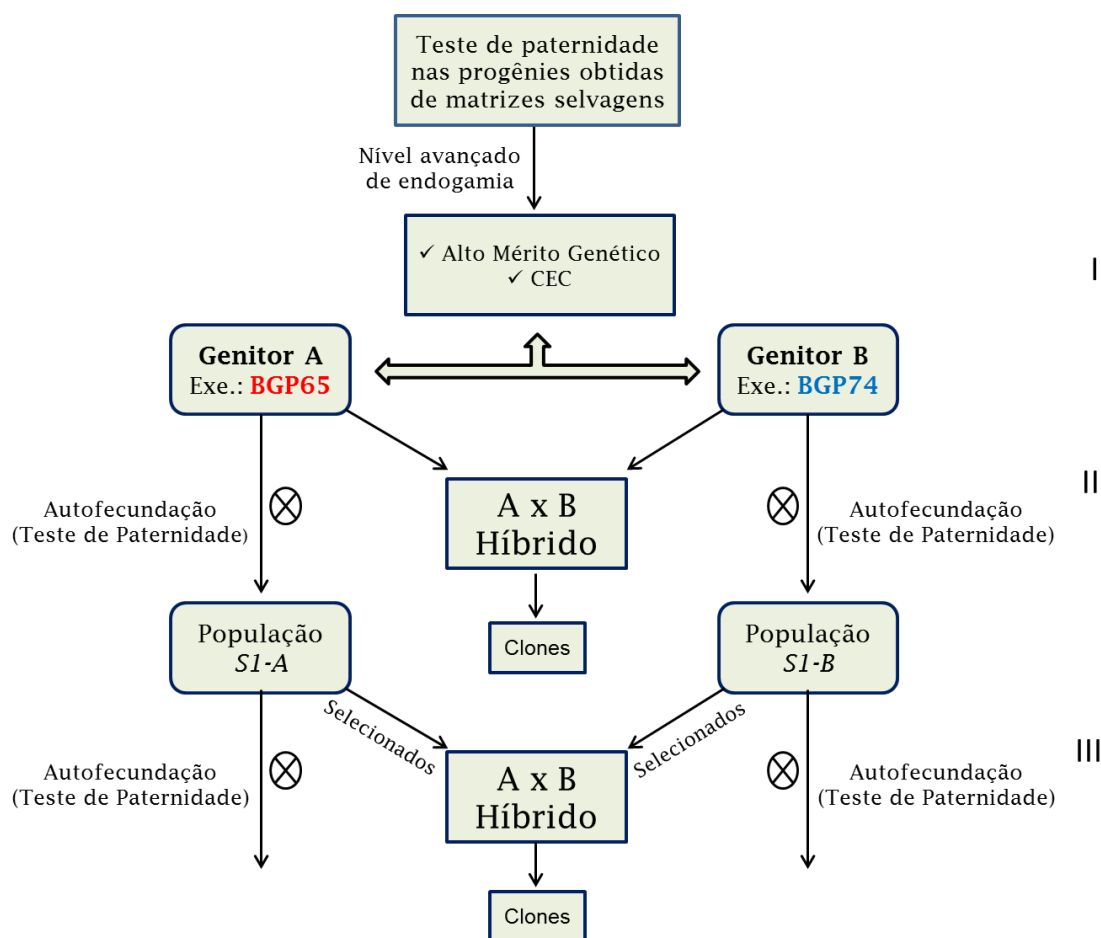


Apêndice D – Método de Seleção Recorrente Recíproca ou Interpopulacional (SSR), sugerido para o programa de melhoramento da *Acrocomia aculeata*.



Esquema do método de Seleção Recorrente Interpopulacional ou Recíproca (SSR) para o programa de melhoramento genético da macaúba. (I) Pools gênicos detectados no Banco de Germoplasma de Macaúba relatado no capítulo 1 da tese. Seleção de acessos de *A. aculeata* dentro de cada pool gênico para compor uma das duas populações de base do programa de melhoramento. Cada pool genético será composto por acessos de diferentes origens geográficas, além das informações das características morfoagronômicas disponíveis para cada indivíduo; (II) Os genitores das melhores combinações híbridas obtidas por cruzamentos teste entre as duas populações são autofecundados e usados para a produção de sementes (repetição dos cruzamentos entre os selecionados do teste híbrido). A expectativa é de que ambos os efeitos da CGC e CEC sejam novamente reproduzidos. Os melhores indivíduos dentro das duas populações são autofecundados e cruzados para formar a próxima geração; (III) A partir de cada população melhorada, cruzamentos testes são realizados e o ciclo é novamente repetido. Os híbridos integrarão os ensaios experimentais visando liberar clones comerciais, ou mesmo podem ser lançados como variedades (adaptado de Corley & Tinker, 2003).

Apêndice E – Método de Seleção Recorrente Recíproca Individual (SSRI), sugerido para o programa de melhoramento da *Acrocomia aculeata*.



Esquema do método de Seleção Recorrente Recíproca Individual (SSRI) com auxílio do teste de paternidade para o programa de melhoramento genético da macaúba. (I) Identificação das progênie endogâmicas obtidas de árvores nativas (autofecundação ou de cruzamento entre aparentados), com alto valor genotípico e capacidade específica de combinação; (II) Os indivíduos de alto desempenho são cruzados (A x B) para criar uma combinação híbrida, e ao mesmo tempo submetida à autofecundação para eliminação da carga genética e geração de uma população de autofecundação (S1); (III) Os indivíduos superiores são selecionados de cada população de autofecundação, e cruzados (S1-A x S1-B) para criar uma nova combinação híbrida. Os híbridos integrarão os ensaios experimentais visando liberar clones comerciais, ou mesmo podem ser lançados como variedades. Além disso, a autofecundação pode ser confirmada por teste de paternidade em cada etapa de geração da população de autofecundação. No exemplo do esquema acima, as famílias BGP65 e BGP74 pertencentes a diferentes pools gênicos (respectivamente, Pop2A e Pop2C, conforme o capítulo 1 da tese), tiveram acessos endogâmicos de cada uma destas famílias identificados por teste de paternidade no capítulo 2. Contudo, ainda é necessário confirmar o mérito genético e a capacidade específica de combinação para uso da SSRI (adaptado de Resende & Barbosa, 2005).

Genomic DNA isolation of *Acrocomia aculeata* (Arecaceae) from leaf and stipe tissue samples for PCR analysis

E.C.M. Lanes, C. Nick, K.N. Kuki, R.D. Freitas and S.Y. Motoike

Departamento de Fitotecnia, Universidade Federal de Viçosa,
Viçosa, MG, Brasil

Corresponding author: E.C.M. Lanes

E-mail: edercml@gmail.com

Genet. Mol. Res. 12 (3): 3905-3911 (2013)

Received January 21, 2013

Accepted July 23, 2013

Published September 23, 2013

DOI: <http://dx.doi.org/10.4238/2013.September.23.9>

ABSTRACT

Macaw palm, *Acrocomia aculeata* is an oleaginous species of the Arecaceae family; it has been identified as one of the most promising plants for sustainable production of renewable energy, especially biodiesel. We developed an efficient protocol of genomic DNA extraction for *A. aculeata* using leaf and stipe tissues, based on the cationic hexadecyltrimethylammonium bromide method, and we evaluated the quantity, purity and integrity of the resultant DNA. We also determined whether these procedures interfere with PCR amplification using SSR molecular markers. The lowest concentration of DNA was obtained from stipe tissues (135 ng/μl), while fresh leaf tissues provided the highest concentration of DNA (650 ng/μl). Good quality DNA was obtained from fresh leaf, lyophilized leaf and stipe tissues (relative purity, 1.79 – 1.89 nm). Differences in quantity and quality of DNA extracted from different tissues did not interfere with general patterns of PCR amplification based on SSR markers.

Key words: Molecular marker; Oleaginous species; Optimization DNA; PCR amplification; SSRs; Macaw palm

INTRODUCTION

Macaw palm, *Acrocomia aculeata* (Jacq.) Lodd. ex Martius, is an oleaginous species of the family Arecaceae. Commonly found in the tropical Americas, this palm tree is widespread in Brazil, with the highest densities found in the states of Minas Gerais, Goiás, Mato Grosso and Mato Grosso do Sul (Henderson et al., 1995). This palm has been identified as one of the most promising plants for the sustainable production of renewable energy, especially biodiesel, based on its high yields and resilience to adverse environmental conditions. Furthermore, it shows several possibilities for the food and cosmetic industries (Motoike and Kuki, 2009; Abreu et al., 2011).

However, in areas where *A. aculeata* is exploited, it is often carried out in a rudimentary way, resulting in a low yield and poor quality products. The breeding program for this species is one of the key strategies to ensure its production sustainability (Moura et al., 2009; Manfio et al., 2012). In this context, the application of molecular biology techniques has been an essential tool for the studies of germplasm conservation, association mapping, phylogeny, mapping of QTL (quantitative trait loci) and analysis of population and structure. However, the success of these studies is very dependent on an efficient extraction of DNA (Borges et al., 2012).

High quality DNA samples must be free of contaminants, or otherwise, they can ruin the reactions. Most of all, the impurities can interfere in the action of restriction enzymes, Taq polymerase in PCR reactions or electrophoretic separation of DNA fragments (Molinari and Crochemore, 2001). An efficient protocol for DNA isolation provides good conditions for PCR amplification allowing reproducibility and clear bands. Several protocols for DNA extraction have been optimized and successfully applied to plant species (Shahzadi et al., 2010). The lyophilization of the tissue, if used prior to the extraction of DNA, could reduce degradation problems in the samples, which allows storage and preservation of the plant tissue for a longer period of time (Nunes et al., 2011).

Besides, sampling foliar tissue may present some restrictions especially in mature palm trees due to their height. Mature palm trees are usually 16 to 20 m tall, which is an obstacle for efficient leaf sampling especially in the wild. As an alternative to the leaf, samples of stipe tissue can be easily obtained at the ground level. However, stipe tissue is more susceptible to oxidation due to its high phenolic contents.

The objective of this study was to standardize an efficient protocol for genomic DNA extraction from leaf (fresh and lyophilized) and stipe tissues of *A. aculeata* for genetic studies by using simple sequence repeat (SSR) molecular markers.

MATERIAL AND METHODS

Plant material

Foliar and stipe samples of *A. aculeata* were collected from the Zona da Mata and Central part of Minas Gerais State, as described below:

Foliar material: Leaflets were collected from non-senescent healthy fronds. They were excised from the rachis and immediately placed in liquid nitrogen and sent to the Laboratório de Biotecnologia e Melhoramento Vegetal - DFT/ UFV, in Viçosa, MG. As soon as they arrived in the laboratory, a group of samples were randomly selected for immediate DNA extraction. Another group of samples was lyophilized for 72 h and then stored for 30 days in the freezer at -20°C.

Stipe material: Stipe material was obtained after a cleanup of the trunk to eliminate debris, thorns and petiole remains until the stipe surface became visible. With a sharp disinfected blade the outer fibrous layer of the stipe was removed exposing the inner part of the organ. A sample of approximately 0.5 cm in thickness and 0.8 cm in diameter was then withdrawn with a cork cutter, transferred to an Eppendorf tube and placed in liquid nitrogen (Figure 1).



Figure 1. A. Macaw palm plant in its natural habitat. B. Exposed trunk after removal of debris. C. Stipe sample harvest with cylindrical cork cutter. D. and E. Stipe sample in Eppendorf tube. F. Storage of samples in liquid nitrogen.

Protocol

The protocol for DNA extraction was based on that of Doyle and Doyle (1990) with main modifications that followed the works of Faleiro et al. (2002) and Dellaporta (1983). The modifications involved the concentration of cationic hexadecyltrimethylammonium bromide (CTAB); the exclusion of 2-mercaptoethanol and the addition of sodium dodecyl sulfate (SDS). The modified CTAB extraction buffer consisted of 0.11 M Tris-HCl (pH 8.0), 55 mM EDTA (pH 8.0), 1.4 M NaCl, and CTAB (2.8%, w/v). A supplement of SDS (1.82%, v/v) was added to each sample after the buffer. The steps for both foliar and stem samples followed very similar steps, and the final protocol is described below.

Procedures for DNA extraction, quantification and quality determination

For the procedures, 200 mg leaf or stipe tissue were used.

1. The leaflets, previously cut into small pieces, were placed in 2.0-mL microtubes and hand mashed using a glass pestle with liquid nitrogen, until a powdery material was obtained. Immediately after maceration, 1.0 mL of extraction buffer was added to each microtube.
2. For stipe tissue, grinding it with a mortar and pestle before the addition of liquid nitrogen, helped to obtain a fine powder.
3. After grinding, 100 μ L 20% SDS (w/v) were added with constant mixing to obtain a homogeneous mixture.
4. The mixture samples, containing leaf or stem tissue, were incubated at 65°C for 60 min; the microtubes were inverted every 10 min.
5. After incubation, the samples were cooled down to room temperature for about 5 min.
6. A volume of 450 μ L chloroform-isoamyl alcohol (24:1, v/v) was added; microtubes were inverted for 10 min.
7. The samples were then centrifuged at 15,682 g for 10 min.
8. The aqueous phase (900 μ L) was transferred to a new microtube.
9. The steps 4, 5 and 6 were repeated once but transferring only 700 μ L of aqueous phase to a new microtube. The remained material was discarded.
10. To precipitate the nucleic acids, cold isopropanol (1 volume of 700 μ L) was added to the collected aqueous phase and then placed in a freezer at -20°C for at least 2 h.

11. After standing in the freezer, the mixture was centrifuged at 15,682 g for 10 min and the supernatant discarded. (Note: stipe samples were centrifuged at 16,168 g for 15 min to allow pellet formation and its fixation to the bottom of tube).
12. The pellet was washed twice with (300 μ L) cold 70% (v/v) ethanol and once with cold 95% (v/v) ethanol.
13. The pellet was air-dried for 1 h at room temperature
14. After cooling, the pellet was dissolved in 70 μ L 0.1X TE buffer with RNase (4 mL 10 mg/mL RNase solution per mL 0.1X TE) and incubated at 65°C for 10 min.
15. An aliquot (2 μ L) of each sample was submitted to electrophoresis on a 1% agarose gel in 1X TBE buffer containing 10 mg/mL ethidium bromide.
16. The concentration of DNA was measured using spectrophotometry, assuming an equivalency of 50 μ g/mL for 1 U absorbance at 260 nm.
17. DNA quality was evaluated by the A260/A280 nm ratio.
18. The DNA samples were stored at -20°C until further use.

PCR amplification and gel electrophoresis

Amplification reaction was conducted using an Applied Biosystems Veriti™ thermal cycler (Applied Biosystems). The SSR locus used were: i) Aacu07 (F: ATCGAAGGCCCTCCAATACT and R: GAAATAAGGGGACCCTCCAA) and ii) Aacu12 (F: GAATGTGCGTGCTCAAAATG and R: AATGCCAAGTGACCAAGTCC) (Nucci et al., 2008). The total volume of each reaction was 20 μ L containing 30 ng genomic DNA, 100 μ M dNTPs, 0.1 μ M primer SSR, 0.5 U Taq polymerase, and 1X PCR buffer. The thermocycler was programmed for one initial denaturation step of 5 min at 94°C and 30 cycles of 1 min at 94°C, 1 min at the specific annealing temperature for each primer pair, and 1 min at 72°C for extension, and a final extension at 72°C for 8 min. The resulting DNA fragments from the amplification were separated by electrophoresis on 6.0% denaturing polyacrylamide (w/v) gels [(19:1) acrylamide/bis-acrylamide (Bio-Rad), 7.5 M urea, 5.0X TBE], using the electrophoresis system "Sequi-Gen GT" (Bio-Rad) in 1X TBE buffer (0.09 M Tris base, 0.09 M boric acid and 2 mM EDTA, pH 8.0). The gels were visualized by silver nitrate staining.

RESULTS AND DISCUSSION

The DNA was successfully obtained from leaf and stipe samples of *A. aculeata*. The average concentration of DNA determined in a spectrophotometer obtained from fresh leaf samples was 650 ng/ μ L and from stipe was 135 ng/ μ L (Table 1).

Table 1. Quantification and purity of genomic DNA using 200 mg from leaf (lyophilized and fresh) and stipe tissues of *Acrocomia aculeata*.

Samples	Quantification (ng/ μ l)	Total of DNA (μ g)	Ratio (A260/A280 nm)
Stipe	135	9.45	$1.89 \pm 0.17^*$
Fresh Leaf	650	45.50	1.79 ± 0.20
Lyophilized Leaf	603	42.18	1.86 ± 0.24

*SD for n = 4

The leaf sample yielded 4.8 times more DNA than stipe tissue. The DNA yield difference between fresh leaf and stipe was confirmed in the analyses of the band pattern of DNA on agarose gel electrophoresis (Figure 2) where the bands of stipe DNA appeared less intense than the bands of leaf DNA (Figure 2a). Although less DNA had been extracted from stipe samples, the amount of DNA recovered was more than that expected. Usually CTAB DNA extraction protocols yield DNA with a concentration of 10 - 200 ng/ μ L extracted from 50 - 200 mg of leaf tissue (Ferreira and Grattapaglia, 1998). This result demonstrated the efficiency of the suggested protocol for leaf and stipe tissues.

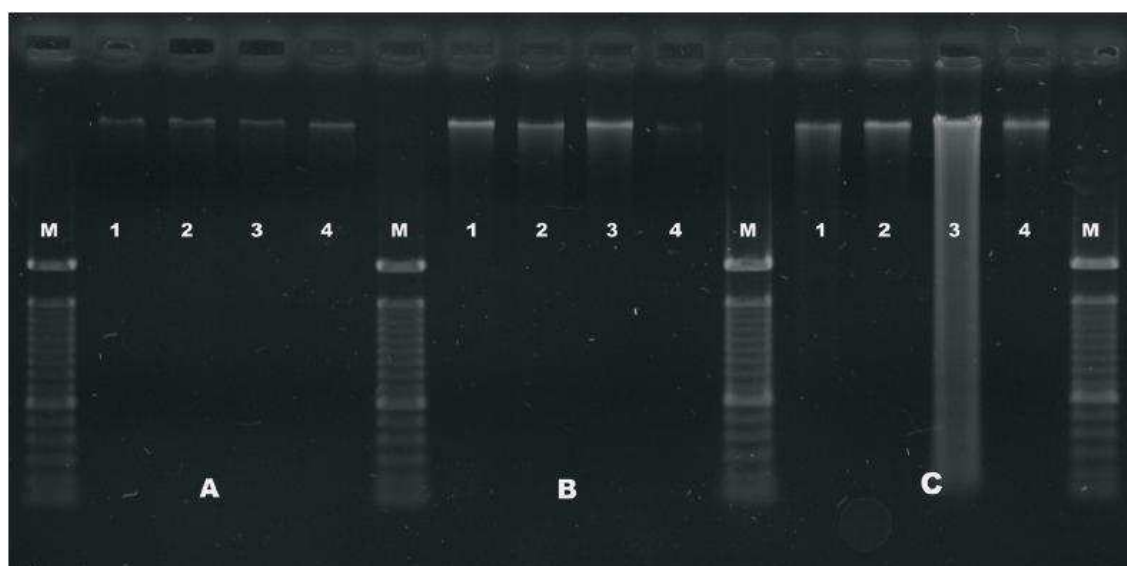


Figure 2. Genomic DNA extracted from specimens of *Acrocomia aculeata*, using Doyle and Doyle protocols modified. **A.** Stipe tissues, **B.** Fresh leaf tissues, and **C.** lyophilized leaf tissues. Lanes 1-4 of **A.**, **B.** and **C.** correspond to their respective tissues. Lane M = 200-bp DNA ladder.

Regardless of quantity, the quality of the DNA extracted from the fresh leaf and the stipe was high, with a relative purity of the DNA varying from 1.79 to 1.89. Pure DNA is supposed to show an A260/A280 nm ratio of 1.8–2.0. A lower ratio (<1.6) indicates contamination with protein and/or other contaminants, whereas a ratio higher than 2.0 indicates that the samples may be contaminated with phenol or chloroform (Sambrook and Russel, 2001; Barbosa, 1998).

The quality and quantity of DNA recovered from lyophilized leaf tissue determined with a spectrophotometer were similar to that obtained with DNA from fresh leaf tissue (Table 1). However, in agarose gel electrophoresis, the DNA from lyophilized leaf tissue showed vertical drags indicating possible DNA fragmentation (Figure 2c). However, this pattern did not interfere with PCR since all three sources of DNA obtained by the modified CTAB extraction protocol were successfully amplified by PCR using SSR primers (Figure 3).

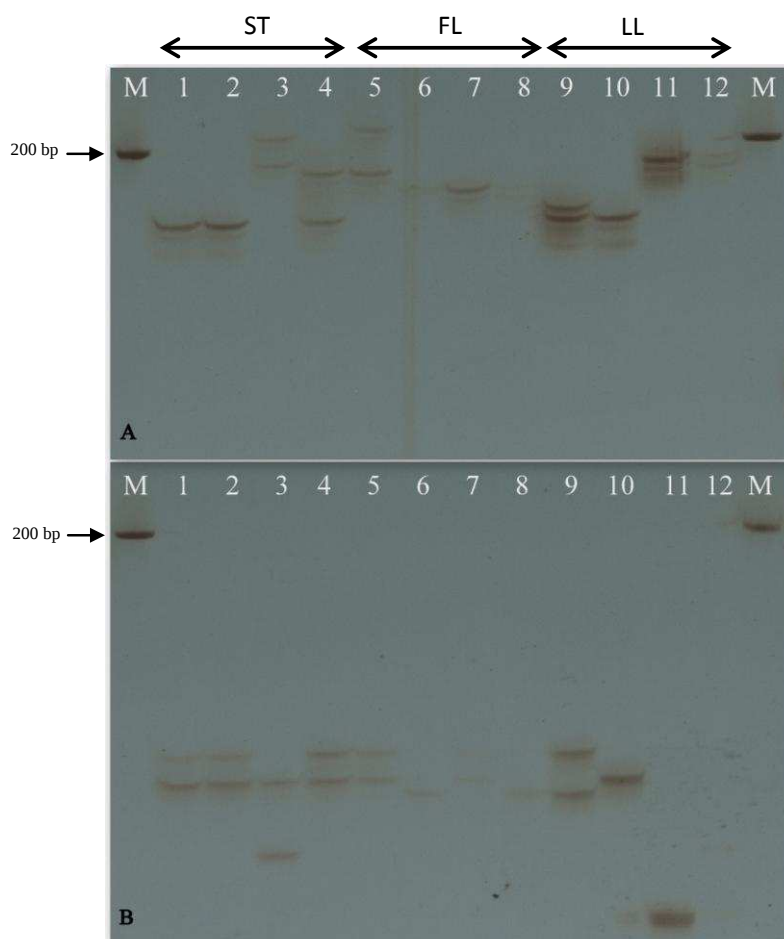


Figure 3. Standard SSR bands generated with primers Aacu12 (A) and Aacu07 (B). ST = Stipe tissues (Lanes 1-4); FL = fresh leaf tissues (Lanes 5-8); LL = lyophilized leaf tissues (Lanes 9-12). M = 200-bp DNA ladder.

These results confirmed that the proposed modified CTAB protocol is efficient for the extraction of *A. aculeata* DNA from different sources: fresh leaf, lyophilized leaf and stipe. Also, the proposed method has a high repeatability and accuracy which can be seen by the low values of standard deviation observed among DNA samples analyzed by spectrophotometry (Table 1). Moreover, the method is less laborious and safer, since 2-mercaptoethanol was eliminated from this protocol. 2-Mercaptoethanol is

volatile and toxic and requires additional care when handled (Molinari and Crochemore, 2001; Sambrook and Russel, 2001).

The stipe is a good source of DNA for molecular biology and genetic studies due to its accessibility, especially in mature wild plants (Novaes et al., 2009). Moreover, lyophilization is a helpful tool when a large number of samples are collected. It allows tissues to be stored for a long period of time for later processing.

ACKNOWLEDGMENTS

Research supported by Petrobras and CNPq.

REFERENCES

- Abreu IS, Carvalho CR, Carvalho GMA, et al. (2011). 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.* 59, p.149 - 155.
- Barbosa MM (1998). Quantificação e Controle da Qualidade do DNA Genômico. In: *Marcadores Moleculares em Plantas* (Milach S, ed.). UFRGS, Porto Alegre, 99-106.
- Borges DB, Amorim MB, Waldschmidt AM, et al. (2012). Optimization of DNA extraction from fresh leaf tissues of *Melanoxylon brauna* (Fabaceae). *Genet. Mol. Res.* 11 (2): 1586-1591.
- Dellaporta SL, Wood J and Hicks JB (1983). A plant miniprep: version II. *Plant Mol. Biol. Rep.* v.1, p.19-20.
- Doyle JJT and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus*, 12:13-15.
- Faleiro FG, Araújo IS, Bahia RCS, et al. (2002). Otimização da extração e amplificação de DNA de *Theobroma cacao* L. visando obtenção de marcadores RAPD. *Agrotróp.* 14 (2): 31 - 34.
- Ferreira ME and Grattapaglia D (1998). *Introdução ao Uso de Marcadores Moleculares em Análise Genética*. 3ª ed. Embrapa-Cenargen, Brasília.
- Henderson A, Galeano G, Bernal R (1995) 'Field guide to the palms of the Americas.' (Princeton University Press: New Jersey).
- Molinari, HB and Crochemore ML (2001). Extração de dna genômico de *Passiflora* spp. para análises PCR-RAPD. *Rev. Bras. Frutic.* v. 23, n. 2, p. 447-450.

- Manfio CE, Motoike SY, Resende MDV, et al. (2012). Avaliação de progênies de macaúba na fase juvenil e estimativas de parâmetros genéticos e diversidade genética. *Pes. Flor. Bras.*, v.32, p.63 – 68.
- Motoike SY and Kuki KN (2009). The potential of macaw palm (*Acrocomia aculeata*) as source of biodiesel in Brazil. *I. Re. Ch. E.* 1, 632–635.
- Moura EF, Motoike SY, Ventrella MC, et al. (2009). Somatic embryogenesis in macaw palm (*Acrocomia aculeata*) from zygotic embryos. *Sci. Hortic.* 119, 447–454.
- Novaes RML, Rodrigues JG and Lovato MB (2009). An efficient protocol for tissue sampling and DNA isolation from the stem bark of Leguminosae trees. *Genet. Mol. Res.*8: 86-96.
- Nucci SM, Azevedo-Filho JA, Colombo CA, et al. (2008). Development and characterization of microsatellites markers from the macaw. *Mol. Ecol. Resour.* 8:224-226.
- Nunes CF, Ferreira JL and Fernandes MCN (2011). An improved method for genomic DNA extraction from strawberry leaves. *Ciência Rural*, Santa Maria, v.41, n.8, p.1383-1389.
- Sambrook J and Russel DW (2001). *Molecular cloning: a laboratory manual*. Third edition. Cold Spring Harbor Laboratory Press, New York.
- Shahzadi I, Ahmed R and Hassan A, et al. (2010). Optimization of DNA extraction from seeds and fresh leaf tissues of wild marigold (*Tagetes minuta*) for polymerase chain reaction analysis. *Genet. Mol. Res.* 9 (1): 386-393.

Alternative fuels: Brazil promotes aviation biofuels

Éder Cristian Malta de Lanes, Paulo Mafra de Almeida Costa, Sérgio Yoshimitsu Motoike

Federal University of Viçosa, Minas Gerais, Brazil. edercml@gmail.com

Nature (London), v. 511, p. 31-31, 2014.

DOI: <http://dx.doi.org/10.1038/511031a>

The drive to produce biofuels that are suitable for aviation is starting to look promising (see go.nature.com/ujv5ne). For instance, the state government of Minas Gerais in Brazil has put in place policies to stimulate the cultivation, extraction and processing of the native macaw palm (*Acrocomia aculeata*), a potentially sustainable source of oil for producing biokerosene.

This tree can produce up to 6,200 kilograms of oil per hectare (see T. P. Pires et al. *Ind. Crops Prod.* **44**, 200–210; 2013) and can sequester as much as 4–5 times more carbon than other afforestation species such as *Eucalyptus camaldulensis* (see H. Seganuma et al. *J. Arid Land Stud.* **22**, 69–72; 2012).

To safeguard forests, macaw palms will be planted only on existing pastureland (estimated at 170 million hectares in 2010). Oil production from macaw palms, which could exceed the size of today's global palm-oil market, does not need to compromise food production or natural ecosystems (see go.nature.com/67tfia). The trees can be cultivated next to agricultural crops and their extensive canopy cover provides fauna with food and shelter, helping to restore biodiversity that has been lost by forest clearance.