INSTITUTO AGRONÔMICO

CURSO DE PÓS-GRADUAÇÃO EM AGRICULTURA TROPICAL E SUBTROPICAL

DESENVOLVIMENTO, CARACTERIZAÇÃO E ANÁLISE DA UTILIDADE DE MARCADORES MICROSSATÉLITES EM GENÉTICA DE POPULAÇÃO DE MACAÚBA

STELLA MARIS NUCCI

Orientador: Dr. Joaquim Adelino de Azevedo Filho
Co-orientadora: Dra. Maria Imaculada Zucchi

Dissertação submetida como requisito parcial para obtenção do grau de **Mestre** em Agricultura Tropical e Subtropical Área de Genética, Melhoramento Vegetal e Biotecnologia

Campinas, SP
Fevereiro 2007

Ao meu pai
Osvaldo,
DEDICO

À minha mãe
Elsa,
OFEREÇO

AGRADECIMENTOS

- Ao pesquisador, amigo e orientador Dr. Joaquim Adelino de Azevedo Filho, pela confiança e oportunidade importantes no curso e na minha vida profissional;

- À pesquisadora, amiga e co-orientadora Dra. Maria Imaculada Zucchi, pela confiança e ensinamentos, pelo auxílio, amizade e atenção dispensada durante a realização deste trabalho;

- Ao pesquisador, colaborador e amigo Dr. Carlos Colombo por toda colaboração e incentivo;

- Aos professores da Área de Concentração em Melhoramento Vegetal da PG-IAC, pela amizade, incentivo e pelos conselhos e ensinamentos constantes transmitidos;

- À Rauly, técnica do laboratório de cultura de tecidos pela ajuda e cooperação;

- Aos funcionários da PG-IAC, pelo auxílio e atenção no decorrer do curso;

- A todos os colegas da pós-graduação, cujos nomes não citarei, porque graças a Deus são muitos;

- À pesquisadora Marilene Leão Alves Bovi (*in memoriam*), pela colaboração, incentivo e amizade;

- A FAPESP, por financiar a execução desse projeto;

- Ao meu pai por todo suporte financeiro, por acreditar na minha capacidade e por me incentivar a concluir este trabalho, pelo carinho, amor e atenção;

- A minha mãe pela paciência, dedicação, compartilhando comigo todos os momentos alegres e difíceis no decorrer do curso, pelo carinho, amor e atenção;

- À Regina e ao Thiaguinho pela colaboração e amizade;

- A todos do laboratório de Biologia Molecular do Centro de Pesquisa e Desenvolvimento de Recursos Genéticos Vegetais;

- A todos que colaboraram para a realização e finalização deste trabalho;

- A Deus por todas as bênçãos recebidas e por estar sempre guiando minha vida e iluminando meu caminho.

Lamento da Terra

A Terra geme, grita, espera...
Procura socorro, quem vai socorrer?
Eu olho assustada o Homem, a fera,
Que procura matar e acha o morrer.

Os rios já não correm e nem são transparentes,
Arrastam-se chorando cobertos de espuma.
No fundo só mortes, químicas diferentes,
Vão mortos cansados sem vida nenhuma.

E voam as aves em vôos incertos,
Querendo abrigo, querem procriar.
Em busca de oásis só acham desertos,
No fruto o veneno, onde se alimentar?

O machado trabalha, e é sem descanso
Que ouço os ecos da grande tombada.
O homem contempla o que julga avanço,
É o verde que morre na grande queimada.

Eu choro por ti, Homem louco insensato,
Que cego caminha rumo a destruição.
Lamento por ti natureza, és de fato
Vítima inocente da grande ambição.

Que os céus te vinguem com tempestades...
Que trema a Terra, nua, abatida.
Que o Homem entenda, antes que acabes,
Que és fonte de bênçãos, o sustento da vida!

(*Elsa T. Zarpelon*)

SUMÁRIO

RESUMO..... vii
ABSTRACT..... viii
1 INTRODUÇÃO..... 01
2 REVISÃO DE LITERATURA..... 03
2.1 *Acrocomia aculeata*..... 03
2.1.1 Aspectos taxonômicos e botânicos..... 03
2.1.2 Uso da espécie..... 11
2.1.3 Potencial para exploração econômica..... 13
2.1.4 Programa biodiesel..... 16
2.1.5 Programa de agricultura familiar..... 18
2.2 Estrutura Genética de Populações de Plantas..... 19
2.3 Marcadores Moleculares..... 20
2.3.1 Marcadores microssatélites..... 25
2.3.2 Análise da estrutura genética com marcadores co-dominantes..... 28
3 MATERIAL E MÉTODOS..... 30
3.1 Material vegetal..... 30
3.1.1 Desenvolvimento de bibliotecas enriquecidas com locos microssatélites..... 30
3.1.2 Caracterização da estrutura genética populacional da espécie *Acrocomia aculeata*..... 30
3.2 Extração e Quantificação de DNA..... 32
3.3 Desenvolvimento e Caracterização de Marcadores Microssatélites..... 33
3.3.1 Construção de biblioteca genômica enriquecidas com microssatélites..... 33
3.3.2 Seleção e sequenciamento dos clones 36
3.3.3 Avaliação do polimorfismo dos locos microssatélites..... 41
3.3.4 Condições de amplificação dos locos microssatélites..... 42
3.3.5 Metodologia de análise estatística dos dados..... 43
4 RESULTADOS 44
4.1 Desenvolvimento de Bibliotecas Enriquecidas com Locos Microssatélites..... 44
4.2 Análise das Seqüências, Desenvolvimento e Utilização de *Primers*..... 44
4.3 Variação Genética..... 54
4.4 Estrutura Genética..... 58

5 DISCUSSÃO..... 61
6 CONCLUSÕES..... 68
7 REFERÊNCIAS BIBLIOGRÁFICAS..... 69
ANEXOS..... 78

NUCCI, Stella Maris. **Desenvolvimento, caracterização e análise da utilidade de marcadores microssatélites em genética de população de macaúba.** 2007. 84f. Dissertação (Mestrado em Genética, Melhoramento Vegetal e Biotecnologia) – Pós-Graduação – IAC.

RESUMO

Acrocomia aculeata é uma espécie de palmeira monocaule de ampla distribuição geográfica e com qualidades que a tornam importante do ponto de vista natural, ecológico e principalmente sócio-econômico, pois seus frutos são uma valiosa fonte de óleo vegetal com potencial para produção de biodiesel, proporcionando vantagens ambientais, econômicas e sociais. Este trabalho teve como objetivo o desenvolvimento de bibliotecas enriquecidas com locos microssatélite visando a análise da diversidade genética da espécie, utilizando nove populações e quatro regiões de ocorrência espontânea nos Estados de São Paulo e Minas Gerais. Foi observado que, apesar do potencial econômico, a espécie ainda é pouco estudada. Através da construção da biblioteca enriquecida com locos microssatélites foram desenvolvidos 77 *primers* para a espécie, dos quais foram escolhidos 36 para a genotipagem e identificados oito locos microssatélites utilizados para a caracterização e estudo da estrutura genética populacional. Foi encontrada uma taxa de cruzamento aparente média (t_a) igual a 0,471, sugerindo que esta espécie tenha sistema reprodutivo misto. As estimativas das estatísticas *F* de Wright (1965) revelaram o valor médio de F_{IT} = 0,51 e o F_{IS} = 0,35, confirmando que o sistema de reprodução é misto, a diversidade genética entre populações foi F_{ST} = 0,25 e R_{ST} = 0,21; a variação entre populações dentro de regiões foi de F_{SR} = 0,27 e, entre regiões, F_{RT} = 0,16, indicando maior importância na coleta de futuros genótipos entre populações dentro de reegiões do que entre regiões. A maior ocorrência de alelos exclusivos foi observada em populações mais antropizadas ou com alta taxa de endogamia As distâncias genéticas de Nei (1978) entre as populações variaram de – 0,0482 a 0,9061 e o dendrograma apresentou três agrupamentos distintos que apresentaram uma distribuição coerente com a ecologia, geografia e clima em que se encontram essas populações. Foram encontrados importantes bancos de sementes nos locais de coleta. O estudo da estrutura genética destas populações forneceu informações importantes que auxiliarão na identificação e conservação da variabilidade genética da espécie e auxiliando ainda nas diferentes etapas do pré-melhoramento dessa palmeira.

Palavras-Chave: *Acrocomia aculeata*, SSR, óleo vegetal, estrutura genética, sistema reprodutivo.

NUCCI, Stella Maris. **Development, characterization and analyses of the use of microsatellite markers in macaw palm population genetics.** 2007. 84f. Dissertation (Masters Degree in Genetic, Plant Breeding and Biotechnology) – Graduated School – IAC.

ABSTRACT

Acrocomia aculeata is a single-stemed palm tree with wide geographic distribution that has important natural, ecological and mainly social and economical features, due to its fruits that are a valuable source of plant oil, potentially useful in biofuel production in economically and socially advantageous systems. The present work aimed to develop microsatellite loci-enriched libraries in order to analysis of genetic diversity of the species employing nine populations and four regions of spontaneous occurrence of macaw palm in the States of São Paulo and Minas Gerais. Although potentially important for economical development, scientific research in macaw palm is still scarce. The generation of a microsatellite loci-enriched library has allowed us to obtain 77 primers for the species and from this total, 36 were chosen for plant genotyping and for the identification of eight microsatellite loci, which were employed to characterize and to investigate the population genetic structure. We have found an apparent rate of crossing (t_a) of 0.471, suggesting that the species has a mixed mating system. The estimates of Wright (1965) *F*-statistics have revealed average values F_{IT}= 0.51 and F_{IS} = 0.35, thus confirming the occurrence of a mixed mating system, genetic diversity between the populations was F_{ST} = 0.25 and R_{ST} = 0.21; the variation between the populations within an area was F_{SR}= 0.27 and between regions F_{RT}= 0.16, demonstrating that it is more important to collect future genotypes among the populations in the same region than among the geographic regions. The higher occurrence of exclusive alleles was observed in populations under higher anthropic influence or with higher inbreeding rates. Nei (1978) genetic distances between populations ranged from -0.0482 to 0.9061 and dendrogram analysis demonstrated the existence of three distinct groups coherent with ecological, geographical and climatic features of the sites where the populations occur. Important seed banks were found on harvesting sites. Studies of the genetic structure of these populations have provided important information to the identification and conservation of the genetic diversity of the species and to pre-improvement stages of breeding research for the species.

Keywords: *Acrocomia aculeata*, SSR, plant oil, genetic structure, system reproductive.

1 INTRODUÇÃO

A macaúba (*Acrocomia aculeata* (Jacq.) Lood. ex Mart.) é uma palmeira arborescente perene, frutífera, nativa de florestas tropicais, tipicamente brasileira e com ampla distribuição geográfica. No Brasil ocorre nos estados do Ceará, Minas Gerais, Mato Grosso, Mato Grosso do Sul, São Paulo e por toda região Sul. São palmeiras monóicas que apesar de autocompatíveis realizam fecundação cruzada entre indivíduos diferentes, revelando sistema reprodutivo misto (SCARIOT et al., 1995).

Esta espécie apresenta ampla utilidade e mais recentemente seu fruto tem despertado grande interesse sócio-econômico por sua capacidade de produção de óleo vegetal, considerando-se que a macaúba é citada como umas das principais fontes deste óleo no Brasil. Em meio a suas várias utilidades são relatados usos medicinais, alimentícios, cosméticos, entre outros. Entretanto a mais promissora e que mais se destaca é sua capacidade de produção de óleo vegetal que pode chegar a quatro mil litros por hectare por ano, lembrando que através de plantios racionais e de programas de melhoramento este valor pode aumentar consideravelmente, oferecendo potencial para produção de biodiesel, proporcionando vantagens ambientais, econômicas e sociais.

A utilização do óleo de macaúba para a produção de biodiesel com a participação da agricultura familiar elevaria a demanda por esta planta e ajudaria a promover o desenvolvimento sócio-econômico de muitas regiões problemáticas no país, através de programas de geração de empregos e renda contribuindo para erradicação da miséria, freando correntes migratórias internas por contribuir com a fixação das famílias a terra (BATISTA, 2006).

A macaúba foi escolhida para o desenvolvimento desse estudo por sua importância como uma das principais fontes de óleo vegetal no Brasil, que tem potencial para produção de biodiesel. Poucos estudos foram, até então, realizados com a espécie, sendo que a maioria relata a ocorrência de pragas, ou sua distribuição geográfica. Para garantir a conservação das populações naturais e possibilitar o pré-melhoramento da macaúba, sendo uma espécie pouco estudada, alguns estudos se fazem necessários. O estudo da estrutura genética de populações e da taxa de cruzamento aparente em macaúba é inédito e é de suma importância, pois traz relevantes subsídios

para a identificação e a conservação da variabilidade genética da espécie, auxiliando ainda nas diferentes etapas do pré-melhoramento dessa palmeira.

Para quantificar a variabilidade genética existente entre e dentro de populações de plantas, o uso de marcadores moleculares tem se revelado uma importante e eficiente ferramenta, cada dia mais acessível. Para este estudo a técnica de microssatélites foi escolhida por suas características intrínsecas de marcador estável, co-dominante, multialélico, robusto, altamente confiável, pela simplicidade operacional e rapidez de análise quando comparado com outras metodologias.

Este estudo teve como objetivos principais o desenvolvimento de bibliotecas enriquecidas com marcadores moleculares microssatélites e caracterização de locos em macaúba (*Acrocomia aculeata* (Jacq.) Lood. ex Mart.) visando estudos populacionais, tais como caracterização da diversidade genética de populações, determinação da taxa de cruzamento aparente e parâmetros afins, visando gerar resultados para a conservação e para o pré-melhoramento da espécie.

Os objetivos específicos deste estudo foram: a) isolar e seqüenciar locos que contenham regiões microssatélites no genoma de macaúba (*Acrocomia aculeata* (Jacq.) Lood. ex Mart.); b) desenhar *primers* específicos visando a amplificação dos locos de microssatélites isolados e selecionados; c) caracterizar cada loco de microssatélite isolado com relação ao tipo de repetição, número e tamanho de alelos amplificados, temperatura de reassociação, heterozigosidade observada e esperada em materiais divergentes de macaúba; d) estudo preliminar de caracterização e estrutura genética populacional da espécie.

2 REVISÃO DE LITERATURA

2.1 *Acrocomia aculeata*

2.1.1 Aspectos taxonômicos e botânicos

A espécie *Acrocomia aculeata*, ex-família Palmae, está classificada atualmente dentro da família Arecaceae, sendo o gênero *Acrocomia* constituído por apenas duas espécies – *A. aculeata* (Jacq.) Lodd. ex Mart. e *A. hassleri* (B. Rodr.) W. J. Hahn. Estas espécies diferem basicamente pelo tamanho dos indivíduos e por sua localização geográfica. Conforme HENDERSON et al. (1995), a *A. aculeata* possui porte maior, estando amplamente distribuída nas regiões secas da América Tropical (Figura 1). São palmeiras nativas de florestas tropicais, ocorrendo desde o sul do México até ao sul do Brasil, Paraguai e Argentina (MORCOTE-RIOS & BERNAL, 2001). No Brasil é abundantemente encontrada nos estados do Ceará, Minas Gerais, Mato Grosso, Mato Grosso do Sul, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul.

Esta palmeira é citada como a única espécie de porte arbóreo de áreas de pastagens (SCARIOT et al., 1991; 1995), ocorrendo principalmente em florestas latifoliadas semidecíduas (LORENZI, 1992). As informações sobre as necessidades edáficas da espécie são por vezes conflitantes. RATTER et al. (1996) cita a espécie como indicadora de solos mesotróficos do Brasil Central, MOTTA et al. (2002), concluem que a ocorrência de macaúba acompanha áreas de solos com maior fertilidade natural e vegetação primitiva de fisionomia florestal, mostrando que a espécie avança como pioneira, evitando extremos de deficiência de nutrientes e de água, enquanto que NOVAES (1952) afirma que a palmeira ocorre no Estado de São Paulo em solos pobres, cobertos pelo cerrado. A espécie *A. hassleri* está restrita à região de cerrado no Brasil e Paraguai. O termo *Acrocomia* tem origem grega e descreve a forma como as folhas estão dispostas na planta, "*Akron*" (cima) e *"Kome*" (cabeleira) sugerindo que as folhas estão dispostas no formato de uma coroa (NOVAES, 1952; HENDERSON et al., 1995).

Em termos taxonômicos, *Acrocomia aculeata* (Jacq.) Lodd. ex Mart. está classificada em:

- Reino: Plantae
- Divisão: Magnoliophyta
- Classe: Liliopsida
- Subclasse: Arecidae
- Ordem: Arecales
- Família: Arecaceae
- Subfamília: Arecoideae
- Tribo: Cocoeae
- Subtribo: Bactridinae
- Gênero: *Acrocomia*

- Espécie: *Acrocomia aculeata* (GRIN, 2006).

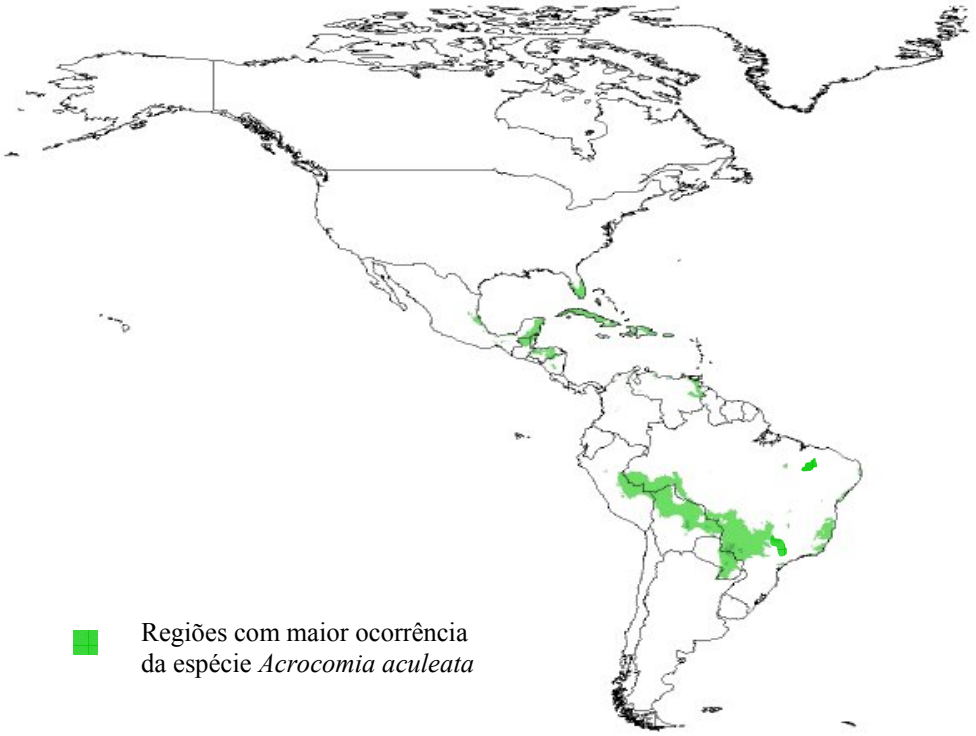


Figura 1 – Distribuição geográfica da espécie *Acrocomia aculeata* pela América Tropical (Fonte: NEW WORLD FRUITS, 2006 – adaptado).

Esta espécie apresenta como sinonímia botânica os seguintes nomes (HENDERSON et al., 1995): *Acrocomia antiguana* L.H. Bailey, *Acrocomia antioquensis* Posada-Arango, *Acrocomia belizensis* L.H. Bailey, *Acrocomia christopherensis* L.H. Bailey, *Acrocomia chunta* Covas & Ragonese, *Acrocomia erisacantha* Barb. Rodr., *Acrocomia fusiformis* Sweet, *Acrocomia glaucophylla* Drude, *Acrocomia grenadana* L.H. Bailey, *Acrocomia hospes* L.H. Bailey, *Acrocomia ierensis* L.H. Bailey, *Acrocomia intumescens* Drude, *Acrocomia karukerana* L.H. Bailey, *Acrocomia lasiospatha* Mart., *Acrocomia media* O.F. Cook, *Acrocomia mexicana* Karw. ex Mart., *Acrocomia microcarpa* Barb. Rodr., *Acrocomia mokayayba* Barb. Rodr., *Acrocomia odorata* Barb. Rodr., *Acrocomia panamensis* L.H. Bailey, *Acrocomia pilosa* Leon, *Acrocomia quisqueyana* L.H. Bailey, *Acrocomia sclerocarpa* Mart. *Acrocomia sclerocarpa var. wallaceana* Drude, *Acrocomia spinosa* (Mill.) H.E. Moore, *Acrocomia subinermis* León ex L.H. Bailey, *Acrocomia totai* Mart., *Acrocomia ulei* Dammer, *Acrocomia viegasii* L.H. Bailey, *Acrocomia vinifera* Oerst., *Acrocomia wallaceana* (Drude) Becc., *Bactris globosa* Gaertn., *Cocos aculeatus* Jacq., *Cocos fusiformis* Sw., *Palma spinosa* Mill.

Apresenta, também, uma vasta sinonímia popular no Brasil (LORENZI, 1992): macaúba, mucajá, mocujá, mocajá, macaíba, macaiúva, bacaiúva, bocaiúva,

umbocaiúva, imbocaiá, coco-de-catarro ou coco-de-espinho; o nome pode ainda variar de acordo com a região de distribuição da palmeira (FRUITS FROM AMERICA, 2005; NEW WORLD FRUITS, 2006): coco baboso, coco de catarro, coquito, macaúba, bocaiuva, mocayá (Argentina), corozo (Colômbia, Venezuela), coyol (América Central), catey, totaí (Bolívia), cayara, cayiete, ocori, tamaco (Colômbia); corosse (Haiti); groo groo, grugru palm, macaw palm, macaya oil (inglês); corosse (francês) e bocaiuva, cachinobí, catey, cayara, cayiete, coco baboso, coco de catarro, coquito, corozo, coyol, gru gru, macaúba, mocayá, ocori, palma de vino, tamaco, totaí (espanhol).

Devido à falta de estudos citogenéticos realizados com a espécie, citam-se alguns dados encontrados relacionados com a subfamília, e algumas informações encontradas sobre o gênero ou a espécie. SHARMA & SARKAR (1956) realizaram estudos citogenéticos com 50 espécies de vinte e oito gêneros de palmeiras, abrangendo algumas espécies de gêneros da subfamília Cocosoideae (atual Arecoideae). Neste estudo, os autores relatam que foram encontrados números de cromossomos variados em diferentes gêneros que pertencem às mesmas e diferentes subfamílias. Contudo, há uma homogeneidade geral e uma uniformidade de cromossomos claramente marcada dentro de cada subfamília e, todas as espécies estudadas são diplóides. Eles consideram a subfamília Cocosoideae como um bom exemplo de estabilidade evolutiva, pois em seis espécies diferentes, de gêneros diferentes, ele encontrou o mesmo número de cromossomos 2n = 32, e uma semelhança total nos cariótipos. Todos os cariótipos são caracterizados por cromossomos aproximadamente médios com gradação no tamanho para cromossomos curtos. O número de cromossomos com constrições secundárias varia de dois a três pares no complemento e, concluem, a semelhança entre eles é tão acentuada que podem ser considerados membros de uma única linha de evolução. No estudo da meiose, reunindo estas espécies de gêneros diferentes, revelou a formação de bivalentes, sugerindo possível origem destas por alopoliploidia. READ (1966), divulga nova contagem de cromossomos em Palmae (atual Arecaceae), que divide a subfamília Cocosoideae em dois grupos: espécies com n = 15, na qual se inclui a *Acrocomia aculeata* (Jacq.) Lood. ex Mart., e espécies com n = 16, portanto, dentro de toda subfamília foram encontradas apenas espécies 2n = 30 e espécies 2n = 32.

Conforme LORENZI (1992), TASSARO (1996), GRUPO VERDE (2006), FRUITS FROM AMERICA (2005) e MOBOT (2006) a espécie *Acrocomia aculeata* (Jacq.) Lood. ex Mart. é representada por palmeiras arborescentes (Figura 2), perenifólias, heliófitas, monóicas, cujo estipe pode atingir acima de 15 m de altura com 20 a 30 cm de diâmetro. Algumas plantas conservam no tronco por muitos anos os remanescentes da base das bainhas foliares, geralmente cobertos de espinhos escuros e compridos, cerca de 10 cm de comprimento (Figura 3).

As folhas são pinadas apresentando quatro a cinco metros de comprimento, geralmente em número de 20 a 30, coloração verde escuro, contemporâneas, distribuídas em diferentes planos dando um aspecto plumoso à copa. Os folíolos são aproximadamente em número de 130 para cada lado da região central da folha que é recoberta de espinhos (Figura 4).

Suas inflorescências amareladas apresentam-se agrupadas em cachos pendentes de até 80 cm de comprimento e protegidas por uma espata que pode atingir até dois metros de comprimento (Figura 5), são pequenas, unissexuais (Figura 6) e ambos os sexos estão presentes numa mesma inflorescência, as flores masculinas nascem no topo da inflorescência, enquanto as femininas situam-se na base. A floração ocorre entre

agosto e novembro. Estas inflorescências são andróginas com marcada protoginia. A polinização ocorre principalmente por besouros, com o vento desempenhando um papel secundário. Os principais polinizadores são *Andranthobius* sp. (Curculionidae), *Mystrops cf mexicana* (Nitidulidae) e *Cyclocephala forsteri* (Scarabaeidae). Mesmo a espécie sendo autocompatível e a gitonogamia respondendo por uma significativa porcentagem da produção de frutos, a xenogamia, polinização cruzada entre indivíduos diferentes, é responsável por muito do sistema reprodutivo.

Figura 2 – a.Vista geral da palmeira *Acrocomia aculeata*; b. Vista geral da copa da palmeira.

Figura 3 – a. Base das bainhas foliares cobertas de espinhos escuros e compridos; b. Visão aproximada das bainhas foliares.

Figura 4 – a. Vista geral da copa da palmeira; b. Vista da folha de macaúba.

Em um estudo realizado no Distrito Federal, no Brasil Central foi observado que cada inflorescência produz em média 60 frutos, mas esse número pode variar de 0 a 271, conforme a região (Figura 7), (SCARIOT et al., 1995). O fruto é globoso com 2,5 a 5,0 cm de diâmetro, liso, de coloração marrom-amarelada quando maduro (Figura 8), possui um epicarpo que se rompe facilmente quando maduro, apresenta mesocarpo branco ou amarelado, fibroso e mucilaginoso, cada fruto contém uma amêndoa oleaginosa, contendo de um a três embriões viáveis, envolvida por um endocarpo rígido. Frutifica durante o ano todo, principalmente de setembro a janeiro (Figura 9), (TASSARO, 1996; LORENZI, 1992; SCARIOT et al., 1991).

A Fundação Centro Tecnológico de Minas Gerais – CETEC, realizou estudos que avaliaram macaubais nativos em solos de média fertilidade, com produção aproximada de 100 kg por árvore por ano, e concluiu que, em plantios racionais com espaçamento 7 m x 7 m, pode se esperar uma produtividade mínima de 20.000 kg/ha/ano de frutos, considerando que esta produtividade pode aumentar muito através de manejo adequado, com a possível redução do espaçamento para 6 m x 6 m e programas de melhoramento da espécie. Outro fato importante a ser considerado é que ambos os espaçamentos permitem a implantação de consorciamentos de culturas alimentares nos três primeiros anos de implantação do macaubal, e o tempo necessário para a primeira colheita é de quatro a cinco anos, dependendo das formas de manejo que forem utilizadas. Em levantamento realizado em maio de 2006, o valor estimado para produção de um hectare de macaúba é de R$ 7.000,00 por ano, e a produção por hectare inclui 4.000 litros de óleo vegetal, 1.200 quilos de carvão vegetal, e 5.300 quilos de farelos para rações (OLIVEIRA, 2006).

Figura 5 – Inflorescências de macaúba protegidas pela espata.

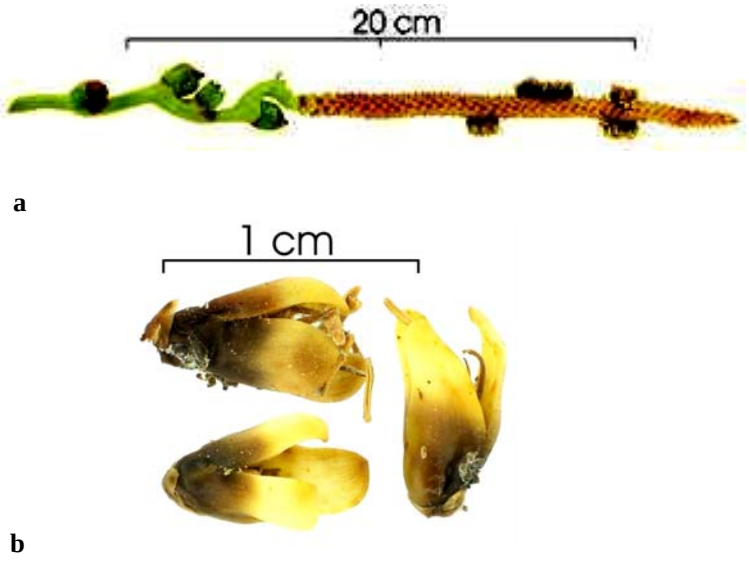
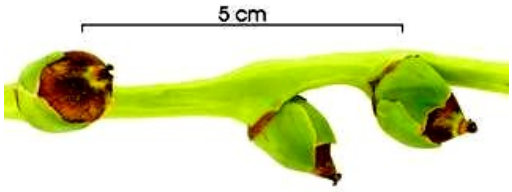


Figura 6 – a. Inflorescência em formação; b. Flores de macaúba (Fonte: DISCOVER LIFE, 2006).

Figura 7 – a. Cacho com frutos de macaúba; b. Macaúba produzindo vários cachos de frutos.

Figura 8 – Frutos maduros de macaúba. a. fruto com e sem epicarpo, evidenciando mesocarpo de coloração amarelada; b. fruto com e sem epicarpo, evidenciando mesocarpo de coloração esbranquiçada; c. fruto seco, destaque para endocarpo rígido; d. frutos cortados ao meio; e. amêndoas de macaúba; f. amêndoa e embrião.

a



b

Figura 9 – a. Cacho de frutos de macaúba em formação; b. frutos imaturos, (Fonte: DISCOVER LIFE, 2006).

2.1.2 Uso da espécie

LIMA et al. (2003) destacam a importância das palmeiras do ponto de vista natural, econômico e ecológico. São rotineiramente utilizadas como alimento pelo homem na forma de palmito, frutos e até produtos elaborados como doces, bebidas e óleos. Artesanato feito com os gêneros *Orbygnia*, *Syagrus*, *Acrocomia* e *Mauritia* é habitualmente comercializado em feiras e mercados de muitas cidades do Brasil Central. As palhas das palmeiras são utilizadas para artesanato e para construção de casas pelos índios Xavantes, a polpa e as sementes são utilizadas como alimento e ingredientes para pinturas, e os estipes em suas tradicionais corridas de tora. Várias espécies da fauna silvestre alimentam-se de suas folhas, polpa dos frutos (canídeos) e sementes (roedores de pequeno e médio porte, psitacídeos, etc.). A estipe é empregado em construções rurais, na confecção de ripas e calhas para água. Do miolo do estipe obtém-se uma fécula nutritiva. As folhas, além de forrageiras, fornecem fibras têxteis para a confecção de redes e linhas de pescar. O fruto é a parte mais importante da planta, cuja polpa é consumida *in natura*, ou usada para extração de óleo comestível; a amêndoa fornece óleo claro com qualidades semelhantes ao da oliveira (LORENZI, 1992). Dada a sua ampla utilidade, essas palmeiras vêm sendo utilizadas pelo homem desde tempos pré-históricos (cerca de 9.000 anos a.C.), como relatado por MORCOTE-RIOS & BERNAL (2001). LORENZI (2006) relata nove diferentes categorias de usos associadas a distintas partes de *Acrocomia aculeata*, a partir de extrativismo por comunidades locais, das quais se destacam os usos medicinal e alimentício (Tabela 1).

Tabela 1 – Usos associados a distintas partes de *Acrocomia aculeata*.

Partes da planta usadas	Categoria de usos	Finalidade de usos
plântulas (parte aérea)	medicinal	diurético, hipotensor
raízes	medicinal	diurético
estipe	estipe	parede, caibro, ripas, calhas para água, mourão, estacas
estipe (medula)	alimento	fécula nutritiva
estipe (meristema apical)	alimento	palmito
estipe (seiva)	alimento	vinho
estipe (seiva)	medicinal	febrífuga
folha	forragem	gado bovino, eqüino,ração animal
folha	fibra	chapéu, balaio, linha de pesca, redes
folha	outro	cobertura de casas
fruto (mesocarpo)	medicinal	fortificante
fruto (mesocarpo)	alimento	fruta, goma de mascar, doces, paçocas, geléias e cocadas
fruto (óleo do mesocarpo)	alimento	licor, sorvete
fruto (óleo do mesocarpo)	medicinal	analgésico (dor de cabeça e nevralgias)
fruto (óleo do mesocarpo)	cosmético	hidratante capilar
semente	outro	Substitui a brita no concreto
semente	artesanato	confecção de botões
semente (amêndoa)	alimento	coco, paçoca
semente (óleo da amêndoa)	alimento	óleo de cozinha

Continua

Tabela 1 – Usos associados a distintas partes de *Acrocomia aculeata*.

Partes da planta usadas	Categoria de usos	Finalidade de usos
semente (óleo da amêndoa)	combustível	lamparina, produção de energia

semente (óleo da amêndoa)	cosmético	hidratante capilar
semente (óleo da amêndoa)	medicinal	laxante
semente (óleo da amêndoa)	outro	ingrediente de sabão

2.1.3 Potencial para exploração econômica

Um bom exemplo da potencialidade econômica do uso da macaúba aliada ao manejo sustentável foi dado pelo produtor rural Cinéias Campelo, 76 anos, que já na década de 60, muito antes de se pensar em fazer pesquisas sobre a macaúba, despertou para a possibilidade de exploração econômica do óleo desta palmeira, abundante na fazenda Montes Verdes, Jequitibá (MG), onde vive desde a infância (AVELAR, 2006). Após saber da instalação de uma fábrica de óleo em Jaboticatubas, criou um maquinário para extração de óleo para macaúba (Figura 10), e vendendo este óleo para uma indústria de cerâmica da região criou seus sete filhos, a maioria deles formados e com curso superior. Com uma produção diária de 100 litros de óleo, Cinéias atraiu a atenção de investidores interessados em conhecer sua usina para implantar projetos semelhantes, recebeu a visita de indígenas que viram no processo de extração de óleo de macaúba uma alternativa de renda. Também estiveram na fazenda Montes Verdes estudiosos dos combustíveis biológicos renováveis, professores universitários e representantes da Fundação Rockfeller (E.U.A.).

Figura 10 – Maquinário criado pelo produtor rural Cinéias Campelo. Um velho britador adaptado com pequenos pinos quebra o coquinho da macaúba e separa a casca da amêndoa (a); em seguida, a amêndoa segue para o torrador (b); e depois disso para a prensa (c), de onde é retirado o óleo (d) e (e).

O óleo de soja utilizado na culinária pode ser perfeitamente substituído pelo óleo de macaúba, e os resíduos gerados podem ser usados na produção de ração para suínos, aves e caprinos. Este óleo ainda pode ser usado como lubrificante para máquinas, combustível, em substituição ao óleo diesel, como carvão, piche, entre outros, reforçando que a palmeira macaúba apresenta grande potencial para produção de óleo com vasta aplicação nos setores industriais e energéticos, com vantagens sobre outras oleaginosas, principalmente com relação à sua maior rentabilidade agrícola e produção total de óleo. O maior desafio é transformar a atividade hoje extrativista em um cultivo racional (PANORAMAM BRASIL, 2006).

SILVA et al. (2006) enfatizam que os extensos palmares existentes em várias regiões brasileiras constituem um verdadeiro patrimônio, em estado potencial, à espera de exploração agrícola-industrial bem orientada. Imensas reservas naturais de palmáceas como no caso do dendê, do indaiá, do babaçu e da macaúba, apresentam-se muito promissoras pelo alto rendimento energético por unidade cultivada, e por se desenvolverem em áreas competitivas com a agricultura de sobrevivência. Os autores afirmam que trabalhos realizados demonstraram que a carbonização do endocarpo de algumas palmáceas pode fornecer um carvão de excelente qualidade quando comparado com o carvão mineral, notadamente no que se refere ao seu baixo teor de cinzas, ausência de enxofre, elevada densidade e controlável teor de carbono fixo e de matérias voláteis. Com base nos resultados por eles obtidos, concluiu-se que o carvão do endocarpo das palmáceas, como a macaúba, é superior ao carvão da madeira de eucalipto para usos, tais como: gasogênios, operações metalúrgicas e siderúrgicas e uso doméstico.

A maioria das palmeiras é rica em óleo. A macaúba é citada como uma das principais fontes de óleo vegetal, no Brasil, que tem potencial para produção de biodiesel. De acordo com estimativas, essa oleaginosa produz até quatro mil litros de óleo por hectare por ano. Convém lembrar que a planta começa a produzir após quatro ou cinco anos que foi plantada (Figura 11). A título de comparação deve ser citado que culturas anuais como a soja produz apenas 420 litros, o girassol 890 litros e mesmo a mamona, somente 1.320 litros (CESAR, 2003). O óleo da macaúba cada vez mais está sendo valorizado pelo mercado nacional e internacional, pois trata-se de um óleo de altíssima qualidade, tem valores nutricionais próximos ao azeite de oliva e seu potencial começou a ser descoberto pelas indústrias de cosméticos de alimentos. Após o processo de extração são obtidos dois tipos de óleo: o da polpa – que é esverdeado e aproveitado como biodiesel ou na indústria de cosméticos; e o óleo da amêndoa – amarelado, é mais nobre, podendo ser utilizado na indústria alimentícia e também em produtos de cosméticos (CIPRIANO, 2006). Soma-se a isso o fato da macaúba ser endêmica, espontânea e perene, com benefícios para o manejo e a conservação de solo, visto que uma vez plantada começa a produzir com quatro ou cinco anos e sua vida útil econômica excede 50 anos.

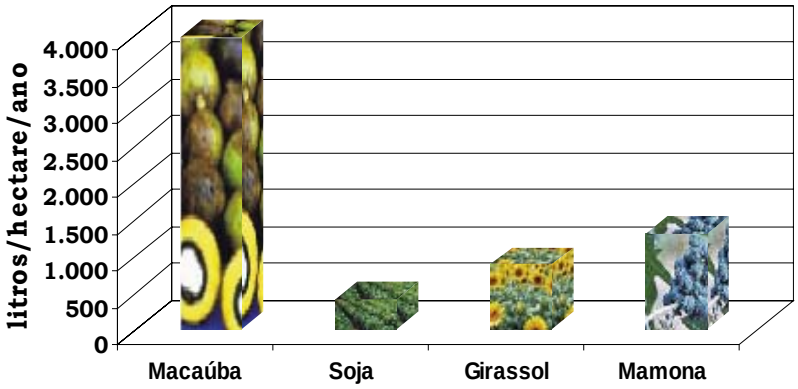


Figura 11 – Comparação entre capacidade de produção de óleo vegetal de algumas oleaginosas.

Pesquisadores de oleaginosas têm apresentado estudos que comprovam que a macaúba, planta até então pouco estudada e quase desconhecida, tem um futuro promissor na produção de biocombustível, na produção de óleos vegetais dietéticos e na oleoquímica. A macaúba atende a três premissas ainda pouco consideradas pela maioria dos envolvidos com a cadeia produtora dos biocombustíveis. A primeira delas é que diante da expectativa de uma demanda cada vez maior, provocada pelas crescentes dificuldades de exploração dos combustíveis fósseis e também por sua utilização como arma política, um programa para os biocombustíveis terá cada vez mais que levar em conta a utilização de plantas de alta produtividade de matéria-prima.

Outra consideração é que assim como as fontes energéticas, a questão da água também adquire importância cada vez maior, lembrando vários indicadores preocupantes de problemas climáticos e escassez de água. A terceira premissa mencionada é a necessidade de frear as correntes migratórias internas, garantindo emprego às populações em suas regiões de origem.

A macaúba oferece ainda outras vantagens sobre outras plantas que ocupam hoje posição de destaque no Brasil na produção de biodiesel, como a soja e o dendê. A macaúba, sendo uma palmeira rústica, necessita de pouca água. Com relação à mão-de-obra, a macaúba evitará o inchaço das cidades das regiões produtoras de cana-de-açúcar, outra matéria prima de combustível. Esses locais costumam receber um número elevado de trabalhadores na época da colheita. Como a macaúba exige trabalho manual para ser colhida, a cultura contribuiria para garantir a fixação das famílias à terra (FERNANDES, 2006). A utilização de seu óleo para a produção de biodiesel elevaria a demanda por essa planta, e ajudaria a promover o desenvolvimento sócio-econômico de algumas regiões problemáticas do estado de São Paulo, como o Vale do Ribeira e o Pontal do Paranapanema, onde o cultivo e a exploração da macaúba poderiam ser incentivados (BATISTA, 2006).

2.1.4 Programa Biodiesel

A capacidade de empregar óleos vegetais como óleo combustível em substituição ao óleo diesel, despertou o interesse do governo brasileiro que criou, através do Ministério de Ciência e Tecnologia – MCT, o Programa Brasileiro de Biodiesel – PROBIODIESEL, regulamentado em 30 de outubro de 2002 pela Portaria MCT nº 702. Este programa visa o desenvolvimento integrado, em rede, das tecnologias de produção, industrialização e uso do biodiesel e de misturas com diesel, a partir de óleos vegetais puros e residuais, produzidos regionalmente (CARNEIRO, 2003).

O biodiesel, combustível derivado de fontes renováveis e biodegradável, pode ser obtido de várias formas tais como craqueamento, esterificação ou transesterificação que é o processo mais utilizado atualmente. Consiste na produção do biodiesel por meio da reação química de óleos vegetais ou óleo animal com um álcool (metanol ou etanol), sempre na presença de uma substância química chamada catalisador, que é fundamental para promover a transformação química, da qual também se extrai a glicerina, produto com aplicações diversas na indústria química. Além da glicerina, a cadeia produtiva do biodiesel gera uma série de outros sub-produtos, como torta e farelo, que podem agregar valor e se constituir em outras fontes de renda importantes para os produtores. Este processo utiliza como matéria-prima, principalmente, óleos de várias espécies de plantas oleaginosas, tais como os extraídos da soja, da mamona, do girassol, do amendoim e de outras sementes, assim como aquele extraído de frutos e sementes de algumas palmeiras, tais como coqueiro, dendezeiro, babaçu e macaúba. O biodiesel substitui total ou parcialmente o óleo diesel de petróleo em motores ciclodiesel automotivos ou estacionários. Pode ser usado puro ou misturado ao diesel em diversas proporções. A mistura de 2% de biodiesel ao diesel de petróleo é chamada de B2 e assim sucessivamente, até o biodiesel puro, denominado B100 (BIODIESEL, 2006).

Um programa de geração de emprego e renda a partir da produção de biodiesel pode ser sustentado pelo grande mercado energético brasileiro. O Ministério da Agricultura, Pecuária e Abastecimento, Ministério do Desenvolvimento Agrário, Ministério da Integração Nacional, juntamente com o Ministério das Cidades desenvolveram estudos que mostram que a cada 1% de substituição de óleo diesel por biodiesel produzido com a participação da agricultura familiar podem ser gerados cerca de quarenta e cinco mil empregos no campo, com uma renda média anual de aproximadamente R$ 4.900,00 por emprego. Admitindo-se que para um emprego no campo são gerados três empregos na cidade, seriam criados, então, cento e oitenta mil empregos. Dados como estes evidenciam claramente a importância de priorizar a agricultura familiar na produção de biodiesel. A agricultura familiar associada à produção de plantas oleaginosas torna o biodiesel uma ferramenta relevante para auxiliar a erradicação da miséria no país. Mais de dois milhões de famílias, que vivem na região semi-árida no nordeste brasileiro, sofrendo com as péssimas condições sócio-econômicas podem ser beneficiadas através de projetos para inclusão social e o desenvolvimento regional, especialmente via geração de emprego e renda. Estes devem ser os princípios orientadores básicos das ações direcionadas ao biodiesel, o que implica dizer que sua produção e consumo devem ser promovidos de forma descentralizada e não-excludente em termos de rotas tecnológicas e matérias-primas utilizadas (BIODIESELBR, 2006).

2.1.5 Programa de agricultura familiar

No Estado do Pará, a Secretaria de Agricultura Familiar do Ministério do Desenvolvimento Agrário está apoiando a inclusão de mais de 100 produtores rurais para a produção de dendê, em parceria com a Companhia Refinadora da Amazônia com a finalidade de produção de biodiesel. Estes agricultores familiares irão plantar dendê em uma área de seiscentos hectares no município de Moju. Além da melhora na qualidade de vida, haverá um aumento significativo na renda dos mesmos. Estima-se que em um ano a renda total desses cem agricultores será de quase R$ 2 milhões (MDA, 2006).

No Estado do Rio Grande do Sul, a Petrobrás está investindo em parceiras com cooperativas de pequenos agricultores, para implantação de duas usinas de biodiesel. Esta parceria envolverá mais de oitenta mil produtores de cento e quarenta municípios, que poderão fornecer matéria-prima para as usinas, gerando aumento da renda familiar promovendo a inclusão social (PRESTES, 2006).

A utilização de biodiesel como combustível tem apresentado um potencial promissor no mundo inteiro. Vários países vêm investindo na produção e na viabilização comercial do biodiesel através de unidades de produção com diferentes capacidades, distribuídas particularmente na Europa (França, Áustria, Alemanha, Bélgica, Reino Unido, Itália, Holanda, Finlândia e Suécia), na América do Norte (Estados Unidos) e na Ásia (Japão), (COSTA NETO et al., 2000).

Conforme MEIRELLES (2005) e CARNEIRO (2006), o biodiesel é um combustível renovável e, portanto uma alternativa aos combustíveis tradicionais obtidos do petróleo, e sua utilização traz uma série de vantagens ambientais, econômicas e sociais:

- ambientais: redução de gases poluentes;
- econômicas: o país poderia enquadrar o biodiesel nos acordos estabelecidos no protocolo de Kyoto e nas diretrizes dos Mecanismos de Desenvolvimento Limpo (MDL);
- sociais: geração de postos de trabalho, especialmente no setor primário; baixo risco de explosão; aumento da oferta de espécies oleaginosas, que são um importante insumo para a indústria de alimentos e ração animal.

Dada a distribuição generalizada da macaúba no Brasil e ao seu potencial para produção de biodiesel, alguns estudos se fazem necessários. Antes de tudo, é de fundamental importância o conhecimento e a exploração da variabilidade genética presente na macaúba, em vista da escassez de estudos sobre a espécie. Em levantamento bibliográfico realizado em outubro de 2006, nas bases de dados *Web of Science* foram resgatados apenas doze artigos científicos realizados com a espécie, e a maior parte deles relatando a ocorrência de alguns insetos, sendo os demais referentes à sua extensa distribuição geográfica.

2.2 Estrutura Genética de Populações de Plantas

Uma população, do ponto de vista genético, é definida como um conjunto de indivíduos de uma mesma espécie, que compartilham o mesmo pool gênico, isto é, indivíduos que compartilham seus genes através de reprodução (FALCONER & MACKAY, 1989). Os parâmetros descritivos para avaliar a estrutura genética de uma população são as freqüências alélicas e genotípicas, que se referem, respectivamente, à quantidade observada de um determinado alelo frente ao total de alelos da população e a quantidade observada de indivíduos com um determinado genótipo frente ao total de indivíduos da população. As heterozigosidades esperada (H_e) e observada (H_o) e a diversidade genética são as medidas da variabilidade genética da população, que permitem identificar os alelos que permitem identificar o número de indivíduos heterozigotos, bem como a quantidade de indivíduos homozigotos diferentes (WEIR, 1996).

Para quantificar a variabilidade genética em populações de plantas, os parâmetros genéticos mais utilizados são o número de alelos por loco, a heterozigosidade esperada e a observada (PINTO, 2001). A alteração genética de uma população pode ser avaliada pela mudança nas suas freqüências gênicas, o que torna a estimativa da freqüência de heterozigotos fundamental nos estudos evolutivos (NEI, 1978). Um determinado alelo, em um grupo de indivíduos diplóides, tem como freqüência duas vezes o número de indivíduos homozigóticos para o alelo, mais o número de indivíduos no grupo. Em amostras grandes é maior a chance de se detectar alelos raros, pois o número de alelos observados por loco aumenta em função do tamanho da amostra. Conforme WEIR (1996), conhecer a freqüência dos heterozigotos é importante na medida em que cada heterozigoto carrega alelos distintos, demonstrando a existência de variação genética na população.

Forças evolutivas tais como: mutação, migração, seleção e deriva genética atuam dentro do contexto de cada espécie influenciando a estrutura genética da população, ou seja, interferindo na distribuição heterogênea (não aleatória) dos alelos e genótipos no espaço e tempo (HAMRICK, 1982). A migração ou fluxo gênico atua promovendo homogeneidade genética, enquanto que mutação, deriva genética e seleção atuam favorecendo adaptações às condições ambientais locais que podem levar à diferenciação genética de populações (p.ex. FUTUYMA, 1998).

O sistema pelo qual a espécie se reproduz influencia diretamente em suas freqüências alélica e genotípica. Uma população de tamanho infinito, que pratica cruzamentos ao acaso, está em equilíbrio de Hardy-Weinberg quando suas freqüências

alélicas e genotípicas não se alteram nas sucessivas gerações, e há total ausência de migração, mutação, seleção e deriva genética. Através deste princípio é possível o cálculo da freqüência dos genótipos, a partir das freqüências dos alelos (p. ex. FUTUYMA, 1993).

A principal ferramenta que alavancou o estudo da genética de populações, possibilitando a caracterização da estrutura genética entre e dentro de populações é o uso de marcadores moleculares.

2.3 Marcadores Moleculares

Segundo MATIOLI & PASSOS-BUENO (2001), a variabilidade genética é um instrumento de investigação muito importante para os pesquisadores em seus estudos quando desejam verificar afinidades e limites entre as espécies, detectar modos de reprodução e estrutura familiar, estimar níveis de migração e dispersão nas populações e até mesmo para ajudar na identificação de restos de espécies ameaçadas de extinção. Os dados básicos para esses estudos, denominados marcadores moleculares, são locos gênicos que apresentam alguma variabilidade.

Por meio do uso de marcadores moleculares, técnicas avançadas têm sido empregadas em pesquisas, permitindo que os pesquisadores alcancem maior eficiência em estudos de diversidade genética. As tecnologias de análise molecular da variabilidade do DNA permitem determinar pontos de referência nos cromossomos, tecnicamente denominados “marcadores moleculares”. Sendo assim, marcador molecular é todo e qualquer fenótipo molecular proveniente de um gene expresso, como no caso de isoenzimas, ou de um segmento específico de DNA (correspondendo a regiões expressas ou não do genoma). A revolução neste plano se iniciou com o descobrimento e utilização de marcadores isoenzimáticos a partir do ano de 1959, ampliando vastamente o número de marcadores genéticos e possibilitando a aplicação da técnica praticamente todas as espécies de plantas (FERREIRA & GRATTAPAGLIA, 1998).

Marcadores moleculares podem ser utilizados habitualmente em programas de melhoramento genético, tanto para caracteres quantitativos como qualitativos, seja em investigação básica ou aplicada. Diversos empregos de marcadores em melhoramento de plantas podem ser distribuídos em aplicações cujos resultados apresentam expectativas de curto, médio e longo prazo. No melhoramento clássico as aplicações em

curto prazo incluem, basicamente, a identificação e a discriminação de genótipos. Nas aplicações analíticas de médio e longo prazo, os marcadores permitem quantificar a variabilidade genética existente ao nível de seqüência de DNA e correlacioná-la com a expressão fenotípica em procedimentos de mapeamento genético. Nas aplicações sintéticas, a informação gerada na fase analítica, é integrada às metodologias de seleção e recombinação de genótipos, como uma ferramenta adicional para promover o avanço genético. Os marcadores moleculares tornaram-se ferramentas fundamentais na construção de mapas genéticos, permitindo a clonagem de genes que podem ser utilizados também no melhoramento via transgênica. Além de sua importância no melhoramento, os marcadores também são importantes nos estudos genéticos das plantas, detectando variações no genoma, aumentando o poder da análise genética. Uma grande variedade de marcadores moleculares está disponível para as diferentes espécies de vegetais (BORÉM & CAIXETA, 2006).

Os marcadores isoenzimáticos foram desenvolvidos pelos pesquisadores MARKET & MOLLER em 1959. São formas distintas de uma enzima, não precisam estar num mesmo indivíduo, com função catalítica idêntica ou similar que diferem entre si na seqüência de aminoácidos. O princípio básico da técnica consiste na utilização de eletroforese em gel e na visualização da atividade enzimática por métodos histoquímicos. Trata-se de uma técnica de fácil manipulação, ainda muito utilizada, custo relativamente baixo, permite analisar vários sistemas isoenzimáticos e grande número de indivíduos e permite visualizar marcadores co-dominantes. Tem como propriedade mais expressiva a base genética simples envolvida na expressão destas enzimas (SOLTIS & SOLTIS, 1989), entretanto apresenta como limitações uma cobertura reduzida dos genomas investigados, devido ao pequeno número de locos que podem ser detectados, e o baixo nível de polimorfismo identificado por loco (ALFENAS, 1998; BORÉM & CAIXETA, 2006).

A partir da década de 70, com a descoberta das enzimas de restrição na década anterior, surgiu o primeiro marcador molecular que possibilitou detectar diferenças entre indivíduos através de cortes no DNA estudado, o RFLP – *Restriction Fragment Lenght Polymorphism* (polimorfismo no tamanho de fragmentos de restrição). Com esta técnica, o DNA genômico de um indivíduo pode ser isolado e digerido com enzimas de restrição, os fragmentos obtidos são separados em um gel num processo denominado eletroforese, gerando fragmentos de tamanho e número variáveis que evidenciam o polimorfismo genético. Os fragmentos são transferidos para uma membrana de náilon

ou nitrocelulose – técnica denominada *Southern blot*. A próxima etapa consiste na hibridização do DNA das amostras já imobilizado em membranas com uma sonda radioativa de DNA complementar ao fragmento de interesse. A última etapa é a auto-radiografia, ou seja, exposição da membrana hibridizada com a sonda radioativa a um filme de Raios-X, revelando a presença de bandas, que são os marcadores RFLP (HELENTJARIS et al., 1986). Uma das principais vantagens desse marcador é a sua expressão co-dominante, que permite distinguir indivíduos homozigóticos de indivíduos heterozigóticos, outra vantagem que merece destaque é a possibilidade de utilização de sondas heterólogas, permitindo o mapeamento comparativo entre as espécies. As principais desvantagens são o trabalho que envolve a execução de várias etapas, o alto custo e o uso de radioatividade (FERREIRA & GRATTAPAGLIA, 1998).

Na década de 80, foi desenvolvido pelo pesquisador Kary Mullis uma tecnologia simples e eficiente que é a reação em cadeia da polimerase – PCR (Polymerase Chain Reaction). Através deste processo é possível multiplicar *in vitro* e em escala exponencial cópias de um segmento específico de DNA, na presença da enzima DNA polimerase. A reação de PCR se baseia na reassociação e extensão enzimática de um par de *primers* que delimitam a seqüência de DNA de fita dupla, os *primers* são sintetizados artificialmente e são complementares às regiões que flanqueiam a região alvo. Cada ciclo de PCR envolve três etapas, desnaturação através da elevação da temperatura de 92°C para 95°C, reassociação, em que a temperatura é rapidamente reduzida para 35°C a 60°C, e extensão que ocorre quando a temperatura é elevada para 72°C. Este ciclo é repetido por várias vezes e, ao final da reação, há uma grande quantidade de DNA de uma seqüência específica de interesse. Os produtos da PCR podem ser visualizados num gel de agarose. Esta visualização é possível com auxílio do brometo de etídeo, que quando presente no gel se intercala entre as bases nitrogenadas no DNA e o torna visível com absorção de um comprimento de onda menor e emissão de um comprimento de onda maior de luz ultravioleta (MULLIS & FALOONA, 1987). O desenvolvimento desta técnica levou Kary Mullis a ganhar o prêmio Nobel de química em 1993, devido ao impacto causado pela PCR e dos métodos dela derivados (NOBELPRIZE, 2006). A técnica de PCR promoveu também o surgimento de novos marcadores moleculares tais como os marcadores RAPD, STS, SCARS, AFLP, SNPs e Microssatélites. Estes são os mais importantes e serão discutidos a seguir.

O RAPD – *Random Amplified Polymorphic DNA* (DNA polimórfico amplificado arbitrariamente) – é basicamente uma variação do protocolo de PCR, difere

em duas características, utiliza um *primer* único e menor ao invés de um par de *primers* e este tem seqüência alvo desconhecida. Esta técnica permite a detecção de fragmentos aleatórios de DNA. Em relação aos RFLPs, os RAPDs são mais baratos, requerem pouco tempo e não necessitam de radioisótopos, entretanto são marcadores dominantes. Outra desvantagem deste marcador é que trata-se de uma técnica muito sensível que não oferece reprodutibilidade entre laboratórios (WILLIAMS, 1990).

PARAMN & MICHELMORE (1993) foram os primeiros pesquisadores a utilizar os marcadores STS – *Sequence Tagged Sites* (sítios marcados por seqüências), em plantas, eles também definiram estes marcadores como DNAs de seqüências curtas e de cópia única, que podem ser identificados mediante amplificação via PCR e, ou hibridização. BORÉM & CAIXETA (2006) citam os STS como uma conversão de RFLP em marcadores baseados em PCR, que oferecem grande vantagem em relação aos marcadores RAPD por apresentarem co-dominância, ou seja, podem distinguir homozigoto de heterozigoto e são também mais reproduzíveis que os RAPDs. Entretanto, as STS requerem algum conhecimento prévio da seqüência do DNA-alvo.

Os marcadores SCAR – *Sequence Characterized Amplified Regions* (regiões amplificadas caracterizadas por seqüências), a exemplo de STS são outra categoria de marcador molecular desenvolvida por conversão de um determinado marcador em outro. Conforme PARAMN & MICHELMORE (1993), os SCARs podem conter seqüências de DNA repetitivo e são usados como pontos de referência física no genoma e como marcadores genéticos. Os pesquisadores propuseram originalmente que SCAR é a conversão de marcadores RAPD, que são obtidos com *primers* de seqüência arbitrária, para marcadores de seqüência específica. XU et al. (2004) também converteram os marcadores AFLP em marcadores do tipo SCAR. Para desenvolver marcadores SCAR é necessária a identificação de um *primer* que confere polimorfismo a dois *bulks* de DNA com fenótipos contrastantes, em seguida o isolamento e a clonagem do fragmento amplificado em um vetor, sequenciamento do fragmento isolado, desenho dos pares de *primers* específicos de 16 a 24 nucleotídeos complementares às extremidades do fragmento seqüenciado, e por fim validar estes *primers* em reações de PCR, para verificar se eles amplificam DNAs genômicos da espécie para qual eles foram desenhados (BÓREM & CAIXETA, 2006).

AFLP – *Amplified Fragment Length Polymorphisms* (polimorfismos de comprimento de fragmentos amplificados), são marcadores moleculares desenvolvidos por ZABEAU (1993) citado por VANTOAL et al. (1996), cuja técnica combina a

digestão de fragmentos de DNA com enzimas de restrição de corte raro e enzimas de corte freqüente e a amplificação desses fragmentos por PCR. Assim sendo, o DNA de um indivíduo é clivado com duas enzimas de restrição, são ligados às extremidades dos fragmentos digeridos adaptadores específicos de DNA fita dupla, que servem de sítios de reassociação dos *primers* para posterior reação de PCR, os produtos dessa amplificação são separados em gel de poliacrilamida desnaturante e o polimorfismo é identificado pela presença ou ausência de bandas (VOS et al., 1995). Assim como os marcadores RAPD, estes marcadores apresentam dominância e são mais caros que RAPD. Apresenta como vantagens a visualização em um único gel de um grande número de fragmentos gerados, proporcionando uma amostragem ampla e simultânea de um genoma, grande poder de detecção de variabilidade genética, apresenta maior robustez quando comparado com RAPD.

Alterações de uma única base no genoma, ocorrendo a cada 600 pares de bases aproximadamente, ao se comparar segmentos correspondentes de um genoma, são denominados SNPs – *Single Nucleotide Polymorphisms* (polimorfismo de um único nucleotídeo), quando correspondem a posições onde existe uma alternância dos nucleotídeos A, C, G, T em uma freqüência alélica mínima de 1% em uma dada população (BROOKES, 1999). Os SNPs são, portanto, marcadores genéticos baseados na detecção de polimorfismos resultantes destas alterações, que devido à freqüência em que ocorrem tornam estes marcadores importantes ferramentas para obtenção de mapas de alta resolução. A detecção de SNPs emprega algumas metodologias; inicialmente a detecção consistia em amplificar e seqüenciar fragmentos genômicos equivalentes, de regiões gênicas específicas do DNA de vários indivíduos, e comparar suas seqüências em busca de variações. A via direta para identificação de SNP é o sequenciamento direto de fragmentos de DNA, amplificados por PCR, com *primers* desenhados para este fim. A identificação também pode ser feita por análise eletrônica de variação pontual (KWOK, 2001).

Microssatélite foi o marcador molecular utilizado como ferramenta neste estudo, então para melhor entendimento da metodologia adotada e das formas de análises, será abordado de forma mais detalhada a seguir.

2.3.1 Marcadores microssatélites

Segundo HAMADA et al. (1982) e TAUTZ & RENZ (1984), os genomas dos eucariotos estão densamente povoados por diferentes classes, mais ou menos complexas, de seqüências repetidas, origem do termo SSR – *Simple Sequence Repeats*, ou STR – *Short Tandem Repeats*, mais tarde chamados de marcadores Microssatélites. São repetições de pequenas seqüências de um a seis nucleotídeos adjacentes repetidos em linha, sendo encontrados amplamente distribuídos pelo genoma da maior parte dos eucariotos (FERREIRA & GRATTAPAGLIA, 1998; TOTH et al., 2000, ZANE et al, 2002). Foram descritos em plantas pela primeira vez por CONDIT & HUBBELL (1991), que detectaram a presença das repetições $(AC)_n$ e $(AG)_n$. MORGANTE & OLIVIERI (1993) constataram em 34 espécies vegetais a presença de microssatélites, sendo o dinucleotídeo AT o elemento repetido mais comum, também constataram que estes marcadores estão em menor freqüência que nos vertebrados, mas em maior freqüência que nos invertebrados e fungos. Estudos apontam para o *slippage*, ou o deslizamento da polimerase, como o principal mecanismo por trás do surgimento e amplificação destas seqüências repetitivas nos genomas, gerando alta taxa de mutação dos microssatélites, numa freqüência de 10^{-3} e 10^{-4} por loco por gameta em cada geração. Acredita-se que durante a replicação de uma região repetitiva, as fitas de DNA separam-se e unem-se novamente de forma incorreta, o que geraria cópias de trechos de DNA (alelos) com diferentes tamanhos ou números de repetições de um determinado motivo no próximo ciclo de replicação, por meio da inserção ou deleção de uma unidade de repetição. O problema também pode estar associado ao sistema de reparo pela DNA polimerase ou ainda, ser conseqüência do processo de recombinação (SCHLOTTERER & TAUTZ, 1992; FIELD AND WILLS, 1996). Durante o processo de recombinação o *crossing-over* desigual pode ser responsável pela alta taxa de polimorfismo destes marcadores, que por problemas no pareamento dessas seqüências durante o quiasma, aumenta a taxa de mutação das regiões microssatélites e são estas mutações que tornam estes marcadores tão informativos (STRAND et al., 1993; SCHLÖTTERER et al., 1998). O polimorfismo de um loco microssatélite é resultado do número de repetições do motivo que o caracteriza. Portanto, os diferentes alelos que um loco microssatélite pode apresentar distinguem-se uns dos outros pelo número de repetições que apresentam. Adotemos como exemplo, um loco em que o dinucleotídeo "CG" é a unidade de repetição, o polimorfismo que distinguirá os alelos deste loco dependerá da presença de números variáveis dessa unidade de repetição, então um alelo pode ser definido, por exemplo, pela ocorrência de 20 unidades de "CG" adjacentes

repetidos lado a lado, enquanto que outro alelo deste loco pode apresentar apenas 18 repetições.

CHAMBERS & MACAVOY (2000) sugerem que são necessários, para iniciar um processo de origem de um microssatélite, no mínimo oito nucleotídeos repetidos em *tandem*.

Os microssatélites são classificados conforme a composição das seqüências repetidas: (a) repetições perfeitas, quando não apresentam nenhuma interrupção, por exemplo: GTGTGTGTGTGTGT; (b) repetições imperfeitas, quando são interrompidas por bases que não correspondem ao motivo, por exemplo: GTGTGTGTaGTGTGTT; (c) repetições compostas, quando duas ou mais repetições (classes) de microssatélites estão dispostas adjacentes, por exemplo: GTGTGTGTGTGTCACACACACACA (BORÉM & CAIXETA, 2006).

Em geral, apesar das seqüências microssatélites variarem de um indivíduo para o outro, as seqüências de DNA que as flanqueiam são bastante conservadas entre indivíduos da mesma espécie. Conhecendo as regiões é possível desenhar *primers* específicos para essas regiões adjacentes às seqüências microssatélites, possibilitando a amplificação via PCR dos fragmentos que contêm o DNA repetitivo em diferentes genótipos. Os produtos da amplificação podem ser visualizados em gel de poliacrilamida desnaturante ou não desnaturante, em gel de agarose de alta resolução ou por meio de *primers* fluorescentes em seqüenciador automático, e o polimorfismo entre as bandas será decorrente dos tamanhos diferentes de elementos simples repetidos. Cada segmento de DNA representa um alelo daquele loco específico (SOUZA, 2001; ALVES, 2002; BORÉM & CAIXETA, 2006). O alto poder discriminatório deste marcador é uma característica importante que merece destaque, motivo pelo qual é amplamente utilizado em estudos de genética de populações, análises de paternidade e forense (OLIVEIRA et al., 2006).

Segundo SLATKIN (1995) são marcadores ideais para estudos de genética e evolução de populações naturais devido ao alto poder discriminatório, à robustez, confiabilidade, praticidade operacional e por serem marcadores mais informativos geneticamente. Os microssatélites apresentam características altamente desejáveis (SUGANUMA & CIAMPI, 2005):

- são marcadores co-dominantes, ou seja, ambos os alelos de um indivíduo heterozigoto podem ser visualizados no gel;

- estão ampla e uniformemente distribuídos pelo genoma dos eucariotos;
- são altamente multialélicos;
- são amplificados via PCR, o que facilita sua obtenção mesmo com poucas quantidades de DNA;
- uma vez desenvolvidos os *primers* que amplificam tais regiões do genoma, estes podem ser facilmente compartilhados entre laboratórios;
- *primers* marcados com fluorescência apresentam a vantagem de possibilitarem sistema multiplex, o que permite avaliar rapidamente um grande número de indivíduos para um grande número de locos em pouco tempo.

O emprego da tecnologia dos microssatélites envolve algumas limitações, como a grande quantidade de trabalho necessário para o desenvolvimento de *primers* específicos para os locos de microssatélites de cada espécie. Entretanto, ocorre conservação de sítios microssatélites entre espécies relacionadas, possibilitando aproveitar *primers*, denominados heterólogos, que foram desenvolvidos para uma espécie e empregá-los nas demais espécies do gênero (FERREIRA & GRATTAPAGLIA, 1998).

2.3.2 Análise da estrutura genética com marcadores co-dominantes

Através do uso de marcadores co-dominantes a caracterização da estrutura genética entre populações, isto é, o grau de diferenciação populacional pode ser abordado de três modos diferentes, sendo eles: estatísticas *F* de Wright; análise da diversidade gênica em populações subdivididas; e os coeficientes de co-ancestralidade de Cockerham. As estatísticas F de Wright possibilitam a caracterização da distribuição da variabilidade genética entre as populações (F_{ST}), bem como dos níveis médios de endogamia ao nível populacional (F_{IS}) e total (F_{IT}), (WRIGHT, 1965). A avaliação da divergência em diferentes níveis de hierarquia e a obtenção de estimativas de endogamia, a partir de uma base não enviesada, é obtida através dos coeficientes de co-ancestralidade de Cokerham (θ), (COCKERHAM, 1969; VENCOVSKY, 1992; WEIR, 1996). E através da análise da divergência gênica em populações subdivididas é possível a comparação dos níveis de heterozigosidade entre e dentro das populações, bem como a obtenção de uma estimativa de divergência, a partir de uma base diferente daquela que fundamenta as estimativas de F_{ST} e θ (NEI, 1977).

São três as formas de estimar as estatísticas F:

- De acordo com WRIGHT (1965):

$$1-\hat{F}_{IT} = (1-\hat{F}_{IS})(1-\hat{F}_{ST})$$

$\hat{F}_{IT}$ =índice de fixação ou coeficiente de endogamia para o conjunto das populações (devido ao sistema reprodutivo e subdivisão);

$\hat{F}_{IS}$ =índice de fixação ou coeficiente de endogamia intrapopulacional (devido ao sistema reprodutivo);

$\hat{F}_{ST}$ = índice de fixação ou coeficiente de endogamia entre populações (devido à subdivisão);

- Coeficientes de co-ancestralidade de Cockerham (COCKERHAN, 1969):

F de Cockerham ≈ F_{IT} de Wright: é a correlação entre alelos dentro de indivíduos em todas as populações. É o coeficiente de endogamia que se refere aos indivíduos em relação ao conjunto de populações, reunindo a informação dos índices de fixação F_{ST} e F_{IS}.

θ de Cockerham ≈ F_{ST} de Wright: é a correlação dos alelos dentro de indivíduos na mesma população. É o coeficiente de co-ancestralidade.

f de Cockerham ≈ F_{IS} de Wright: correlação dos alelos dentro de indivíduos dentro das populações.

- De acordo com NEI (1977):

$$\hat{F}_{IT} = 1-\frac{\hat{H}_{ot}}{\hat{H}_{et}}; \qquad \hat{F}_{IS} = 1-\frac{\hat{H}_{ot}}{\hat{H}_{ei}} \qquad \hat{F}_{ST} = 1-\frac{\hat{H}_{ei}}{\hat{H}_{et}}$$

em que:

$$\hat{H}_{ot} = \frac{\sum\sum \hat{x}_{il}}{ls} \qquad \hat{H}_{ei} = 1-\frac{\sum\sum \hat{x}_{ilk}^2}{ls}$$

$$\hat{H}_{et} = 1-\frac{\sum\left(\frac{\sum \hat{x}_{ilk}}{s}\right)^2}{l}$$

sendo:

$\hat{H}_{et}$ = heterozigosidade esperada total;

$\hat{H}_{ot}$ = heterozigosidade observada total;

$\hat{H}_{ei}$ = heterozigosidade esperada média total;

$\hat{x}_{il}$ = freqüência de heterozigotos do loco l na população i;

$\hat{x}_{ilk}$ = freqüência do alelo k do loco l na população i;

s =número de populações;

l = número de locos.

Recentemente, uma estatística análoga ao θ_p de Cockerham tem sido utilizada em estudos com marcadores microssatélites. Trata-se da estatística R_{ST} que comporta em seu cálculo o *stepwise model* proposto por SLATKIN (1995) em que a mutação para um novo alelo de um determinado marcador microssatélite ocorre em passos, assim alelos com tamanhos mais próximos são considerados mais similares do que os alelos com tamanhos mais distantes. A estatística R_{ST} é calculada da seguinte forma:

$$R_{ST} = \frac{\overline{S} - S_w}{\overline{S}}$$

sendo:

S_w = duas vezes a média da variância estimada do tamanho do alelo encontrado dentro de cada população;

$\overline{S}$ = duas vezes a variância estimada do tamanho do alelo encontrado em todas populações.

Locos que apresentam freqüências semelhantes em diferentes populações apresentam baixos valores de R_{ST}, enquanto que aqueles que apresentam freqüências gênicas muito distintas entre populações apresentam valores altos.

O uso de marcadores co-dominates possibilita a obtenção de informações que fornecem subsídios para entender a dinâmica evolutiva das espécies auxiliando na adoção de estratégias de conservação e melhoramento genético, manejo de populações naturais, podendo determinar o impacto da influência antrópica na variabilidade das espécies (HAMRICK, 1989; SEOANE et al., 2001).

3 MATERIAL E MÉTODOS

3.1Material vegetal

3.1.1 Desenvolvimento de bibliotecas enriquecidas com locos microssatélites

Um genótipo de *Acrocomia aculeata* foi escolhido ao acaso, entre os que foram coletados, para obtenção do DNA genômico total a fim de construir uma biblioteca enriquecida com locos microssatélites para a espécie *Acrocomia aculeata*.

3.1.2 Caracterização da estrutura genética populacional da espécie *Acrocomia aculeata*

Visando a obtenção de informações da estrutura genética populacional utilizando locos microssatélites em macaúba, foram utilizados 47 genótipos de onze localidades do Estado de São Paulo e duas do Estado de Minas Gerais, em função da ocorrência da espécie nesses locais, representando quatro grandes regiões (Figura 12) e totalizando nove populações (Tabela 2). Trata-se de indivíduos de ocorrência espontânea nas localidades em que foram coletados.

Tabela 2 – Regiões e localidades de coleta, indivíduos de macaúba escolhidos para este estudo.

Região	População	Localidade de coleta	N I
I	1	Campinas	4
		Jaguariúna	3
	2	Amparo	1
		Serra Negra	2
		Lindóia	2
II	3	Mococa	5
	4	Casa Branca	5
III	5	Dourado	5
	6	Vicentinópolis	2
		Barbosa	4
	7	SPTurvo	4
IV	8	Ibituruna/Gupiara	6
	9	Ibituruna/Lagoa	4

NI = número de individuos.

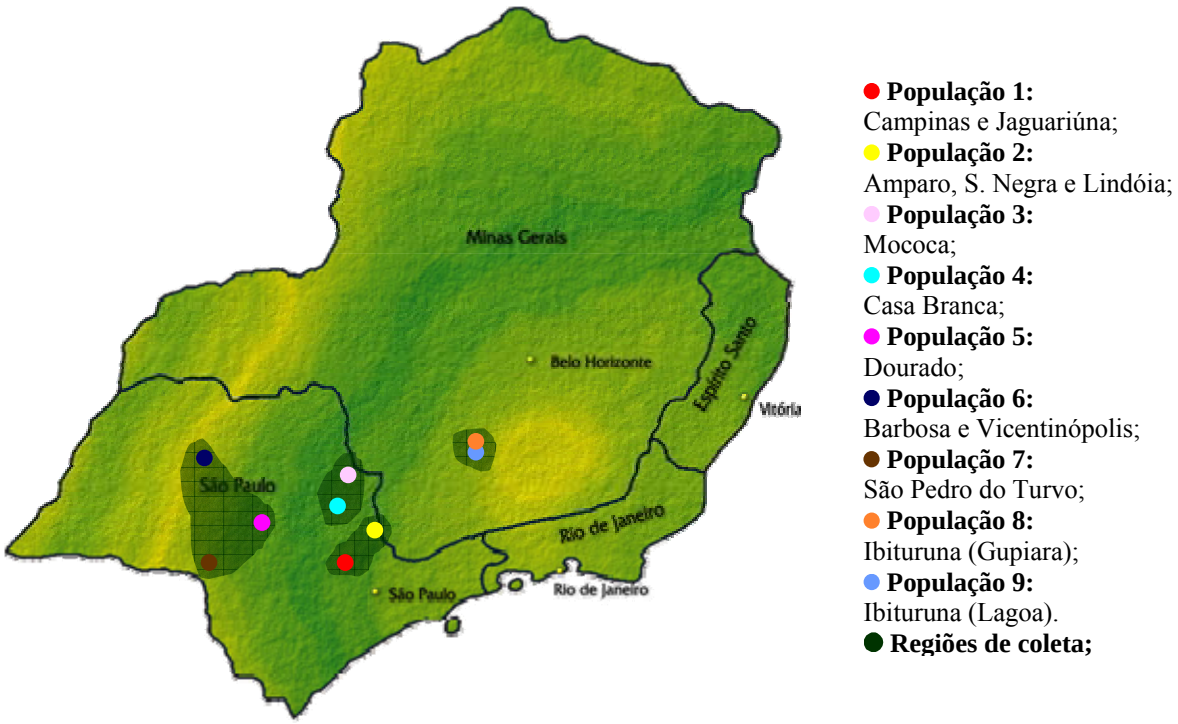


Figura 12 – Localidades de coleta dos genótipos de macaúba utilizados neste estudo.

Os materiais vegetais coletados dos diferentes genótipos foram armazenados separadamente em freezer a –20ºC e posteriormente macerados com nitrogênio liquido, em almofariz, e armazenados para posterior extração de DNA.

3.2 Extração e Quantificação de DNA

Tanto para construção da biblioteca enriquecida com locos microssatélites, como para a caracterização e estudo da estrutura genética populacional da macaúba, o procedimento para extração de DNA adotado foi o mesmo e está detalhado a seguir (Figura 13). O DNA foi extraído a partir de 500 mg de tecido vegetal macerado utilizando o protocolo CTAB descrito por DOYLE & DOYLE (1990) com modificações: (1) Em eppendorffs de 1,5 mL identificados para cada amostra foram adicionados 30 mg do tecido macerado; (2) foram adicionados 800 µL de tampão de extração (pré-aquecido a 65ºC), os tubos foram fechados e agitados para ressuspender o tecido no tampão; (3) foram levados ao banho-maria (65ºC) por 60 minutos, agitando-os manualmente a cada 10 minutos; (5) após serem retirados do banho-maria e com a mistura em temperatura ambiente foram adicionados 700 µL de clorofórmio-álcool

isoamílico (24:1) e os tubos foram agitados por 5 minutos invertendo-os no mínimo 20 vezes até fazer uma emulsão homogênea; (6) os tubos foram centrifugados por 5 minutos a 12000 rpm; (7) o sobrenadante foi transferido para novos tubos também marcados; (8) foram adicionados 70% (v/v) do volume (≈ 500 µl) de isopropanol gelado, misturado suavemente até formar um precipitado; (9) os tubos foram centrifugados durante 10 minutos a 12000 rpm; (10) suavemente foi descartado o máximo de sobrenadante invertendo os tubos sem perder o pellet; o precipitado secou a temperatura ambiente; (11) foram adicionados 500 µL de etanol 70% (v/v) para lavar o precipitado, deixando-o imerso por 5 a 10 minutos, invertendo-o suavemente; (12) os tubos foram centrifugados durante 5 minutos a 12000 rpm; (13) foi retirado o máximo do etanol sem perder o pellet; (14) foi adicionado 500 µL de etanol 95% (v/v) para lavar o precipitado, invertendo-o suavemente; (15) os tubos foram centrifugados durante 5 minutos a 12000 rpm; (16) foi retirado o máximo do etanol e o pellet secou em temperatura ambiente; (17) cada precipitado foi dissolvido em 100 µL de tampão TE (10 mM Tris-HCl, 1 mM de EDTA, pH 8,0) acrescido de 1 µL RNAse (10mg/mL) e deixado sobre bancada *overnight*. As amostras foram então armazenadas a –20ºC.

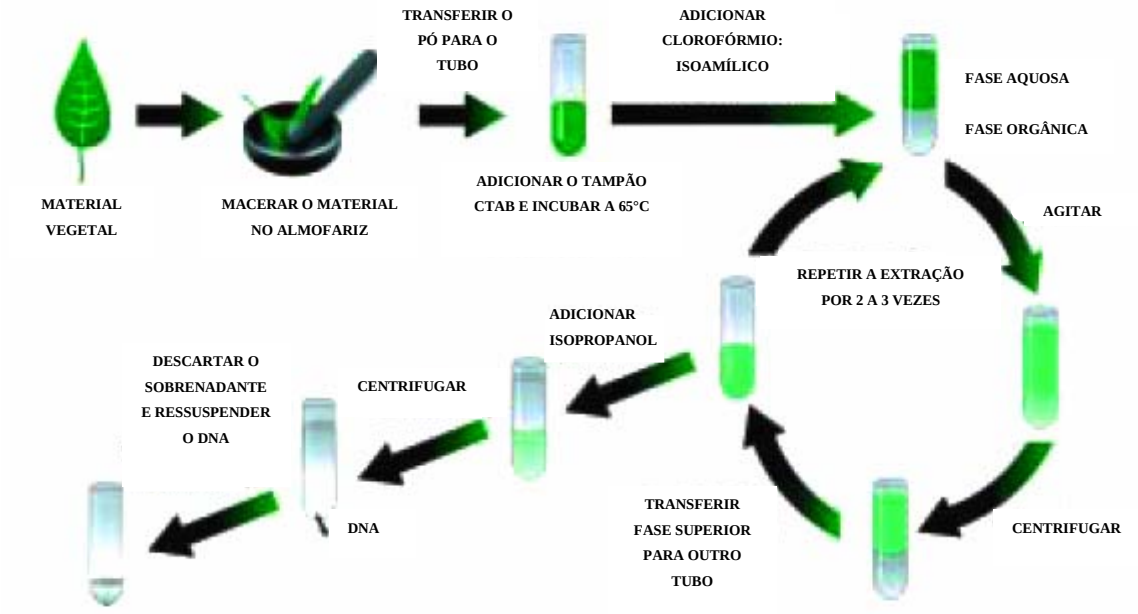


Figura 13 – Esquema representativo das etapas de extração de DNA pelo método CTAB (Fonte: ROMANO & BRASILEIRO, 2005).

Após eletroforese a quantificação das amostras foi estimada pela intensidade de fluorescência emitida pelo brometo de etídeo sob UV em géis de agarose a 0,8% (p/v). Essa intensidade foi comparada à de padrões com pesos moleculares e concentrações específicos e conhecidos (DNA do fago λ).

3.3 Desenvolvimento e Caracterização de Marcadores Microssatélites

3.3.1 Construção de biblioteca genômica enriquecida com microssatélites

Foi testada a transferibilidade de 18 *primers* desenvolvidos para *Euterpe edulis* em macaúba, conforme GAIOTTO et al., (2001), antes de iniciarmos o desenvolvimento da biblioteca genômica enriquecida com locos microssatélites para macaúba.

A construção da biblioteca genômica enriquecida com locos microssatélites foi feita utilizando o protocolo descrito por BILLOTTE et al. (1999) com modificações (Figura 14), conforme os passos a seguir:

Restrição do DNA

Aproximadamente 6,0 μg de DNA genômico foram digeridos utilizando 5,0 μL enzima *Rsa*I (10 U/μL), adicionando-se na reação 10,0 μL de tampão de reação (50 mM Tris-HCl, pH 7,8; 10 mM $MgCl_2$; 10 mM DTT; 25 μg/ml BSA) e água milliQ autoclavada até completar 100 μL. O material foi incubado a 37^{o}C por 18 horas.

Ligação dos adaptadores

Os adaptadores *Rsa*21 (5′ - CTC TTG CTT ACG CGT GGA CTA - 3′) e *Rsa*25 (5′ - TAG TCC ACG CGT AAG CAA GAG CAC A - 3′) foram ligados aos fragmentos digeridos de extremidade abrupta com a enzima T4 DNA ligase (Amershan Pharmacia Biotech). Foram utilizados 2 μL do adaptador *Rsa*21 (10 μM) e 2 μL do adaptador *Rsa*25 (10 μM), 6,0 μL do DNA genômico digerido no passo anterior, 10,0 μL do tampão 5X (50 mM Tris-HCl, pH 7,8; 10 mM MgCl2; 10 mM DTT; 25 μg/ml BSA), 4,0 μL de T4 DNA ligase e água milliQ autoclavada para completar 50,0 μL. Esta reação foi incubada a 20^{o}C por 2 horas.

Pré-Amplificação via PCR

Para gerar uma maior quantidade de DNA para a seleção, os fragmentos digeridos foram submetidos a uma PCR utilizando como *primer* um dos adaptadores, o *Rsa*21. Após a ligação os fragmentos foram amplificados utilizando 5 μL do produto da ligação; 2,0 μL do *primer Rsa*21 (10 μM); 5,0 μL de tampão 10 X (50 mM KCl; 10 mM de Tris-HCl, pH 8,9); 2,0 μL de $MgCl_2$ (50,0 mM); 4μL de dNTP (2,5 mM), 1,0 μL de *Taq* polimerase (5 U) e água milliQ autoclavada para completar 50,0 μL. Esta reação foi submetida a uma etapa inicial de desnaturação a 95°C por 4 minutos, seguida de 20 ciclos de 94°C por 30 segundos, 60°C por 1 minuto e 72°C por 1 minuto. Um passo prolongado de extensão de 72°C por 8 minutos foi adicionado após o último ciclo. Os fragmentos amplificados foram purificados utilizando o Kit Concert (Invitrogen Life Sciences).

Seleção dos fragmentos contendo SSR

O processo de enriquecimento foi realizado através da hibridização de oligonucleotídeos conjugados a biotinas, biotIIIII$(TTC)_{10}$, biotIIIII$(CG)_{10}$ e biotIIIII$(GT)_{10}$, complementares a seqüências repetitivas GA, GC e AAG. Os fragmentos contendo tais repetições foram recuperados por estreptavidina ligada a microesferas magnéticas utilizando o kit Streptavidine-Magnesphere (Promega). Para realização das lavagens, as esferas magnetizadas foram atraídas por um imã posicionado lateralmente ao tubo em um suporte.

A fração enriquecida contendo os fragmentos de DNA previamente ligados aos adaptadores foi incubada a 95°C por 15 minutos para a desnaturação da dupla fita. Ao DNA desnaturado foram adicionados 3 μL de cada oligonucleotídeo marcado com as biotinas. Esta solução de hibridização foi incubada a temperatura ambiente sob constante agitação e após 20 minutos foi adicionada às esferas magnetizadas previamente lavadas seguindo as recomendações do fabricante. A suspensão contendo as esferas magnetizadas e o complexo DNA-sonda foi incubada por 10 minutos à temperatura ambiente sob suave agitação.

Após a incubação, foram realizadas lavagens conforme descrito por BILLOTTE et al. (1999) e os fragmentos selecionados foram recuperados através de lavagem com 250 μL de água milliQ autoclavada.

Amplificação dos fragmentos selecionados via PCR

Após a seleção dos fragmentos enriquecidos, estes foram submetidos a uma PCR utilizando como *primer* o adaptador *Rsa*21. A amplificação foi realizada utilizando 20 μL do produto da ligação; 4,0 μL do primer *Rsa*21 (20 μM); 10 μL de tampão 10 X (50 mM KCl; 10 mM de Tris-HCl pH 8,9); 4,0 μL de $MgCl_2$ (2,0 mM); 8 μL de dNTP (2,5 mM), 5U de *Taq* polimerase (Invitrogen) e água milliQ autoclavada para completar 100,0 μL. Esta reação foi submetida a uma etapa inicial de desnaturação a 95°C por 1 minuto, seguida de 25 ciclos de 94°C por 40 segundos, 60°C por 1 minuto e 72°C por 2 minutos. Um passo prolongado de extensão de 72°C por 5 minutos foi adicionado após o último ciclo.

Clonagem em vetor pGEM-T e transformação em células supercompetentes

Após a amplificação, 3 μL dos produtos da PCR foram ligados a 1 μL do vetor pGEM-T (kit Promega), segundo o protocolo fornecido com o vetor plasmidial, e foram transformados em células XL1-Blue competentes. As células transformadas foram plaqueadas em meio LB sólido contendo ampicilina (100 μg/ml), 60 μL IPTG (24 mg/ml) e 60 μL X-Gal (20 mg/ml). As placas foram incubadas invertidas a 37 °C por 18 horas em estufa para o crescimento de colônias.

Manutenção dos clones

Para garantir que cada construção (vetor + fragmento) fosse mantida em condições apropriadas para análise posterior, foram utilizadas placas ELISA com fundo em U, com 80 μL de meio LB contendo ampicilina (100μg/ml) por poço. As colônias brancas foram repicadas com a ajuda de palitos estéreis. As placas foram incubadas a 37 °C *overnight* em estufa para o crescimento de colônias. Após esse período as placas foram armazenadas em freezer -20°C por 30 minutos e então foram armazenadas em freezer -80°C com glicerol 50% (p/v).

3.3.2 Seleção e sequenciamento dos clones positivos

Amplificação dos insertos clonados

Com o objetivo de identificar os clones contendo microssatélites e verificar a eficiência do procedimento de enriquecimento e clonagem, foi realizada uma reação de amplificação dos insertos diretamente das colônias obtidas na etapa anterior, utilizando

o *primer Rsa*21. Colônias individuais foram transferidas com um palito estéril para o tubo de PCR contendo 30,5 µL de água milliQ estéril. Esta suspensão de células foi utilizada em uma reação com volume final de 45,0 µL contendo as seguintes concentrações dos reagentes: 5,0 µL de tampão 10X (50 mM KCl; 10mM de Tris-HCl, pH 8,9), 4,0 µL de $MgCl_2$ (2,5 mM), 4,0 µL de dNTP (2,5 mM), 2,5 µL de *Rsa*21 (10 mM), 1,0 µL de *Taq* DNA polimerase (5 U/µL). Esta reação foi submetida a um passo inicial de desnaturação a 95ºC por 4 minutos, seguido de 30 ciclos (94ºC por 30segundos, 52ºC por 45segundos, 72ºC por 1 minuto e 30 segundos) e um passo final de extensão a 72ºC por 8 minutos. Para controle, 10 µL do volume da reação foi utilizado na eletroforese em gel de agarose 1,5% (p/v).

Inoculação e extração plasmidial

Com o objetivo de isolar o DNA plasmidial das colônias recombinantes para posterior sequenciamento foi colocado 1 ml de meio Circle Grow contendo 100 µg/ml de ampicilina em cada pocinho da microplaca; foram inoculadas colônias individuais com o auxílio de pipeta multicanal. A placa foi selada com adesivo, foram feitos furos em cima com uma agulha para aeração durante o crescimento, foi incubada a 37ºC (shaker) a 300 rpm durante 22h; o adesivo da placa foi trocado e placa foi centrifugada por 6 minutos a 3000 rpm, para sedimentar as células; foi descartado o sobrenadante e a placa foi mantida invertida sobre papel absorvente por 1 minutos; foi adicionado a cada pocinho 240 µL de solução GTE [Glicose 20%, Tris 1 M (pH 7,4), EDTA 0,5 M (pH 8,0)], a placa foi selada com adesivo e houve a ressuspensão das células agitando-as no vortex por 2 minutos; a placa foi centrifugada por 6 minutos a 4000 rpm e o sobrenadante foi descartado; foi adicionado a cada pocinho 80 µL de GTE, a placa foi selada com adesivo para etanol e agitada no vortex por 5 minutos; foi transferido 60 µL de cada suspensão de células para placa com fundo em U contendo 5 µL de RNAse (10 mg/mL); foi adicionado a cada pocinho 60 µL de NaOH 0,2M – SDS 1% (p/v), a placa foi selada com adesivo e a solução foi misturada 10 vezes por inversão e incubada 10 minutos à temperatura ambiente; após esse período a placa foi centrifugada brevemente; foi adicionado a cada pocinho 60 µL de KOAc 3 M, a placa foi selada com adesivo e a solução foi misturada 10 vezes por inversão, centrifugada brevemente; o adesivo foi removido e a placa incubada em estufa a 90ºC por exatos 30 minutos; a placa foi esfriada em gelo por 10 minutos e centrifugada por 4 minutos a 4000 rpm; foi fixada com fita adesiva uma placa filtro no topo de uma placa de fundo em U, foi transferido

todo o volume para a placa filtro, e centrifugada por 4 minutos a 4000 rpm, a 20ºC; foi removida a placa filtro e adicionado ao filtrado 110 µL de isopropanol; a placa foi selada com adesivo e a solução foi misturada 20 vezes por inversão e centrifugada por 45 minutos a 4000 rpm a 20ºC; o sobrenadante foi cuidadosamente descartado e foram adicionados 200 µL de etanol 70% (v/v) gelado; a placa foi centrifugada por 5 minutos a 4000 rpm a 20ºC, e o sobrenadante descartado; a placa foi invertida sobre papel absorvente, centrifugada invertida até 600 rpm. A placa secou por 60 minutos à temperatura ambiente e foi ressuspendida em 60 µL de água milliQ (*overnight*).

Seleção dos insertos e sequenciamento

Este método garante que somente insertos que contenham microssatélites sejam amplificados usando *primer* que se reassocia ao sítio T7 como à seqüência de microssatélites presente no inserto. A utilização do *primer* T7 leva à amplificação do sítio inteiro. As reações de sequenciamento foram realizadas em um termociclador PTC-200 (MJ Research, Inc.) em um volume final de 10 µL contendo 2,0 µL de tampão Save Money (2,5 µL de $MgCl_2$ 2 M, 200 µL de Tris-HCl 1 M pH 9,0), 0,5 µL de *primer* T7 (5 pMol/µL), 0,4 µL de Big Dye e de 3 µL de DNA (1µg).

Esta reação foi submetida a uma etapa inicial de desnaturação a 96ºC por 2 minutos, seguida de 25 ciclos (96ºC por 45 segundos, 50ºC por 30segundos, 60ºC 4 minutos) e um passo final a 4ºC até retirar a reação da máquina. Para eliminar possíveis interferentes ou excessos de reagentes da reação de sequenciamento e deixar as amostras prontas para a eletroforese, foi realizada a purificação das amostras. Foram adicionados nas amostras 80 µL de isopropanol 65% (v/v) e foi agitado suavemente; ficou na bancada por 15 minutos e foi centrifugado a 3000 rpm por 45 minutos; o sobrenadante foi descartado e a placa foi mantida invertida por 1 minuto em papel absorvente para secar. Foram adicionados 200 µL de etanol 60% (v/v) e centrifugado a 3000 rpm por 10 minutos; o sobrenadante foi descartado e a placa foi mantida invertida por 1 minuto em papel absorvente para secar. A lavagem foi repetida. Foi dado um *spin* com placa invertida. A placa ficou no fluxo laminar por 1 hora para secar.

Depois de feita a extração e purificação, a seqüência de cada clone selecionado foi determinada através da análise em um seqüenciador automático modelo 3100 (*Applied Biosystem*).

Análises das seqüências

Os cromatogramas produzidos pelo seqüenciador foram analisados através do programa Chromas 2.21 (Technelysium Pty Ltd). As seqüências foram obtidas através do *3730/3730xl Data Collection Software v3.0 (Applied Biosystems)*, foram processadas com o auxílio do conjunto de *softwares Phred, Consed* e *Cross_Match* (*Laboratory of Phil Green, Genome Sciences Department, University of Washington* – http://www.phrap.org/index.html). Por meio de ferramentas disponíveis nestes programas foram retiradas das seqüências as regiões correspondentes ao vetor e ao adaptador, restando apenas a seqüência do inserto. Trechos de má qualidade e seqüências cujas bases não puderam ser identificadas com segurança (parâmetro *Phred* > 20), também foram excluídos. Com isso, arquivos do tipo "FASTA" foram obtidos, contendo apenas seqüências de interesse e com boa qualidade. Estas seqüências "FASTA" foram então processadas pelo *software CAP3 Sequence Assembly Program* (HUANG & MADAN, 1999), para identificar "contigs" e "singlets". Para construção dos "contigs" foi necessária uma sobreposição de no mínimo 20 pares de base entre duas seqüências, assim os "contigs" e "singlets" puderam ser analisados pelo programa TROLL (CASTELO & GAO, 2002).

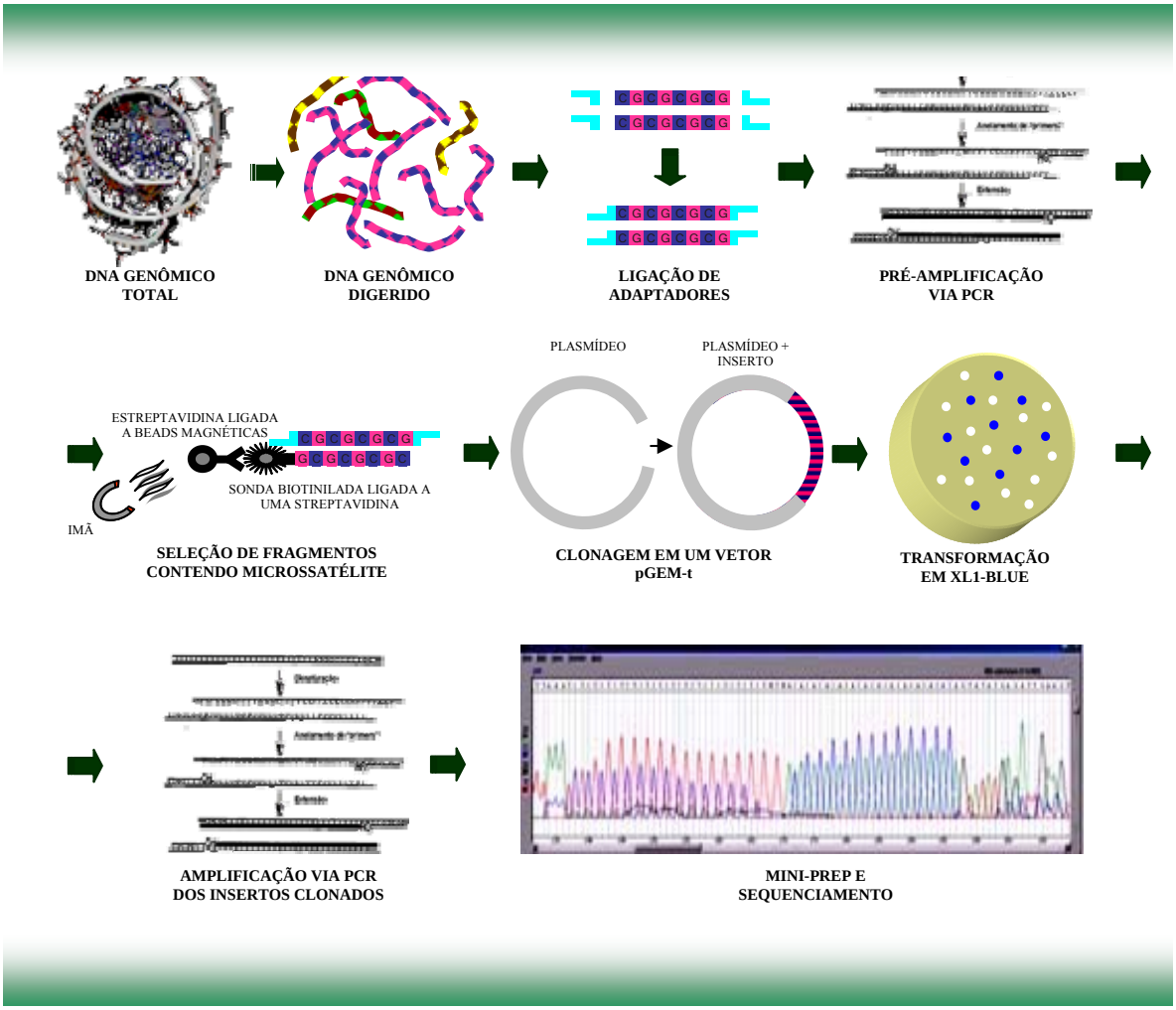


Figura 14 – Esquema representativo das etapas da construção da biblioteca enriquecida com microssatélites.

Foram adotados parâmetros para identificação das regiões de interesse, tais como: microssatélites mononucleotídeos deveriam ter no mínimo dez repetições do motivo, dinucleotídeos deveriam ter no mínimo cinco repetições do motivo, microssatélites constituídos por trinucleotídeos deveriam ter no mínimo quatro repetições do motivo, microssatélites constituídos por tetranucleotídeos deveriam ter no mínimo três repetições do motivo e pentanucleotídoes deveriam ter no mínimo duas repetições do motivo. Em virtude destes parâmetros, algumas seqüências foram excluídas da análise, restando apenas seqüências que continham microssatélites com características desejáveis. A partir daí, os pares de *primers* complementares às seqüências que flanqueiam as ilhas de repetição microssatélites, foram desenhados utilizando o programa TROLL (http://www.bioinformatica.ucb.br/troll.html). Foram

utilizados os critérios estabelecidos pelo programa, como: conteúdo mínimo de 50% e máximo de 60% do motivo microssatélite; diferença de temperatura de reassociação dos *primers* variando entre 55 a 70^{o}C e diferença de temperatura de reassociação entre os pares de *primers* variando no máximo de 3^{o}C (CASTELO & GAO, 2002). Para evitar o desenho de pares de *primers* redundantes todas as seqüências obtidas foram alinhadas entre si. O alinhamento foi realizado com auxílio do programa de alinhamento múltiplo Clustal X (THOMPSON et al., 1997). Visando facilitar a detecção de polimorfismo, os *primers* foram desenhados, sempre que possível, de forma que o produto de amplificação tenha no mínimo 100 e no máximo 250 pb.

As seqüências dos locos polimórficos foram anotadas e submetidas ao GeneBank, com o auxílio da ferramenta Banklt (http://www.ncbi.nlm.nih.gov/Banktl).

Amplificações iniciais, visualização dos produtos de PCR e interpretação do gel

Para cada par de *primers* desenhados, foram realizadas séries de amplificações visando alcançar as melhores condições para amplificação dos locos. Para as amplificações foi utilizada a enzima *Taq* DNA polimerase (Invitrogen). As amplificações foram padronizadas para cada loco, variando a temperatura de reassociação e as concentrações de $MgCl_2$, DNA molde, dNTP e enzima. *Tailed primers* foram adicionados à reação, conforme metodologia a seguir.

Os *primers* microssatélites obtidos no desenvolvimento desse estudo foram sintetizados por *Gene Link, Inc*. A metodologia de marcação de *primers* com fluorocromos foi a do *primer-tail*, descrita por MISSIAGLIA & GRATTAPAGLIA (2006). Seqüências de *primers* não especificas ao genoma estudado são acopladas a extremidade 5' do *primer* reverso de cada loco. Além disso, estas mesmas seqüências são sintetizadas com diferentes fluorocromos (FAM e JOE). Na reação de PCR as seqüências complementares são sintetizadas pela *Taq* DNA polimerase, juntamente com o *primer* que pareou com a fita de DNA. Deste modo, a partir do terceiro ciclo a seqüência marcada com fluorocromo complementar à cadeia sintetizada, pode também ser sintetizada, marcando com fluorocromo toda a cadeia reverso do loco microssatélite (Figura 15). Para este estudo, foram utilizadas duas seqüências denominadas OSD12 e ORY5437, provenientes do cromossomo Y humano e sem complementaridade com o genoma de *Acrocomia aculeata*. As seqüências de oligonucleotídeos marcadas com

fluorocromos, bem como os respectivos locos microssatélites a que foram acopladas estão na tabela 4, que está disponível em resultados.

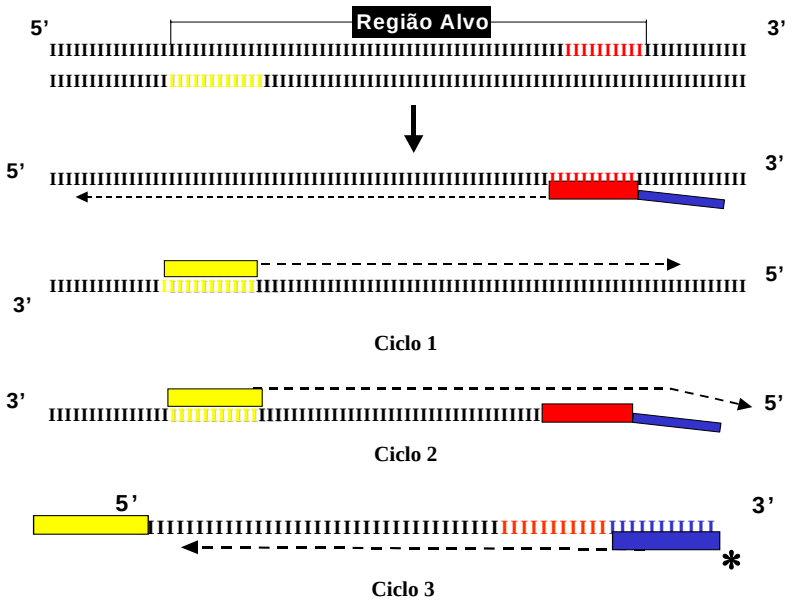


Figura 15 – Esquema representativo mostrando o princípio da estratégia do "*primer*-tail". A barra amarela representa o *primer* reverso que flanqueia o microssatélite. A barra vermelha representa o *primer forward* que flanqueia o microssatélite. A barra azul acoplada à barra vermelha representa a seqüência de um *primer* não específico para o genoma estudado ("tail"). A barra azul com um asterisco representa o *primer* não específico marcado com fluorocromo (Fonte: MISSIAGLIA , 2005).

3.3.3 Avaliação do polimorfismo dos locos microssatélites

A avaliação de polimorfismo foi feita com dois indivíduos de cada população objetivando-se pelo menos oito locos microssatélites polimórficos para a espécie estudada. Estes locos selecionados foram utilizados para a genotipagem dos indivíduos das nove populações de macaúba, visando o estudo populacional.

3.3.4 Condições de amplificação dos locos microssatélites

Para este trabalho foram testados 36 de pares de *primers* dos 77 desenvolvidos para a espécie e foram selecionados oito pares de *primers* para a genotipagem. Uma reação de PCR foi realizada para seis dos oito locos microssatélites e 47 genótipos avaliados, em um volume final de 20 μL contendo cerca de 30 ηg do DNA; 2,67 μL do *primer* direto (10 μM) e 2,67 μL do *primer* reverso (10 μM); 0,53 μL do *primer* marcado com fluorocromo; 250 μM de cada desoxirribonucleosídeo trifosfato (dNTP); solução tampão de Buffer 1 X (50 mM de KCl; 10 mM de Tris-HCl, pH 8,9); 1,20 μL de $MgCl_2$ (3 mM) e 1 U de *Taq* DNA polimerase. Para seis dos oito locos estudados, o ciclo total das amplificações foi realizado em um termociclador PTC200 (MJ Research, Inc.) programado para iniciar com 5 minutos a 94ºC para uma desnaturação inicial, seguido por 30 ciclos com 1 minuto de desnaturação a 94ºC, 1 minuto de pareamento a 56ºC e 1 minuto de extensão a 72ºC, finalizando com 8 minutos a 72ºC para extensão final.

Os outros dois locos que não amplificaram satisfatoriamente nas condições anteriores, foram submetidos a uma reação de PCR que foi realizada em um volume final de 20 μL contendo cerca de 30 ηg do DNA; 1,34 μL do *primer* direto (10 μM) e 1,34 μL do *primer* reverso (10 μM); 0,27 μL do *primer* marcado com fluorocromo; 250 μM de cada desoxirribonucleosídeo trifosfato (dNTP); solução tampão de Buffer 1 X (50 mM de KCl; 10 mM de Tris-HCl, pH 8,9); 1,20 μL de mM de $MgCl_2$ (3 mM) e 1 U de *Taq* DNA polimerase. Para estes dois locos dos oito locos estudados, o ciclo total das amplificações foi realizado em um termociclador PTC200 (MJ Research, Inc.) programado para iniciar com 5 minutos a 94ºC para uma desnaturação inicial, seguido por 30 ciclos com 1 minuto de desnaturação a 94ºC, 1 minuto de pareamento a 52ºC e 1 minuto de extensão a 72ºC, finalizando com 8 minutos a 72ºC para extensão final.

Os produtos de amplificação foram separados sob condições desnaturantes em gel 5% de poliacrilamida (v/v), contendo 8 M de uréia e 1 X TBE (0,045 M Tris-borato e 0,001 M EDTA) em seqüenciador automático (ABI, modelo 377) por cerca de duas horas e trinta minutos a 70 Watts de potência. De acordo com o funcionamento do equipamento, os sinais emitidos pela incidência do raio laser sobre os diferentes fluorocromos presentes nos *primers* foram captados e convertidos em sinais na forma de eletroferograma. Por meio de programas computacionais específicos, GeneScan® Analyses software versão 2.1.1 (ABPEC, Foster City, CA) e Genotyper® software, versão 2.0 (Applied Biosystems, Foster City, CA), os locos foram retidos para a realização de análises posteriores.

3.3.5 Metodologia de análise estatística dos dados

As freqüências gênicas ou alélicas e genotípicas em cada loco foram obtidas a partir da leitura dos dados nos géis de poliacrilamida. Estas freqüências foram submetidas a um teste de aderência (teste exato de Fisher) às proporções de equilíbrio de Hardy-Weinberg, conforme definido por WEIR (1996), utilizando o programa TFPGA (MILLER, 1997). O teste exato de Fisher foi feito por meio do método convencional de Monte Carlo utilizando 10 *batches* com 1000 permutações por *batch*.

A diversidade genética e as estatísticas F foram estimadas sob modelo aleatório de acordo com WEIR (1996), em que as populações amostradas são consideradas como representativas da espécie e com uma história evolutiva comum. As freqüências alélicas, o número de alelos por loco (A), a heterozigosidade observada (H_0) e esperada (H_e) e as estatísticas F de Wright ($F_{IS,}$ F_{ST} e F_{IT}), foram estimadas utilizando o programa GDA (LEWIS & ZAYKIN, 2006). As populações também foram agrupadas em regiões, conforme tabela 2 para cálculo das estatísticas F de Wright ($F_{IS,}$ $F_{SR,}$ F_{RT} e F_{IT}), sendo que para esta situação a estimativa do F_{ST} foi decomposta em F_{SR} e F_{RT}.

Em locos microssatélites, o processo mutacional não está de acordo com o que se admite no modelo de alelos infinitos com baixas taxas de mutação. Por essa razão, foi utilizada, além da estatística de F_{ST}, uma análoga a ela, denominada de R_{ST} (SLATKIN, 1995), desenvolvida especificamente para dados de microsatélites. Parâmetros como o R_{ST} e o fluxo gênico (Nm) foram estimados com o auxílio do programa R_{ST}Cal (Goodman, 1997). A estruturação da variabilidade foi visualizada através de dendrogramas construídos pela matriz de distâncias genéticas de NEI (1978) e pelo critério de agrupamento UPGMA, e para isto utilizou-se o programa NTSYS (ROHLF, 1989). A estabilidade dos agrupamentos também foi testada, empregando o procedimento de reamostragem por 10000 bootstrap, utilizando programa TFPGA (MILLER, 1997).

4 RESULTADOS

4.1 Desenvolvimento de Bibliotecas Enriquecidas com Locos Microssatélites

A extração do DNA genômico a partir de 200 g do material vegetal de um genótipo de macaúba forneceu uma concentração de aproximadamente 300 ηg/μL e foram utilizados 6 ηg na reação. Na digestão com a enzima *Rsal* o DNA genômico foi digerido não gerando nenhum tamanho preferencial de fragmento e produzindo o perfil adequado para a construção da biblioteca. A ligação dos adaptadores foi importante para garantir que cada fragmento tivesse uma terminação comum e conhecida A pré-amplificação foi realizada com sucesso, confirmando a ligação dos adaptadores aos fragmentos e gerando uma maior quantidade de DNA para a seleção. Não houve a amplificação aparente de fragmentos preferenciais, o que poderia ser prejudicial nas próximas etapas, levando a clonagem e sequenciamento de fragmentos preferenciais, aumentando o tempo e os custos para a identificação de microssatélites não redundantes.

A eficácia do enriquecimento foi verificada através dos produtos de PCR obtidos diretamente de colônias individuais e foi confirmada com a técnica de sequenciamento. Cerca de 88,78% dos clones analisados apresentaram presença de regiões microssatélites, sendo que muitas das seqüências apresentavam mais de uma região microssatélite.

4.2 Análise das Seqüências, Desenvolvimento e Utilização dos *Primers*

Através de uma interface web o programa TROLL localizou, nos 196 insertos seqüenciados, 174 que possuíam pelo menos um microssatélite. Nestas 174 seqüências foram encontradas um total de 303 motivos microssatélites, sendo 135 motivos dinucleotídeos, dois motivos trinucleotídeos, seis motivos tetranucleotídeos e 160 motivos pentanucleotídeos (Figura 16). Foram usados critérios estabelecidos pelo programa para localizar as regiões de elementos repetitivos

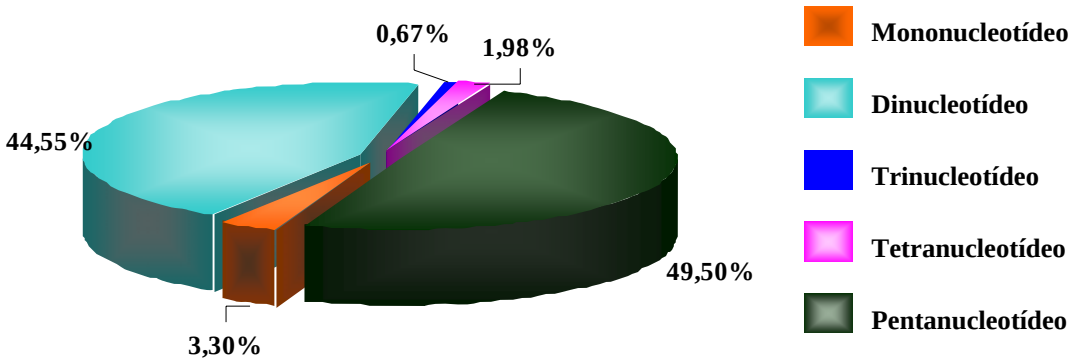


Figura 16 – Porcentagem de motivos microssatélites encontrados nos 174 insertos seqüenciados que apresentaram pelo menos uma região microssatélite.

Todas as seqüências obtidas foram alinhadas entre si e cerca de 5% apresentaram redundância, as que eram complementares e/ou sobrepostas foram excluídas. Setenta e sete destas seqüências puderam ser utilizadas para o desenho de *primers* complementares às regiões flanqueadoras dos microssatélites. As seqüências restantes não foram utilizadas, pois a região flanqueadora era também um microssatélite ou estava muito próximo de uma das extremidades da seqüência impossibilitando o desenho dos *primers*. Dentre as 77 seqüências escolhidas para o desenho de *primers* estão presentes diferentes motivos (Tabela 3).

Tabela 3 – *Primers* construídos a partir das seqüências microssatélites.

Primers	Motivo	Classificação	Produto (pb)	TA °C	Seqüência (*Primer*)	
1	$(AC)_{10}$	PERFEITO	265	59,80	F	TCAGCGAATTCAAGAGGACA
					R	GTGGACAAGAAAAGGGCAAA
2	$(TC)_{15}$	PERFEITO	230	60,03	F	TTCTCAGTTTCGTGCGTGAG
					R	GGGAGGCATGAGGAATACAA
3	$(CT)_{11}$ $(CA)_{15}$	COMPOSTO	262	60,02	F	ATTCTCTTGCTTACGCGTGG
					R	TCTTGCTTACGCGTGGACTA
4	$(GT)_7$	PERFEITO	273	59,23	F	TGGCATTTTGTTGGTGTGTT
					R	TCCCTCGTCATGACAATCTC
5	$(CA)_{10}$	PERFEITO	193	59,95	F	ACCGAGCTTGTTGACTTCGT
					R	

						AGCAAGATAGGTGGGTGGTG

Continua

Tabela 3 – *Primers* construídos a partir das seqüências microssatélites.

Primers	Motivo	Classificação	Produto (pb)	TA °C		Seqüência (*Primer*)
6	$(AT)_7$ $(TG)_5$ A $(TG)_{12}$	COMPOSTO/INTERROMPIDO	208	59,98	F	TGGATGCGTAGGGATGTGTA
					R	AGGATGCTGTCTGGCTCAGT
7	$(TATC)_7$ $(TG)_{29}$	COMPOSTO/INTERROMPIDO	243	56,60	F	TTATGGGGCTGAACTTGGAC
					R	GGTCACTTGACCCGTAATCC
8	$(CA)_8$	PERFEITO	112	59,59	F	CCCCATGGTTTAGCATTCAC
					R	CTTTGCTTACGCGTGGAATA
9	$(TG)_{10}$	PERFEITO	144	60,03	F	CTTGCTTACGCGTGGACATA
					R	TAAGATCTTCGCCCCCTTCT
10	$(TG)_8$	PERFEITO	265	59,95	F	TTCCTGACCATCGGAAAGAG
					R	CTGGACTGGAACGGGAGTAG
11	$(AC)_7$	PERFEITO	214	59,88	F	CTAGCGGTATGGTGGAGGAG
					R	ACGCGTGGACTAACCGTAAC
12	$(TG)_{18}$	PERFEITO	255	59,60	F	CGTCGAATGTAATTGTGTTGG
					R	GAGCCTGGTCACCGAAACT
13	$(CGAC)_5$	PERFEITO	225	60,06	F	CAGACTACCAGGCTTCCAGC
					R	TCATCATCGCAGCTTGACTC
14	$(CGCA)_3$	PERFEITO	264	59,98	F	TACCCCTCTGACGAATCCAG
					R	AGAGCGCGTGGACTAACAAT
15	$(AC)_{10}$	PERFEITO	269	59,69	F	TGTTTAAAGCTTTTGGCGC
					R	GGAAAAAGACATTGCCATCC
16	$(AT)_6$ AC $(GT)_{19}$	COMPOSTO/INTERROMPIDO	238	59,71	F	TTCACCATTCCATGCCATTA
					R	GCTTACGCGTGGACTAGCTT
17	$(TG)_9$	PERFEITO	226	59,94	F	GCTTACGCGTGGACAATACA
					R	GAACCCGGTCGACAGAACTA
18	$(GT)_8$	PERFEITO	191	60,14	F	GTGTGTGTGTGTGTGTGCGT
					R	GAACCCGGTCGACAGAACTA
19	$(AC)_{13}$ TG $(AC)_8$	INTERROMPIDO	249	59,80	F	TCATACTTGCTTACGCGTGG
					R	GCAAGTTCAAAATTGCACCA
20	$(AC)_{15}$	PERFEITO	150	60,09	F	GCAAAGACCAAGGCCTACAA
					R	ATGTTGCTCATTTTCCTGGG
21	$(TG)_9$ C $(GT)_6$	COMPOSTO/INTERROMPIDO	228	59,86	F	TCTTGCTTACGCGTGGACTA
					R	GGAAGAAACCGGTTGACAGA
22	$(GT)_8$	PERFEITO	234	60,07	F	GTGTGTGTGCGTGTGTGTGT
					R	GTGGCCGTGAATTCGATAGT
23	$(TG)_{18}$	PERFEITO	272	59,86	F	GCGTGGACTAACTCTCGGAT
					R	ATTCTCTTGCTTACGCGTGG
24	$(GT)_8$	PERFEITO	171	59,81	F	ACCAAGATGAGCATGGGAAG
					R	AAAAATCCTTTGACAGGCCA
25	$(AC)_8$	PERFEITO	262	59,88	F	ATTCTCTTGCTTACGCGTGG
					R	TGCCAAGATGCATCAGAGAT
26	$(TG)_7$	PERFEITO	219	59,95	F	TCTTGCTTACGCGTGGACTA
					R	TGAATTCCCCAAAGGCAATA

27	$(AC)_{13}$ $(AG)_{14}$	COMPOSTO	254	60,05	F R	ACTTGCAGCCCCATATTCAG GCAGGAACAGAGGCAAGTTC
28	$(TG)_8$	PERFEITO	150	60,35	F R	AAGTTGCAGCACAACAGTGG CAAGGGTGCAAACACACAGA

Continua

Tabela 3 – *Primers* construídos a partir das seqüências microssatélites.

Primers	**Motivo**	**Classificação**	**Produto (pb)**	**TA °C**	**Seqüência (*Primer*)**	
29	$(AC)_8$	PERFEITO	226	60,20	F R	GACTGCCGCAAAGGTGTTAT TCAAGTGGTTAGCTTGATCGG
30	$(TG)_8$	PERFEITO	277	60,00	F R	GAGAAAGAGCAACACGCCTC TCTGTCTGATGAGTTGAAGATAAAGAA
31	$(AT)_5$ $(TG)_{15}$	COMPOSTO	237	59,70	F R	TCTTGCTTACGCGTGGACTA AAATACAGCAGAGGCACGGT
32	$(TG)_9$	PERFEITO	278	59,25	F R	TTTCCCACTAGCTGAAAGCA AAAAGAGCGGTTTGTCTCCA
33	$(AGG)_4$	PERFEITO	134	59,96	F R	TATTTTCGTTTTCCGATGCC AGTCAACCTGCGAATCTGCT
34	$(TG)_{15}$ $(TA)_6$	COMPOSTO	257	59,86	F R	TACGCGTGGACTAGCAGATG CGCGTGGACTATCAATTTCA
35	$(TG)_{20}$	PERFEITO	280	59,50	F R	CTTGCTTACGCGTGGACTAA TGTGATTTTCTTACGCGTGC
36	$(GT)_5$ A $(TG)_9$	COMPOSTO/INTERROMPIDO	275	59,82	F R	TCTTGCTTACGCGTGGACTA AAACAACAACCTTGTTCCGC
37	$(TC)_{16}$ $(CA)_6$	COMPOSTO	264	59,84	F R	TCTTGCTTACGCGTGGACTA TTTGTGTGTGTGTGTGTGTG
38	$(AC)_{24}$	PERFEITO	201	59,93	F R	TCTCTCTCTCTCTCTCACACACA CAAACATGCAATACTGGATTTGA
39	$(AAG)_{12}$	PERFEITO	250	59,47	F R	ACACACACACACACACACAC GTGGACTAGCCATCTGCCAT
40	$(AC)_9$	PERFEITO	147	59,84	F R	TTCAGCAAAATCCATTGCAC GCTTTGCTTCTAGTCGGTGG
41	$(TG)_9$	PERFEITO	262	60,01	F R	ATAGCAGCTCAACGATGCCT GCGTGGACTAGCACATCTCA
42	$(TG)_{10}$ C $(GT)_6$	COMPOSTO/INTERROMPIDO	115	60,09	F R	TACGCGTGGACTAGCTGTG CGGCCAATTCAACTCTTGTT
43	$(TG)_{10}$	PERFEITO	273	59,90	F R	TGTCTGAGACCTATGGAGGTAGG GGAGATGTCCCATATGGTCG
44	$(CA)_9$	PERFEITO	272	59,84	F R	TCTTGCTTACGCGTGGACTA TCTTGCTTACGCGTGGACTA
45	$(GA)_{13}$	PERFEITO	149	60,03	F R	ATCGAAGGCCCTCCAATACT GAAATAAGGGGACCCTCCAA
46	$(CA)_{18}$	PERFEITO	167	59,83	F R	TCTCCCTTTTTGTCAGCAGTC TCAATCATTTCAGGAATTCATACG
47	$(TGTA)_3$	PERFEITO	280	59,42	F R	ACACGCACAGACACACACAC TCTTGCTTACGCGTGGACTA
48	$(AT)_8$	PERFEITO	110	59,79	F R	TCTTGCTTACGCGTGGACTA CCATAACATCAGGTTCGCCT

49	$(TG)_{21}$	PERFEITO	206	59,94	F R	CGTGGACTAACAGTGACCGA CCGCAACTAACCTGCAAAAT
50	$(AT)_5$ $(TG)_7$	COMPOSTO	105	60,29	F R	ATTCTCTTGCTTACGCGTGG ACGCACACACACACACACAC
51	$(TG)_{20}$	PERFEITO	253	60,40	F R	ATTCTCTTGCTTACGCGTGG GCCAGAACTAACCTGCGAAA

Continua

Tabela 3 – *Primers* construídos a partir das seqüências microssatélites.

Primers	**Motivo**	**Classificação**	**Produto (pb)**	**TA °C**	**Seqüência (*Primer*)**
52	$(CT)_{18}$	PERFEITO	202	60,09	F CACCATGATCTGTGCCATTC R GCATCTTTTCAGATTGCAACAA
53	$(TG)_{11}$	PERFEITO	122	59,47	F CGCGTGGACTAACTGTGACT R AGGTCACTCGACCCGTAATG
54	$(TC)_{26}$	PERFEITO	197	60,04	F GGAGTCGAAGCTGACGAAAG R GCCTTCTTGCACTGATCCTC
55	$(TG)_{17}$	PERFEITO	229	59,97	F TCCACCTTTAATGGGAGTGC R TAAACAGCGCCAGGTCTTCT
56	$(TG)_{22}$	PERFEITO	270	59,24	F CAGAAGCCGATTTCCTAATTG R TCTTGCTTACGCGTGGACTA
57	$(TA)_5$ $(TG)_{17}$	COMPOSTO	161	59,90	F ATTCGATTTTACGCGTGGAC R GAGTAGATCACTCGACCCGC
58	$(TTTC)_3$	PERFEITO	176	59,85	F AGAAGAACCGTTCAGCAAGC R ATTTCAAAAGGCAGGCAATG
59	$(GT)_{16}$	PERFEITO	132	59,14	F ATTCTCTTGCTTACGCGTGC R CATCAGATTTGCCTAAATACACC
60	$(GT)_{23}$	PERFEITO	266	59,68	F TCTTGCTTACGCGTGGACTA R GGACAAGTAGCGAATACCGC
61	$(TG)_{18}$	PERFEITO	204	59,75	F TCTTGCTTACGCGTGGACTA R CACCAGAACTCACCTGCAAA
62	$(TG)_8$	PERFEITO	183	59,73	F TCTTGCTTACGCGTGGACTA R TCATTGAGCAGAGTGTTGGG
63	$(AC)_{17}$	PERFEITO	233	60,07	F TCTGGACAATAGCGAGGAGC R TCTTGCTTACGCGTGGACTA
64	$(TG)_{10}$ A $(GT)_{15}$	COMPOSTO/INTERROMPIDO	104	60,23	F CGCGTGGACTAACAGAGGAT R GCCATGGCTCTCGTATCTGT
65	$(GT)_{11}$	PERFEITO	238	60,03	F CGTGGACTAACTGTGACCGA R GAACAAATGGCGAAATACCG
66	$(TG)_{17}$	PERFEITO	135	60,05	F TGTGACTGATATGCGTGGGT R GTAGGTCACTCGACCCGAAA
67	$(TGCT)_3$	PERFEITO	158	60,24	F GTATGCAGGGTTGCTATGCC R GTATGCAGGGTTGCTATGCC
68	$(TG)_7$	PERFEITO	142	59,80	F TGCCACATAGAGTGCTTGCT R GCTGGTCCTTGTTTGCTCTC
69	$(AG)_{16}$	PERFEITO	165	59,73	F TGCCACATAGAGTGCTTGCT R CTACCACATCCCCGTGAGTT
70	$(CA)_{15}$	PERFEITO	202	60,02	F TACTGTTGTGCCAAGTCCCA R GAGCACAAGGGGGATATCAA

71	$(TG)_9$	PERFEITO	206	59,90	F R	CGTGGACTTAATGTGGCTGA ATAGTGGCACGATGGAGGAG
72	$(CA)_8$	PERFEITO	211	59,93	F R	CTTACGCGTGGCACTAACAA TCAGGTATGGCCAACATTCA
73	$(GT)_7$	PERFEITO	261	59,69	F R	TCTTGCTTACGCGTGGACTA GCATCCTCCATAATTGGCAT
74	$(TG)_8$ C $(GT)_7$	COMPOSTO/INTERROMPIDO	249	59,90	F R	TAGCAAAATACCGTAGCGGC TCTCTTGCTTAATGCGTGGA

Continua

Tabela 3 – *Primers* construídos a partir das seqüências microssatélites.

Primers	**Motivo**	**Classificação**	**Produto (pb)**	**TA °C**	**Seqüência (*Primer*)**	
75	$(CA)_{18}$	PERFEITO	265	60,06	F R	TTGTGGAAGAAACAGGTCCC TCGCCTTGAGAAATTATGGC
76	$(TC)_{20}$	PERFEITO	191	60,12	F R	GAATGTGCGTGCTCAAAATG AATGCCAAGTGACCAAGTCC
77	$(AC)_{15}$ AG $(AT)_5$	COMPOSTO/INTERROMPIDO	274	60,09	F R	TGGTCACTCTTTGGCCCTAC TTACGCGTGGACTAGCTGTG

pb = pares de bases ; Ta ºC = temperatura de reassociação em ºC.

A figura 17 mostra a distribuição, nos 77 *primers* desenhados, dos motivos microssatélites perfeitos, classificados em dinucleotídeos, trinucleotídeo e tetranucleotídeos. A figura 18 mostra entre os *primers* desenhados qual a porcentagem de motivos microssatélites perfeitos, compostos e compostos interrompidos.

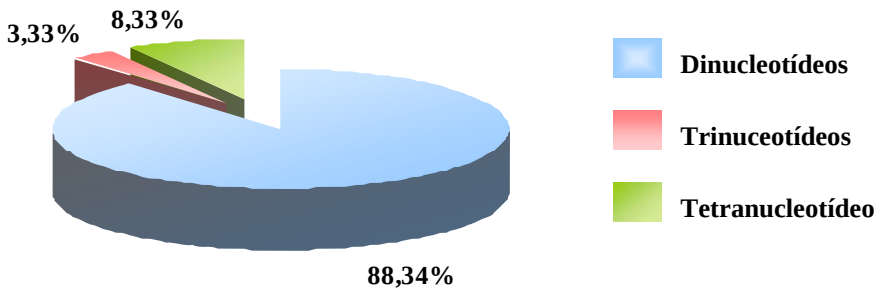


Figura 17 – Motivos microssatélites perfeitos e as freqüências de motivos dinucleotídeos, trinucleotídeos e tetranucleotídeos.

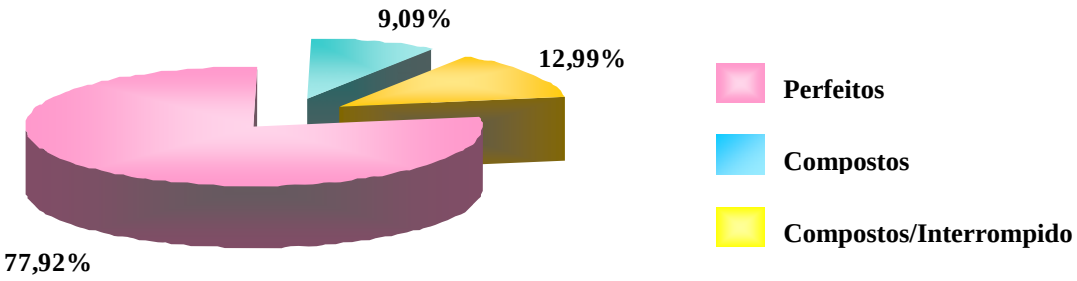


Figura 18 – Freqüência de motivos microssatélites perfeitos, compostos e compostos interrompidos.

Os motivos microssatélites mais freqüentes encontrados nos setenta e sete *primers* desenhados foram os dinucleotídeos TG e AC, outros motivos apareceram em porcentagem pequena e foram agrupados na categoria "outros" e estão representados na figura 19.

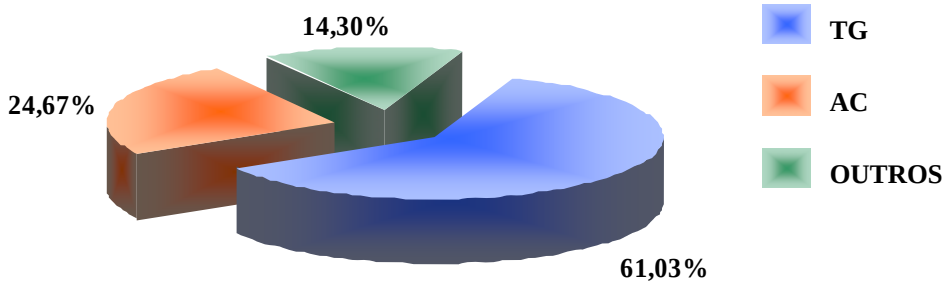


Figura 19 - Motivos microssatélites mais freqüentes identificados em *Acrocomia aculeata*.

Para a genotipagem foram escolhidos 36 pares de *primers* que foram sintetizados e testados (Tabela 4). Foram identificados pela seleção oito pares de *primers* que foram utilizados para caracterização genética e estudo da estrutura genética populacional. Inicialmente, todos os *primers* formam submetidos ao programa básico com 56°C para a reassociação do *primer*. Aqueles que não amplificaram satisfatoriamente foram submetidos a ciclos de amplificação com temperaturas de reassociação mais baixas. Os *primers* Aacu 32 e Aacu 35 produziram amplificação satisfatória a 52°C, conforme descrito em material e métodos.

Na tabela 5 são apresentadas as condições de amplificação utilizadas para os oito locos de microssatélites avaliados e suas respectivas amplitudes alélicas. A maior amplitude alélica observada foi de 430 pares de bases.

Tabela 4 – Relação dos 36 pares de *primers* sintetizados, mais as seqüências de oligonucleotídeos marcadas com fluorocromos.

LOCO	MOTIVO	***PRIMER*** Seqüência *Forward* (F) Seqüência *tail* + Seqüência *Reverse* (R)
Aacu1	$(TG)_{10}$ A $(GT)_{15}$	CGCGTGGACTAACAGAGGAT ACCAACCTAGGAAACACAGTGCCATGGCTCTCGTATCTGT
Aacu2	$(AT)_8(TG)_{21}$	TCTTGCTTACGCGTGGACTA ACCAACCTAGGAAACACAGTCCATAACATCAGGTTCGCCT
Aacu3	$(TG)_{10}$ C $(GT)_6$ CG $(GT)_5$	TTACGCGTGGACTAGCTGTG ACCAACCTAGGAAACACAGTCGGCCAATTCAACTCTTGTT
Aacu4	$(TG)_{11}$	CGCGTGGACTAACTGTGACT ACCAACCTAGGAAACACAGTAGGTCACTCGACCCGTAATG
Aacu5	$(GT)_{16}$	ATTCTCTTGCTTACGCGTGC ACCAACCTAGGAAACACAGTCATCAGATTTGCCTAAATACACC
Aacu6	$(TG)_{17}$	TGTGACTGATATGCGTGGGT ACCAACCTAGGAAACACAGTGTAGGTCACTCGACCCGAAA
Aacu7	$(GA)_{13}$	ATCGAAGGCCCTCCAATACT ACCAACCTAGGAAACACAGTGAAATAAGGGGACCCTCCAA
Aacu8	$(AC)_{15}$	GCAAAGACCAAGGCCTACAA ACCAACCTAGGAAACACAGTATGTTGCTCATTTTCCTGGG
Aacu9	$(TA)_5$ $(TG)_{17}$	ATTCGATTTTACGCGTGGAC ACCAACCTAGGAAACACAGTGAGTAGATCACTCGACCCGC
Aacu10	$(AG)_{16}$	TGCCACATAGAGTGCTTGCT ACCAACCTAGGAAACACAGTCTACCACATCCCCGTGAGTT
Aacu11	$(CA)_{18}$	TCTCCCTTTTTGTCAGCAGTC ACCAACCTAGGAAACACAGTTCAATCATTTCAGGAATTCATACG
Aacu12	$(TC)_{20}$	GAATGTGCGTGCTCAAAATG ACCAACCTAGGAAACACAGTAATGCCAAGTGACCAAGTCC

Aacu13	$(TC)_{26}$	GGAGTCGAAGCTGACGAAAG ACCAACCTAGGAAACACAGTGCCTTCTTGCACTGATCCTC
Aacu14	$(CT)_{18}$	CACCATGATCTGTGCCATTC ACCAACCTAGGAAACACAGTGCATCTTTTCAGATTGCAACAA
Aacu15	$(TG)_{18}$	TCTTGCTTACGCGTGGACTA ACCAACCTAGGAAACACAGTCACCAGAACTCACCTGCAAA
Aacu16	$(AT)_7$ $(TG)_5$ A $(TG)_{12}$	TGGATGCGTAGGGATGTGTA ACCAACCTAGGAAACACAGTAGGATGCTGTCTGGCTCAGT

Continua

Tabela 4 – Relação dos 36 pares de *primers* sintetizados, mais as seqüências de oligonucleotídeos marcadas com fluorocromos.

LOCO	MOTIVO	***PRIMER*** Seqüência *Forward* (F) Seqüência *tail* + Seqüência *Reverse* (R)
Aacu17	$(TG)_9$ C $(GT)_6$	TCTTGCTTACGCGTGGACTA GACTATGGGCGTGAGTGCATGGAAGAAACCGGTTGACAGA
Aacu18	$(TG)_{17}$	TCCACCTTTAATGGGAGTGC GACTATGGGCGTGAGTGCATTAAACAGCGCCAGGTCTTCT
Aacu19	$(AC)_{17}$	TCTGGACAATAGCGAGGAGC GACTATGGGCGTGAGTGCATTCTTGCTTACGCGTGGACTA
Aacu20	$(AT)_5$ $(TG)_{15}$	TCTTGCTTACGCGTGGACTA GACTATGGGCGTGAGTGCATAAATACAGCAGAGGCACGGT
Aacu21	$(AT)_6$ AC $(GT)_{19}$	TTCACCATTCCATGCCATTA GACTATGGGCGTGAGTGCATGCTTACGCGTGGACTAGCTT
Aacu22	$(TATC)_7$ $(TA)_4$ $(TG)_{29}$	TTATGGGGCTGAACTTGGAC GACTATGGGCGTGAGTGCATGGTCACTTGACCCGTAATCC
Aacu23	$(AC)_{13}$ TG $(AC)_8$	TCATACTTGCTTACGCGTGG GACTATGGGCGTGAGTGCATGCAAGTTCAAAATTGCACCA
Aacu24	$(TG)_8$ C $(GT)_7$	TAGCAAAATACCGTAGCGGC GACTATGGGCGTGAGTGCATTCTCTTGCTTAATGCGTGGA
Aacu25	$(TG)_{20}$	ATTCTCTTGCTTACGCGTGG GACTATGGGCGTGAGTGCATGCCAGAACTAACCTGCGAAA
Aacu26	$(AC)_{13}$ $(AG)_{14}$	ACTTGCAGCCCCATATTCAG GACTATGGGCGTGAGTGCATGCAGGAACAGAGGCAAGTTC
Aacu27	$(TG)_{18}$	CGTCGAATGTAATTGTGTTGG GACTATGGGCGTGAGTGCATGAGCCTGGTCACCGAAACTA
Aacu27	$(TG)_{15}$ $(TA)_6$	TACGCGTGGACTAGCAGATG GACTATGGGCGTGAGTGCATCGCGTGGACTATCAATTTCA
Aacu29	$(CT)_{11}$ $(CA)_{15}$	ATTCTCTTGCTTACGCGTGG GACTATGGGCGTGAGTGCATTCTTGCTTACGCGTGGACTA
Aacu30	$(CA)_{18}$	TTGTGGAAGAAACAGGTCCC GACTATGGGCGTGAGTGCATTCGCCTTGAGAAATTATGGC
Aacu31	$(GT)_{23}$	TCTTGCTTACGCGTGGACTA GACTATGGGCGTGAGTGCATGGACAAGTAGCGAATACCGC
Aacu32	$(TG)_{22}$	CAGAAGCCGATTTCCTAATTG GACTATGGGCGTGAGTGCATTCTTGCTTACGCGTGGACTA
Aacu33	$(AC)_{15}$ AG $(AT)_5$	TGGTCACTCTTTGGCCCTAC GACTATGGGCGTGAGTGCATTTACGCGTGGACTAGCTGTG

Aacu34	$(TG)_{18}$	TCTTGCTTACGCGTGGACTA GACTATGGGCGTGAGTGCATCTCTTGCTTACGCGTGGAT
Aacu35	$(TG)_{20}$	CTTGCTTACGCGTGGACTAA GACTATGGGCGTGAGTGCATTGTGATTTTCTTACGCGTGC
Aacu36	$(TC)_{16}$ $(CA)_6$ T $(AC)_{24}$ A $(AAG)_{12}$	TCTTGCTTACGCGTGGACTA GACTATGGGCGTGAGTGCATGTGGACTAGCCATCTGCCAT
OSD12 (JOE)		5' ACCAACCTAGGAAACACAGT 3'
ORY5437 (FAM)		5' GACTATGGGCGTGAGTGCAT 3'

Tabela 5 – Seqüências dos pares de *primers* desenvolvidos para *Acrocomia aculeata* que amplificaram locos microssatélites, com suas respectivas amplitudes alélicas, número de alelos, heterozigosidades esperadas e observadas, temperatura de reassociação, conteúdo do índice de polimorfismo e número de acesso no Genbank.

LOCO	MOTIVO	SEQUÊNCIAS DOS PARES DE *PRIMER* (MENOS O TAIL)	AMPLITUDE ALÉLICA (pb)	A*	H_e	H_o	Ta (°C)	PIC	Nº DE ACESSO GENBANK
Aacu07	$(GA)_{13}$	**F** ATCGAAGGCCCTCCAATACT **R** AAATAAGGGGACCCTCCAA	153-177	6	0,731	0,395	56	0,687	BV703504
Aacu10	$(AG)_{16}$	**F** TGCCACATAGAGTGCTTGCT **R** CTACCACATCCCCGTGAGTT	168-186	5	0,521	0,089	56	0,499	BV703505
Aacu12	$(TC)_{20}$	**F** GAATGTGCGTGCTCAAAATG **R** AATGCCAAGTGACCAAGTCC	190-202	7	0,645	0,333	56	0,618	BV703506
Aacu18	$(TG)_{17}$	**F** TCCACCTTTAATGGGAGTGC **R** TAAACAGCGCCAGGTCTTCT	173-193	2	0,451	0,543	56	0,399	BV703507
Aacu26	$(AC)_{13}$ $(AG)_{14}$	**F** ACTTGCAGCCCCATATTCAG **R** CAGGAACAGAGGCAAGTTC	273-316	8	0,762	0,325	56	0,737	BV703508
Aacu30	$(CA)_{18}$	**F** TGTGGAAGAAACAGGTCCC **R**TCGCCTTGAGAAATTATGGC	148-158	4	0,515	0,187	56	0,466	BV703509
Aacu32	$(TG)_{22}$	**F** AGAAGCCGATTTCCTAATTG **R** TCTTGCTTACGCGTGGACTA	229-233	2	0,053	0,054	52	0,042	BV703510
Aacu35	$(TG)_{20}$	**F** CTTGCTTACGCGTGGACTAA **R** TGTGATTTTCTTACGCGTGC	380-430	6	0,752	0,226	52	0,713	BV703511
Média				5	0,535	0,269			

A*: número de alelos; H_e e H_o: heterozigosidades esperada e observada, respectivamente; Ta (°C): temperatura de reassociação; PIC: conteúdo do índice de polimorfismo.

As freqüências alélicas de cada loco estão representadas graficamente na Figura 20. Foram encontrados 40 alelos distribuídos entre os oito locos analisados nas nove populações.

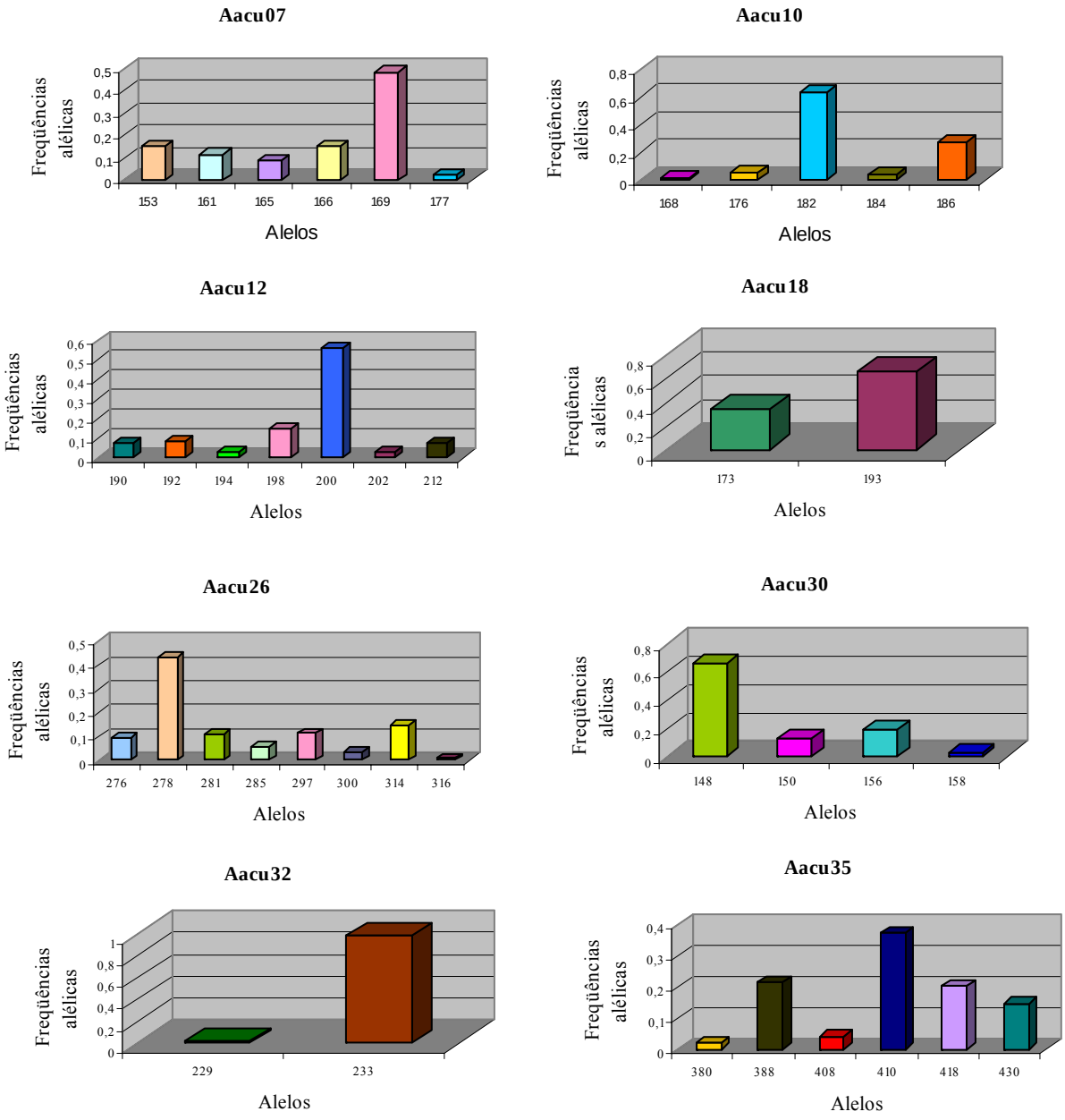


Figura 20 – Histograma das freqüências alélicas dos oito locos microssatélites, estimados para 47 indivíduos de *Acrocomia aculeata*. O eixo Y indica a freqüência alélica e o eixo X indica o número do alelo, em pares de base.

4.3 Variação Genética

O número médio de alelos por loco foi de 2,581, variando de dois alelos (locos Aacu18 e Aacu32) ao máximo de oito alelos (loco Aacu26), como se pode observar na figura 20 e na tabela 6. Na tabela 5 pode-se verificar a maior e a menor

heterozigosidade observada e esperada nos locos. O loco Aacu18 apresentou a maior heterozigosidade observada (0,543) e o Aacu32 a menor (0,054). A maior heterozigosidade esperada foi verificada no loco Aacu26 (0,762) e a menor (0,053) no loco Aacu32. A Figura 21 apresenta a variabilidade alélica existente nos locos Aacu07 e Aacu26.

Tabela 6 – Estimativas de parâmetros genéticos de diversidade em nove populações de *Acrocomia aculeata*.

Pop	N	n_A	Ap	H_e	H_o	f	t_a
1	7	2,125	2,500	0,270	0,202	0,267	0,578
2	5	2,125	2,500	0,365	0,323	0,132	0,766
3	5	2,286	2,500	0,428	0,343	0,202	0,663
4	5	2,250	2,250	0,555	0,467	0,243	0,609
5	5	2,250	3,000	0,364	0,225	0,410	0,419
6	6	2,50	3,000	0,445	0,221	0,523	0,313
7	4	2,375	2,833	0,458	0,229	0,540	0,460
8	6	1,833	2,250	0,320	0,222	0,279	0,563
9	4	1,875	2,400	0,366	0,177	0,550	0,291
MÉDIA		**2,180**	**2,581**	**0,397**	**0,268**	**0,359**	**0,471**

N: número de indivíduos amostrados; n_A: número médio de alelos; Ap: número total de alelos por loco polimórfico; H_o: heterozigosidade observada; H_e: heterozigozidade esperada sob equilíbrio de Hardy-Weinberg; f: índice de fixação, t_a: taxa de cruzamento aparente; pop1: Campinas, Jaguariúna; pop2: Serra Negra, Amparo, Lindóia; pop3; Mococa; pop 4: Casa Branca; pop 5: Dourado; pop 6: Barbosa, Vicentinópolis; pop 7; São Pedro do Turvo; pop 8: Ibituruna – Gupiara; pop 9: Ibituruna – Lagoa.

O índice de fixação intrapopulacional ($\hat{f} \approx \hat{F}_{IS}$) estimado em cada população através das heterozigosidades (H_o e H_e), estão apresentados na tabela 6. O valor médio foi de $\hat{f}$ = 0,359, com variação de 0,132 a 0,550. Pelo processo da análise de da variância (Tabela 9) obteve-se $\hat{f}$ = 0,3490 (IC 95% 0,1586 a 0,5096). Estes resultados, juntamente com os testes exatos de Fisher, indicam que as populações são mistas.

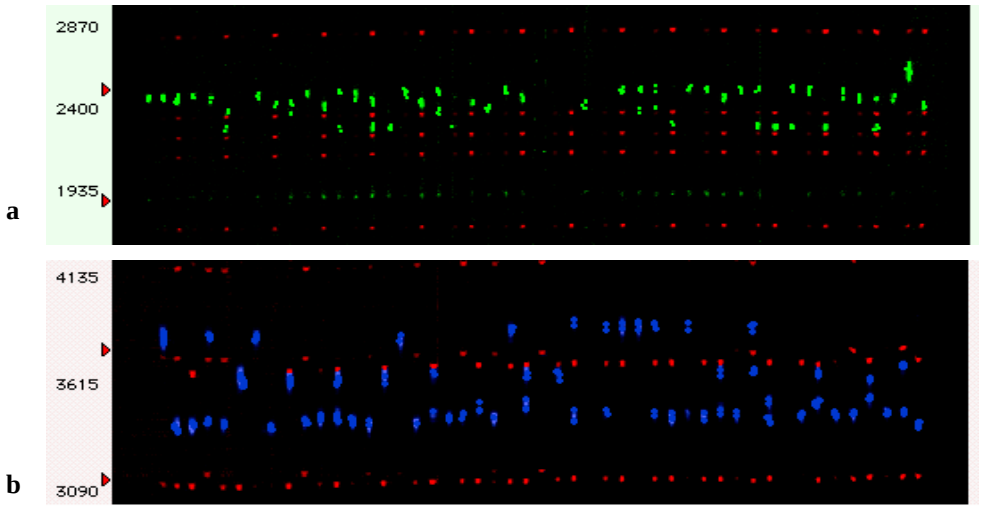


Figura 21 – - Perfil de dois géis de microssatélites. a. Utilizando o *primes* Aacu07 (seis alelos); b. *primer* Aacu26 (oito alelos), ambos em 47 genótipos de macaúba.

Através do teste exato de Fisher (Tabela 7) verificou que a maioria das populações não se encontra em equilíbrio de Hardy-Weinberg, para a maioria dos locos estudados. A população cinco não mostrou significância ao teste exato de Fisher em dois dos oito locos, portanto, em oito locos estudados seis deles não estão no equilíbrio de Hardy-Weinberg nesta população, enquanto que na população seis apenas um loco não mostrou significância ao teste exato de Fisher, portanto, em oito locos estudados sete deles não estão no equilíbrio de Hardy-Weinberg nesta população. É importante ressaltar que os desvios do equilíbrio de Hardy-Weinberg, neste caso, não tem a ver apenas com o sistema reprodutivo, mas também com o tamanho amostral (N).

Tabela 7 – Probabilidade do Teste Exato de Fisher para aderência ao equilíbrio de Hardy-Weinberg.

Loco	Pop1	Pop2	Pop3	Pop 4	Pop 5	Pop 6	Pop 7	Pop 8	Pop 9
Aacu07	1,0000	1,0000	1,0000	1,0000	0,3333	0,0303	0,4286	1,0000	-
Aacu 10	-	-	1,0000	0,4286	0,0476	0,3333	0,1429	1,0000	-
Aacu 12	1,0000	0,4286	-	1,0000	1,0000	1,0000	1,0000	-	0,3333

(-) = Equilíbrio de Hardy-Weinberg não pode ser calculado devido a população apresentar apenas um alelo por loco; pop1: Campinas, Jaguariúna; pop2: Serra Negra, Amparo, Lindóia; pop3; Mococa; pop 4: Casa Branca; pop 5: Dourado; pop 6: Barbosa, Vicentinópolis; pop 7; São Pedro do Turvo; pop 8: Ibituruna – Gupiara; pop 9: Ibituruna – Lagoa.

Continua

Tabela 7 – Probabilidade do Teste Exato de Fisher para aderência ao equilíbrio de Hardy-Weinberg.

Loco	Pop1	Pop2	Pop3	Pop 4	Pop 5	Pop 6	Pop 7	Pop 8	Pop 9
Aacu 18	0,4406	1,0000	0,1270	1,0000	0,4286	1,0000	0,1429	1,0000	1,0000
Aacu 26	0,4406	1,0000	0,1111	1,0000	0,0476	1,0000	0,4286	-	0,1429
Aacu 30	0,2308	1,0000	1,0000	1,0000	-	-	-	0,3333	0,1429
Aacu 32	-	-	-	1,0000	-	-	-	-	-
Aacu 35	1,0000	0,1429	0,2000	0,0857	-	0,2000	1,0000	-	1,0000

(-) = Equilíbrio de Hardy-Weinberg não pode ser calculado devido a população apresentar apenas um alelo por loco; pop1: Campinas, Jaguariúna; pop2: Serra Negra, Amparo, Lindóia; pop3; Mococa; pop 4: Casa Branca; pop 5: Dourado; pop 6: Barbosa, Vicentinópolis; pop 7; São Pedro do Turvo; pop 8: Ibituruna – Gupiara; pop 9: Ibituruna – Lagoa.

Foram identificados entre os 40 alelos existentes na população nove alelos exclusivos nas populações analisadas, sendo um na população um, um na população quatro, três alelos na população cinco, um alelo na população seis e três alelos exclusivos na população nove. Considerando que as populações não compartilham os mesmos alelos, não há um número expressivo de alelos que caracterizam determinadas populações. As populações que apresentam maior número de alelos exclusivos são as populações cinco e nove. As respectivas freqüências destes alelos podem ser observadas na tabela 8.

Tabela 8 – Alelos exclusivos obtidos em 40 alelos de marcadores microssatélites, e freqüência com que ocorrem em nove populações de *Acrocomia aculeata*.

Loco	Alelo (pb)	Freqüência	População
Aacu07	177	0,200	Pop 5
Aacu10	168	0,100	Pop 6
Aacu10	184	0,400	Pop 5
Aacu12	202	0,200	Pop 5
Aacu12	212	0,500	Pop 9
Aacu26	316	0,071	Pop 1
Aacu32	229	0,200	Pop 4
Aacu35	408	0,333	Pop 9

Continua

Tabela 8 – Alelos exclusivos obtidos em 40 alelos de marcadores microssatélites, e freqüência com que ocorrem em nove populações de *Acrocomia aculeata*.

Loco	Alelo (pb)	Freqüência	População
Aacu35	380	0,167	Pop 9

pb = pares de bases

4.4 Estrutura Genética

O índice de fixação da espécie ($\hat{F}_{IS}$) estimado com base nos oito locos de microssatélites, 40 alelos e nas nove populações apresentou valor 0,3490. Este valor foi acima do esperado para uma espécie alógama e abaixo do esperado para uma espécie autógama, sugerindo que a espécie é provavelmente mista.

O valor de R_{ST} estimado foi de 0,2084 e o de F_{ST} igual a 0,2522. Estas estimativas foram congruentes e significativamente diferentes de zero, como pode se observar na tabela 9. O parâmetro *Nm* estimado com base na estimativa de $R_{ST,}$ resultou no valor de 0,9406 indivíduos, indicando uma taxa de migrantes entre as populações de grandeza intermediária. Este valor pode ser considerado como indício que as populações encontram-se em um processo de diferenciação e com possível comprometimento de sua estrutura metapopulacional, porém se faz necessário um estudo com um maior número de indivíduos e populações para confirmar estas estimativas.

Tabela 9 – Estimativa das estatísticas F de Wright, do R_{ST} e do numero de migrantes por geração (Nm) em nove populações naturais de *Acrocomia aculeata*. Intervalo de confiança (IC) de 95% de probabilidade.

	F_{IS}	F_{IT}	F_{ST}	R_{ST}	Nm*
Estimativa	0,3490	0,5132	0,2522	0,2084	0,9406
Limite Superior (IC 95%)	0,5096	0,6720	0,3529		
Limite Inferior (IC 95%)	0,1586	0,2881	0,1437		

* Nm calculado com base no R_{ST}

Agrupando essas populações por região de coleta, conforme tabela 2, foram estimados os índices de fixação F_{IS} , F_{IT} e F_{ST} decomposta em F_{SR} e F_{RT} com oito locos de microssatélites (40 alelos) em nove populações e quatro regiões. Estas estimativas podem ser observadas na tabela 10. O F_{SR} foi de 0,2712 e o F_{RT} foi de 0,1587, sendo estas medidas relacionadas à variação entre populações dentro de regiões e apenas entre regiões. A maior variabilidade como pode ser observado está entre populações indicando que a maior endogamia ocorre entre populações dentro de regiões.

Tabela 10 – Estimativa das estatísticas *F* de Wright em nove populações naturais, agrupadas em quatro regiões, de *Acrocomia aculeata*. Intervalo de confiança (IC) de 95% de probabilidade.

	F_{IS}	F_{IT}	F_{SR}	F_{RT}
Estimativa	0,3532	0,5286	0,2712	0,1587
Limite Superior (IC 95%)	0,5136	0,6851	0,3805	0,3131
Limite Inferior (IC 95%)	0,1451	0,2983	0,1521	0,0494

As distâncias genéticas de NEI (1978) calculadas entre as populações variaram de -0,0482 a 0,9061. O respectivo dendrograma está apresentado na figura 22. A correlação cofenética do agrupamento UPGMA desta matriz foi elevada (0,99). Esta análise revela que a distância entre as populações Serra Negra e Casa Branca é nula, e entre São Pedro do Turvo e as demais foi verificada a maior distância que foi de 0,5651. Convém observar que as populações de Ibituruna – Lagoa e Ibituruna – Gupiara também formaram um grupo distinto. Portanto, esta estruturação que foi visualizada através do dendrograma, evidencia dois grupos distintos formados, sendo um grupo composto pelas populações de Campinas, Serra Negra, Casa Branca, Mococa, Dourado e Barbosa, e o outro composto pelas populações de Ibituruna – Lagoa e Ibituruna – Gupiara, estes dois grupos são distintos geneticamente da população de São Pedro do Turvo.

Tabela 11 – Distâncias genéticas de NEI (1978), calculadas entre as populações.

	Pop 1	Pop 2	Pop 3	Pop 4	Pop 5	Pop 6	Pop 7	Pop 8	Pop 9
Pop 1	0,0000								
Pop 2	0,0102	0,0000							
Pop 3	0,1088	0,0819	0,0000						
Pop 4	-0,0275	-0,0482	0,0042	0,0000					
Pop 5	0,2503	0,2376	0,1084	0,1092	0,0000				
Pop 6	0,3972	0,2661	0,2097	0,2553	0,2195	0,0000			
Pop7	0,4869	0,5084	0,4428	0,5681	0,4883	0,4233	0,0000		
Pop 8	0,3171	0,4223	0,2440	0,1039	0,2335	0,5537	0,6966	0,0000	
Pop 9	0,5407	0,4891	0,6311	0,2723	0,7093	0,7344	0,9061	0,2438	0,0000

Pop1: Campinas, Jaguariúna; pop2: Serra Negra, Amparo, Lindóia; pop3; Mococa; pop 4: Casa Branca; pop 5: Dourado; pop 6: Barbosa, Vicentinópolis; pop 7; São Pedro do Turvo; pop 8: Ibituruna – Gupiara; pop 9: Ibituruna – Lagoa.

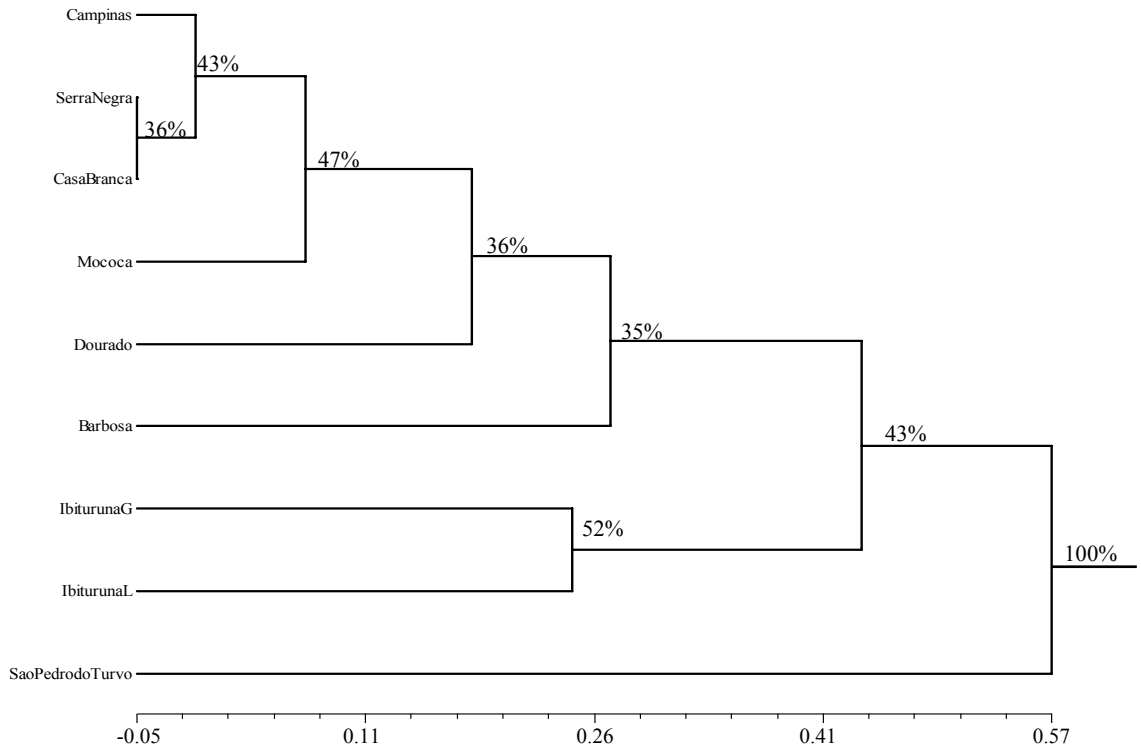


Figura 22 – Padrão de divergência genética entre nove populações de macaúba, definido pelo agrupamento UPGMA, a partir nas das distâncias genéticas de Nei (1978). Correlação cofenética igual a 0,99%. (1000 reamostragens bootstrap para porcentagem de consistência dos nós)

5 DISCUSSÃO

Até pouco tempo o uso de marcadores moleculares microssatélites esbarrava na dificuldade causada pelo alto custo e tempo despendidos no desenvolvimento de *primers* para cada espécie. Porém, atualmente, novos protocolos para desenvolvimento de bibliotecas têm facilitado muito o uso desta técnica, como os protocolos desenvolvidos por BILLOTTE et al. (1999) e KIJAS et al. (1994).

A transferibilidade de microssátelites entre espécies relacionadas é uma conseqüência da homologia da seqüência de DNA das regiões que flanqueiam os microssatélites. No início deste estudo foram testados *primers* desenvolvidos para *Euterpe edulis* em macaúba, porém não houve homologia. BILLOTTE et al. (2004) na construção de uma biblioteca enriquecida para *Bactris gasipaes*, desenvolveram 18 marcadores microssatélites, e ao avaliar a conservação de polimorfismo em experiências de transferibilidade verificaram em 10 acessos de *Bactris setulosa* e em 25 acessos de *Astrocaryum macrocalyx* e *A. urostachys* a transferibilidade potencial e utilidade de marcadores microssatélites de *Bactris gasipaes* para outros gêneros da tribo Cocoeae.

Vários estudos com árvores tropicais têm evidenciado uma alta razão de transferibilidade de locos de microssatélites entre espécies arbóreas taxonomicamente relacionadas, como ocorre dentro das Leguminosae (DAYANADAN et al., 1997), Meliaceae (WRITE & POWELL, 1997) e entre espécies de *Eucalyptus* (BRONDANI et al., 1998). A absoluta transferibilidade (100% de 10 locos) foi verificada em microssatélites desenvolvidos com *Caryocar brasiliense* para cinco outras espécies do mesmo gênero (*C. coriaceum, C.edule, C. glabrum, C. pallidum e C villosum*) indicando um alto nível de homologia do genoma e permitindo, dessa forma, estudos comparativos da estrutura genética em todas estas espécies (COLEVATI et al., 1999). Com base nos resultados desse estudo suspeita-se que a população de São Pedro do Turvo é formada por indivíduos da outra espécie do gênero estudado, conforme é discutido mais adiante. Caso isso seja confirmado temos aqui a transferibilidade dos oitos locos em *A. aculeata.*

Neste estudo, o desenvolvimento da biblioteca enriquecida com locos microssatélites para *Acrocomia aculeata* foi realizado com sucesso, gerando os produtos esperados. Entre os insertos seqüenciados 88,78% apresentaram pelo menos um loco microssatélite, dos quais 39,29% foram utilizados para desenho dos *primers* e, destes

46,75% (36 pares de *primers*) foram sintetizados. Estes locos foram selecionados com base no número de repetições do motivo SSR, pois o uso de repetições maiores facilita a detecção de polimorfismo entre os genótipos analisados. Os oito *primers* utilizados neste estudo foram selecionados a partir da otimização da reação de PCR a partir dos 36 *primers* sintetizados, e demonstraram o polimorfismo existente entre os indivíduos estudados. Os motivos mais freqüentes encontrados na análise foram os pentanucleotídeos, e isto se deve aos critérios de busca próprios do programa utilizado, porém não foram usados para síntese de *primers* por se tratarem de locos com apenas duas repetições dificultando a identificação de polimorfismo entre indivíduos. Foram localizadas poucas seqüências repetitivas de motivos trinucleotídeos e tetranucleotídeos.

A sonda de motivos trinucelotídeos usada neste procedimento não obteve sucesso no resgate de seqüências com tais repetições, isto pode ser devido à falta de homologia destes locos no genoma da espécie ou devido a algum problema com a sonda utilizada no desenvolvimento da técnica, por exemplo: degradação da sonda ou concentração não ideal de oligonucleotídeos, outra explicação para o fato que deve ser considerada é que altas taxas de trinucleotideo são encontradas em regiões transcritas, 5´- UTRs, 3´- UTRs e ORFs (MORGANTE et al, 2002). Evolutivamente regiões codantes têm seleção negativa para todos os motivos microssatélites menos para os trinucleotideos (MORGANTE et al, 2002). No estudo realizado utilizamos DNA genômico, muitos desses locos microssatélites estão em regiões não codantes, onde se encontram a maioria das seqüências de mononucleotídeos, dinucleotídeos, tetranucleotídeos e pentanucleotídeos.

Os motivos dinucleotídeos também foram encontrados em grande quantidade, isto se explica pelo fato de que, das três sondas usadas no enriquecimento da biblioteca, duas eram de repetições de motivos dinucleotídeos.

Outros estudos em palmeiras, conforme GAIOTTO et al. (2001), no desenvolvimento de marcadores microssatélites para *Euterpe edulis*, 475 clones analisados apresentaram repetições de motivos $(AG)_n$; destes, 170 apresentavam tamanho e posição adequados e 64 foram usados para síntese de *primers*. Um total de 18 pares de *primers* foram utilizados para as análises genéticas. BILLOTTE et al. (2004) isolaram, de 34 clones contendo regiões microssatélites, 18 locos $(GA)_n$ para estudo da diversidade genética em *Bactris gasipaes* Kunth.

A escolha de estratégias eficientes para a conservação *ex situ* e *in situ* do germoplasma da espécie *Acrocomia aculeata* depende principalmente do prévio

conhecimento da distribuição de sua variabilidade genética entre e dentro de populações naturais. Os marcadores microssatélites são ferramentas poderosas para este tipo de estudo e foram utilizados para cálculo de estimativas populacionais.

Em relação à variação genética observou-se no presente trabalho um nível considerável de multialelismo nos locos polimórficos estudados. O número médio de alelos por loco foi de 5,0 e a heterozigozidade esperada média foi de 0,535, sendo esta menor do que a encontrada em estudos com outras palmeiras utilizando o mesmo tipo de marcador molecular, como GAIOTTO et al. (2001), que em *Euterpe edulis* encontraram uma estimativa de H_e média de 0,860 estudando duas populações naturais localizadas em regiões de cerrado próximo ao Distrito Federal. BILLOTTE et al. (2004) encontraram uma estimativa de H_e médio (0,858) utilizando sete locos microssatélite desenvolvidos para *Bactris gasipaes* em 85 acessos demonstrando a importância desta técnica para estudos populacionais.

Valores mais altos de endogamia foram encontrados na região formada pelas populações cinco (Dourado), seis (Barbosa e Vicentinópolis), sete (São Pedro do Turvo) e em outra região na população nove (Ibituruna – Lagoa). Nas populações cinco, seis e sete os coeficientes de endogamia (0,410, 0,523 e 0,540 respectivamente) podem ser explicados pelo pequeno número amostral, pela existência de poucas e pequenas populações de macaúba na região, estas populações estão isoladas, cercadas por agricultura ou pastagens, contribuindo para a escassez de polinizadores da espécie, além dos indivíduos estarem muito próximos um do outro. Estes fatores favorecem o cruzamento entre indivíduos aparentados e podem gerar uma tendência de estreitamento da variabilidade genética, pela diminuição do fluxo gênico entre as populações. Entretanto, o valor mais alto de endogamia foi encontrado na população nove (0,550). Este valor era esperado, pois os indivíduos coletados estavam em volta de uma única planta adulta com cerca de 30 anos, num raio de dois quilômetros não havia outra planta adulta e a endogamia deve ser resultado de auto-fecundação ou fecundação entre plantas irmãs, ou seja pelo seu alto grau de parentesco.

Em contrapartida, valores mais baixos de endogamia foram encontrados na região em que se encontram as populações um (Campinas e Jaguariúna), dois (Serra Negra, Lindóia e Amparo), três (Mococa), quatro (Casa Branca). Nas populações um, dois, três e quatro os coeficientes de endogamia (0,267, 0,132, 0,202 e 0,243 respectivamente) podem ser explicados pela maior distância geográfica entre os indivíduos das populações, pela maior ocorrência de plantas mais antigas que não

sofreram tanta influência da fragmentação do ambiente ou da antropização. A população oito (Ibitutuna-Gupiara) que está na mesma região da população nove, também apresentou um baixo valor de endogamia (0,279), e aqui também há maior ocorrência de plantas mais antigas que não sofreram tanta influência da fragmentação do ambiente, trata-se de uma população maior e a distância entre a primeira e a última planta coletada é de dois quilômetros, e estão distribuídas ao longo da margem do Rio das Mortes, estes fatores podem explicar o predomínio de fecundação cruzada entre indivíduos diferentes e, provavelmente maior ocorrência de polinizadores, devido à menor degradação ambiental.

Uma característica relevante observada a ser destacada é que nos locais onde são encontrados os indivíduos de macaúba existem ótimos bancos de sementes importantes para fluxo gênico temporal.

O valor de F_{IS} encontrado em nosso estudo (0,359) pode ser considerado alto nas populações de *A. aculeata.* Este valor sugere que a espécie tem comportamento misto de reprodução e as frequências genotípicas destas populações encontram-se fora das proporções de equilíbrio de Hardy-Weinberg, para este grupo de indivíduos nesta geração.

A taxa de fecundação cruzada aparente, estimada foi considerada intermediária, tendo um valor médio de 0,471. Esta estimativa está coerente com os dados de SCARIOT et al. (1991), pois neste estudo com uso de marcadores microssatélites, confirmamos que em locais onde as populações estão muito fragmentadas e isoladas ocorre maior taxa de endogamia, confirmando que a espécie que é monóica e autocompatível realiza autofecundação e/ou cruzamentos entre indivíduos aparentados, pois os polinizadores não têm muita escolha. Ao contrário, em locais onde as populações estão menos perturbadas e menos isoladas ocorre endogamia, mas prevalece a fecundação cruzada entre indivíduos diferentes, favorecendo o fluxo gênico e aumentando a variabilidade da população.

Quanto à medida da diversidade entre populações, notou-se que as estimativa de R_{ST} e F_{ST} (0,2084 e 0,2522 respectivamente) foram muito semelhantes. Estes valores foram mais elevados que os encontrados por GAIOTTO et al. (2001) com duas populações da espécie *Euterpe edulis* (R_{ST} 0,07) e menor que a estimativa encontrada para três populações de cupuaçuzeiro θ_p = 0,307 (ALVES, 2002). Este parâmetro de diversidade entre populações é importante em estudos de conservação. Esta estimativa é

inédita e vem contribuir para a conservação da espécie e o direcionamento de novas coletas.

As nove populações estudadas foram agrupadas em quatro regiões para se obter a decomposição F_{ST}, para mensurar a variação entre populações dentro de regiões (F_{SR}) e entre as regiões (F_{RT}). Os resultados obtidos indicam que a maior variação ainda se encontra entre populações (F_{SR} = 0,2712) do que entre regiões (F_{RT} = 0,1587). Essas estimativas têm grande valia também no futuro na coleta de genótipos, indicando maior importância na coleta de indivíduos entre populações da mesma região do que entre regiões.

GOVINDAJARU (1989) distinguiu três níveis de fluxo gênico: alto Nm>1, intermediário (0,25<Nm<0,99) e baixo Nm<0,25. De acordo com esta definição o fluxo gênico encontrado neste estudo foi intermediária (0,9406). Sendo a estimativa de fluxo gênico calculada através do parâmetro $R_{ST,}$ o fluxo estimado neste trabalho não deve ser considerado atual, mas baseado na história genética destas populações. Outro motivo que deve ser considerado é o tamanho da amostragem deste estudo que é pequena. Sendo assim, deve-se atentar para este fato nos estudos de conservação genética dessa espécie, bem como em outros estudos que utilizam esta estimativa.

Este estudo foi realizado com poucos indivíduos e é preliminar, entretanto os dados sugerem que fluxo gênico estimado entre as populações foi pequeno e deve ser considerado restrito, provavelmente como conseqüência da antropização das regiões onde se encontram as populações. O fluxo é restrito e provavelmente a deriva genética vêm atuando nestas populações.

Os alelos exclusivos estimados indicam que 22,5% dos alelos se encontram restritos a uma população. Vale ressaltar que estes alelos exclusivos foram estimados a partir de uma amostragem pequena. SLATKIN (1985) descreve uma metodologia para avaliação do fluxo gênico a partir de alelos raros (ou exclusivos). A distribuição das freqüências destes alelos raros (alelos que aparecem em uma única população) é utilizada para estimar a média do número de migrantes permutados entre populações locais. O logaritimo de Nm é aproximadamente relacionado de forma linear com o logaritimo da média de frequência de alelos exclusivos (SLATKIN, 1985; SLATKIN & BARTON, 1989). O método de SLATKIN (1985), considera a relação:

$Ln[p(1)] = a\ ln(Nm) + b$, onde $p(1)$ é a frequência média dos alelos encontrados em apenas uma população, e a e b são constantes que dependem do número de indivíduos amostrados por população. Segundo os autores, a deriva genética resultará em diferenciação populacional se $Nm \leq 1$, mas não resultará, se $Nm \geq 1$.

Neste estudo, dos nove alelos exclusivos, três se encontram na população cinco (Dourado) que está em áreas de pastagem e bordadura de fragmentos, e três alelos exclusivos estão na população nove (Ibituruna – Lagoa) que é a mais endogâmica de todas. As características das áreas em que ocorrem estas populações pode explicar a ocorrência destes alelos, pois nestas situações pode haver uma tendência a diminuição de fluxo gênico.

O dendrograma apresenta agrupamentos coerentes com a ecologia e a geografia dessas populações. As populações da região I e II (Campinas, Serra Negra, Casa Branca e Mococa) formaram um grupo menos divergente devido provavelmente a questões ecológicas, pois os ambientes apresentam clima semelhante e relevo com ocorrência de Serras e Vales, sendo próximas geograficamente. As populações das localidades Dourado e Barbosa (região III) estão agrupadas, porém com maior distância do grupo mencionado anteriormente e isto também pode ser devido a ecologia da região onde predomina o cerrado, grandes pastagens e clima semelhante. A população de São Pedro do Turvo (região III) ficou separada das demais populações, foi a mais divergente, e a provável explicação para essa distância genética (0,5651) é que esses indivíduos podem ser outro ecótipo ou pertencer à outra espécie do gênero – *Acrocomia hassleri*. A identificação visual entre as duas espécies que compreendem esse gênero é difícil, considerando que a principal diferença morfológica entre elas é apenas o tamanho que apresentam. Os indivíduos da população de São Pedro do Turvo apresentam um porte menor e a diferença no tamanho dos frutos é marcante (Figura 23), além do mais estão numa região de cerrado correspondendo às características descritas da espécie *A. hassleri*.

Figura 23 – Comparação entre os frutos coletados, mostrando a diferença de tamanho dos frutos das populações de Ibituruna – Lagoa, Serra Negra e São Pedro do Turvo.

As populações oito e nove (Ibituruna - Lagoa e Gupiara) estão próximas geograficamente e formaram um grupo distinto do grupo formado pelas demais populações (Campinas, Serra Negra, Casa Branca, Mococa, Dourado e Barbosa) e a provável causa desta distância genética é a ocorrência de um obstáculo geográfico natural, a Serra da Mantiqueira que delimita os Estados de São Paulo e Minas Gerais e conseqüentemente as populações, impedindo a troca de alelos entre estas causando a divergência genética visualizada no dendrograma.

6 CONCLUSÕES

Os resultados permitem concluir que:

a) A biblioteca foi eficiente na identificação de locos microssatélites;
b) Os dados sugerem que a espécie tem sistema reprodutivo misto;

c) Existe uma elevada diversidade entre populações de *Acrocomia aculeata*;
d) A diversidade genética é maior entre populações dentro de regiões do que entre regiões, de modo que a coleta dentro de uma região pode ser mais vantajosa que uma coleta em todas as regiões;
e) Sugerem também, um comprometimento da estrutura metapopulacional, indicando que as populações estudadas sofreram deriva genética devido a antropização e têm fluxo gênico restrito. Estes resultados devem ser considerados para a conservação e o pré-melhoramento desta espécie. Porém, é importante ressaltar que é necessário um número maior de indivíduos para confirmar os resultados aqui obtidos;
f) A menor endogamia observada nas macaúbas das regiões de Campinas, Serra Negra, Mococa e Casa Branca pode ser explicada pela maior ocorrência da espécie nessas regiões que são menos antropizadas em oposição àquelas da região do planalto paulista, cujas populações encontram-se restritas a fragmentos isolados;
g) Estes dados são preliminares e um estudo mais completo deve ser realizado, abrangendo maior número de indivíduos, populações e locos.

7 REFERÊNCIAS BIBLIOGRÁFICAS

ALFENAS, A. C. Eletroforese de isoenzimas e proteínas afins: fundamentos e aplicações em plantas e microrganismos. Viçosa: UFV, 1998. 574p.

ALVES, M. R. P. Palmeiras: características botânicas e evolução. Campinas, SP: Fundação Cargil, 1987.

ALVES, R. M. Caracterização genética de populações de cupuaçuzeiro, *Theobroma grandiflorum* (Willd. ex. Spreng.) Shum., por marcadores microssatélites e descritores botânico-agronômicos. 2002, 159p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura "Luiz de Queiroz" da Universidade de São Paulo, Piracicaba. 2002.

AVELAR, A. Criei meus sete filhos com o óleo de macaúba. http://www.revista encontro. com.br/junho05/rural/personagem.asp, (12 agosto 2006).

BATISTA, A. C. F. Biodiesel no tanque, http://www.ambientebrasil.com.br/composer. php3?base=./energia/index.html&conteudo=./energia/artigos/oleo_vegetal.html, (17 julho 2006).

BILLOTTE, N.; COUVERUR, T.; MAESEILAC, N.; BROTTIER, P.; PERTHUIS, B.; VALLEJO, M.; NOYER, J. L.; JACQUEMOUD-COLLET, J. P.; RISTERUCCI, A. M.; PINTAUD, J. C. A new set of microsatellite markers for the peach palm (*Bactris gasipaes* Kunth); characterization and across-taxa utility within the tribe Cocoeae. **Molecular Ecology**, v.4, p.580-582, 2004.

BILLOTTE, N.; LAGODA, P. J. L.; RISTERUCCI, A. M.; BAURENS, F. C. Microssatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. **Fruits**, v.54, p.277-288, 1999

BIODIESEL, O Biodiesel, http://www.biodiesel.gov.br/, (15 agosto 2006).

BIODIESELBR, Agricultura familiar, emprego e o lado social do biodiesel, http://www.biodieselbr. com/biodiesel/social/aspectos-sociais.htm, (15 setembro 2006).

BORÉM, A. & CAIXETA, E. T. Marcadores moleculares. Viçosa – M.G., 374p., 2006.

BRONDANI, R.P.V.; BRONDANI, C.; TARCHINI, R.; GRATTAPAGLIA, D. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophyla*. **Theoretical and Applied Genetics**, v.97, p.816-827, 1998.

BROOKES, A. J. The essence of SNPs. **Gene**, v.234(2), p.177-186, 1999.

CARNEIRO, N. M.C. A substituição de combustível fóssil por biodiesel e o mercado de carbono, http://www.rbb.ba.gov.br/index.php?id=319&prefixo=det&menu=biblioteca, (28 outubro 2006).

CARNEIRO, R. A. F. A produção de Biodiesel na Bahia, http://www.sei.ba.gov.br/ publicacoes/ bahia_analise/conj_planejamento/pdf/c&p112/05.pdf, (15 fevereiro 2005).

CASTELO, A. T.; MARTINS, W.; GAO, G. R. TROLL – Tandem Repeat Occurence Locator. **Bioinformatics**, v.18 (4), p.634-636, 2002.

CESAR, E. Biodiesel no tanque. **Pesquisa Fapesp**, São Paulo, 94 ed., dez. 2003.

CHAMBERS, J. K. & MACAVOY, E. S. Microssatellete: Consensus and controversy. **Comparative Biochemistry and Physiology**, v.126, p.455-476, 2000.

CIPRIANO, R. Descobrindo o poder da macaúba. http://www.embrapa.br/noticias /banco_de_noticias/folder.2006/foldernoticia.2006-04-03.3722359657/noticia.2006-04-26.0919750710/mostra_noticia, (15 setembro 2006).

COCKERHAM, C. C. Variance of gene frequencies. **Evolution**, v.23, p.72-84, 1969.

CONDIT, R. & HUBBELL, S. P. Abundance and DNA frequence of 2-base repeat regions in tropical tree genomes. **Genome**, v.34, p.66-71, 1991.

COSTA NETO, Pedro R., ROSSI, Luciano F. S., ZAGONEL, Giuliano F., RAMOS, Luis P. Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em frituras. **Química Nova**, vol.23, n.4, p.531-537, jul./ago. 2000.

COLLEVATTI, R.G.; BRONDANI, R.V.; GRATTAPAGLIA, D. Development and characterization of microsatellite markers for genetic analysis of a Brazilian endangered tree species *Caryocar brasiliense*. **Heredity**, v.83, n.6, p.748-756, 1999.

DAYANANDAN, S.; BAWA, K. S.; KESSELI, R. et al. Conservation of microsatellites among tropical tree (Leguminosae). **American Journal of Botany**, v.84, n.12, p.1658-1663, 1997.

DISCOVER LIVE – *Acrocomia aculeata*, http://pick4.pick.uga.edu/mp/20m?kind=Acrocomia+aculeata, (10 agosto 2006).

DOYLE, J. J. & DOYLE, J. L. Isolation of plant DNA fresh tissue, **Focus**, v.12, p.13-15, 1990.

FALCONER, D. S. & MACKAY, T. F. C. Introdution to quantitative genetics. New York: Longmam Scientific & Technicals, 1989. 438p.

FERNADES, S. O. Macaúba, matéria-prima alternativa, http://www.iea.sp.gov.br/ OUT/verTexto.php?codTexto=7358, (5 novembro 2006).

FERREIRA, M. E., GRATTAPAGLIA, D. Introducción al Uso de Marcadores moleculares en el Análisis Genético. Brasilia: EMBRAPA-CENARGEN, 220p., 1998.

FIELD, D. & WILLS, C. Long, polymorphic microsatellites in simple organisms. **Proceedings: Biological Sciences**, v.263, n.1367, p.209-215, 1996.

FRUITS FROM AMERICA - An ethnobotanical inventory, http://www.ciat.cgiar.org /ipgri /fruits_from_americas/frutales/Ficha%20Acrocomia%20aculeata.htm, (17 março 2005).

FUTUYMA, D. J. Evolutionary Biology. 3. ed., Sunderland, MA: Sinauer Associates, 1998. 196p.

FUTUYMA, D. J. Biologia evolutiva. 2.ed. Ribeirão Preto: Sociedade Brasileira de Genética, 1993. 631p.

GAIOTTO, F. A.; GRATTAPAGLIA, D.; VENCOVSKY, R. Microssatellite markers for Herat of palm – *Euterpe edulis* and *E. oleracea* Mart. (Arecaceae). **Molecular Ecology**, v.1, p.86-88, 2001.

GOODMAN, S. J. Rst Calc: a colletion of computer programs for calculating estimates of genetic differentition from microsatellite data amd a determining their significance. **Molecular Ecology**, v.6, p.881-885, 1997.

GOVINDAJARU, R. D. Variation in gene flow levels among predominantly self-pollinated plants. **Journal of Evolutionary Biology**, v.2, p.173-181, 1989.

GRIN – Germplasm Resources Information Network, http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?1388#uses, (2 outubro 2006).

GRUPO VERDE – *Acrocomia aculeata*, http://www.grupoverdepalms.com/index2. html, (02 abril 2006).

HAMADA, H.; PETRINO, M. G.; KAKUNAGA, T. A novel repeated element with Z-DNA forming potential is widely found in evolutionarily diverse eukayotic genomes. **Proceedings of National Academy of Sciences of United States of America**, v.79, p. 6465-6469, 1982.
HAMRICK, J. L. Isozymes and analysis of genetic structure in plants populations. In: Soltis, D. E.; Soltis, P. S. (ed.). London: Chapman and Hall, p.335-348, 1989.

HAMRICK, J. L. Plant population genetics and evolution. **American Journal of Botany**, v.69, n.10, p.1685-1693, 1982.

HELENTJARIS, T.; SLOCUM, M.; WRIGHT, S.; SCHAEFER, A.; NIENHUIS, J. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. **Theoretical and Applied Genetics**, v. 72, p.761-769. 1986.

HENDERSON, A.; GALEANO, G.; BERNAL, R. Field Guide to the Palms of the Americas. New Jersey: Princepton University, p.166-167., 1995.

HUANG, X. & MADAN, A. Cap3: a DNA sequence assembly program. **Genome Research**, v.9 (9), p.868-877, 1999.

KIJAS, J. M.; FOWLER, J. C.; GARBETT, C. A.; THOMAS, M. R. Enrichment of microsatellites from the *Citrus* genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *BioTechniques*, v.16, p.656-662, 1994.

KWOK, P. Y. Methods for genotyping single nucleotide polymorphisms. **Annual** Review of Genomics and Human Genetics, v.2, p.235-258, 2001.

LEWIS, P. & ZAYKIN, D. Genetic data analysis: computer program for the analysis of allelic data (software) 2000. version 1.0(d12), http://alleyn.eeb.uconn.edu/gda/ (21 novembro 2006).

LIMA, Edson S., FELFILI, Jeanine M., MARIMON, Beatriz S., SCARIOT, Aldicir. Diversidade, estrutura e distribuição espacial de palmeiras em um cerrado *sensu stricto* no Brasil Central. **Revista Brasileira de Botânica**, São Paulo, v.26, n.3, p.361-370, 2003.

LORENZI, G. M. A. C., *Acrocomia aculeata* (Jacq.) Lodd. ex Mart. – Arecaceae: bases para o extrativismo sustentável. Curitiba, 2006, 156f. Tese (Doutorado em Ciências Agrárias) Universidade Federal do Paraná, Curitiba, 2006.

LORENZI, H. Palmeiras do Brasil: exóticas e nativas. Nova Odessa: Editora Plantarum, p.1-20, 1996.

LORENZI, H. Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. Nova Odessa, SP: Plantarum, 1992.

MARKERT, C.L. & MOLLER, F. Multiple forms of enzymes: issue, ontogenetic and species specific patterns. **Proc Natl Acad Sci** USA, v45, p.753-763, 1959.

MATIOLI, S. R. & PASSOS-BUENO, M. R. S. Métodos baseados em PCR para análise de polimorfismos de ácidos nucléicos. In: Matioli, S. R. (ed). Biologia Molecular e Evolução. Ribeirão Preto: Holos, p. 153-161, 2001.

MDA – MINISTÉRIO DO DESENVOLVIMENTO AGRÁRIO. Mais 100 agricultores do Pará iniciam produção de matéria-prima para biodiesel, http://www.mda.gov.br /index.php?ctuid=10638&sccid=134, (3 novembro 2006).

MEIRELLES, F. S. Biodiesel, http://www.faespsenar.com.br/faesp/economico/ EstArtigos/biodiesel.pdf?, (13 novembro.2005).

MILLER, M. Tools for population genetic analyses (TFPGA) (software), version 1.3: a windows program for analyses of allozyme and molecular population genetic data, 1997, http://herb.bio.nau.edu/~miller/tfpga.htm, (20 novembro 2006).

MISSIAGLIA, A. A. Mapeamento genético de QTL para qualidade da madeira e florescimento precoce e estudos de expressão gênica alelo específica em *Eucaliptus* spp. Piracicaba-SP. 236 p. Tese (Doutorado) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo. 2005.

MISSIAGLIA, A. & GRATTAPAGLIA, D. Plant microssatellite genotyping with 4-color fluorescent detection using multiple-tailed primers. **Genetics and Molecular Research**, v.5 (1), p.72-78, 2006.

MOBOT - MISSOURI BOTANICAL GARDEN - Manual de Plantas de Costa Rica, http://www.mobot.org/MOBOT/research/treat/acoelorraphe.shtml, (15 junho 2006).

MORCOTE-RIOS, G. & BERNAL, R. Remains of palms (Palmae) at archaeological sites in the New Word: a review. **The Botanical Review**, New York, v.67, n.3, p.309-350, 2001.

MORGANTE, M.; HANAFEY, M.; POWELL,W. Microssatellites are preferentially associated with nonrepetitive DNA in plant genomes. **Nature Genetics**, v.30, p.194-200, 2002.

MORGANTE, M. & OLIVIERI, A. M. PCR-amplified microssatellites asmarkers in plant genetics. **The Plant J.**, v.3, p.175-182, 1993.

MOTTA, P. E. F.; CURI, N.; OLIVEIRA-FILHO, A. T.; GOMES, J. B. V. Occurrence of macaúba in Minas Gerais, Brazil: relationship with climatic, pedological and vegetation attributes. **Pesq. agropec. bras.,** Brasília, v.37, n.7, p.1023-1031, jul. 2002.

MULLIS, K. & FALOONA, F. Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. **Methods Enzymol**. V.55, p.335-350. 1987.

NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. **Genetics**, v.89, n.3, p.583-590, 1978.

NEI, M. F-statistics and analysis of gene diversity in subdivided population. **Annals of Human Genetics**, v.41, p.225-33, 1977.

NEW WORLD FRUITS – Database, http://www.ipgri.cgiar.org/Regions/Americas/programmes/TropicalFruits/qryall3.asp?intIDSpecies=154, (24 outubro, 2006).

NOBELPRIZE – Nobelprize.org, http://nobelprize.org/nobel_prizes/chemistry/laureates/1993/index.html, (15 julho 2006).

NOVAES, R. F. Contribuição para o estudo do coco macaúba. Piracicaba, 1952, 85f. Tese (Doutorado em Ciências Agrárias) – Escola Superior de Agricultura "Luiz de Queiroz" da Universidade de São Paulo, Piracicaba. 1952.

OLIVEIRA, F. A. M. A produção de óleos vegetais de macaúba e seus co-produtos na região metropolitana de Belo Horizonte. In: 3º Congresso Brasileiro de Plantas Oleaginosas, óleos, Gorduras e Biodiesel, Varginha, 2006. CD-Rom.

PANORAMA BRASIL – Óleo de macaúba disputa mercado com o de soja. http://www.todafruta.com.br/todafruta/mostra_conteudo.asp?conteudo=3660, (15 outubro 2006).

PARAMN, I. & MICHELMORE, R. W. Development of reliable PCR-based markers linkage to downy mildew resistance genes in lettuce. **Theoretical and Applied Genetics**, v. 85, p.985-993. 1993.

PINTO, L. R. Análise da estrutura genética das populações de milho (*Zea mays* L.) BR-105 e BR-106 e respectivos sintéticos IG-3 e IG-4 por meio de microssátelites, Piracicaba, 2001. 142p. (Tese Doutorado) – Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 2001.

PRESTES, S. Petrobras investe em parceria com cooperativas de pequenos agricultores gaúchos, http://www.biodiesel.gov.br/, (3 novembro 2006).

RATTER, J. A.; BRIDGEWATER, S.; ATKINSON, R.; RIBEIRO, J. F. Analysis of the floristic composition of the Brazilian cerrado vegetation. **Edinburgh Journal of Botany**, Edinburgh, v.53, n.2, p.153-180, 1996.

READ, R. W. New chromosome counts in the Palmae. **Principes**, v.10, p.55-61, 1966.

ROLF, F. J. NTSYS-Pc: Numerical taxonomy and multivariate analysis system. New York: Exeter Publisher, 1989. 210p.

ROMANO, E. & BRASILEIRO, A. C. M. Extração de DNA de plantas, http://www.uefs.br/disciplinas/bot859/extra.pdf , (20 julho 2005).

SCARIOT, A.; LLERAS, E.; HAY, J. D. Flowrering and fruiting phenologies of the palm *Acrocomia aculeata*: patterns and consequences. **Biotropica**, Washington, v.27, n. 2, p. 168-173, 1995.

SCARIOT, A.; LLERAS, E.; HAY, J. D. Reproductive biology of the palm *Acrocomia aculeata* in Central Brazil. **Biotropica**, Washington, v.23, n. 1, p. 12-22, 1991.

SCHLÖTTERER, C.; RITTER, R.; HARR, B.; BREM, G. High mutation rate of a long microssatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates. **Molecular and Biology Evolution**. v.15, n.10, p.1269-1274, 1998.

SCHLÖTTERER, C. & TAUTZ, D. Slippage synthesis of simple sequence DNA. **Nucleic Acids Research**, v.20, n.2, p.211-215, 1992.

SEOANE, C. E. C.; SEBBENN, A. M.; KAGEYAMA, P. Y. Sistema de reprodução em populações de *Esenbeckia leiocarpa* Engl. **Revista do Instituto Florestal**, v.13 (1), p.19-26, 2001.

SHARMA, A. K. & SARKAR, S. K. Cytology of different species of palms and its bearing on the solution of the problems of phylogeny and speciation. **Genetica**, v.28, p.361-488, 1956.

SILVA, J. C.; BARRICHELO, L. E. G.; BRITO, J. O. Endocarpos de babaçu e macaúba comparados à madeira de *Eucalyptus grandis* para produção de carvão vegetal. http://www.ipef.br/publicacoes/ scientia/nr34/cap04.pdf, (3 setembro 2006).

SLATKIN, M. A measure of population subdivision based on microsatellite allele frequencies. **Genetics**, v.130, p.457-462, 1995.

SLATKIN, M. & BARTON, N. H. A comparision of three indirect methods for estimating average levels of gene flow. **Evolution**, v.43, n.7, p.1349-1368, 1989.

SOLTIS, D. E. & SOLTIS, P. S. Isozymes in plant biology. Dioscorides. Portland, OR. 1989.

SOUZA, A. P. Biologia molecular aplicada ao melhoramento. In: NASS, L. L., VALOIS, A. C. C., MELO, I. S., VALADARES-INGLIS, M. C. (ed.). Recursos genéticos & melhoramento: plantas, Rondonópolis: Fundação M.T., p.939-966, 2001.

STRAND, M.; PROLLA, T. A.; LISKAY, R. M.; PETES, T. D. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. **Nature**, v.365, p.274-276, 1993.

SUGANUMA, E. & CIAMPI, A. Y. Análise genética populacional de jatobá (*Hymenaea* ssp Leguminosaea) utilizando microssatélites, http://www.redbio.org/portal/encuentros/enc_2001/posters/03/03pdf/03-007.pdf., (13 novembro. 2005).

TASSARO, H. Frutas no Brasil. São Paulo: Empresa das Artes, 1996.

TAUTZ, D. & RENZ, M. Simple sequences are ubiquitous repetitive componets of eukaryotic genomes. **Nucleic Acids Research**, v.12, n.10, p. 4127-4138, 1984.

THOMPSON, J. D.; PLEWNIAK, F.; POCH, O. A comprehensive comparison of multiple sequence alignment programs. **Nucleic Acids Res.**, v.27, p.2682-2690, 1999.

TOTH, G.; GASPARI, Z.; JURKA, J. Microssatellites in different eukaryotic genomes: survey and analysis. **Genome Research**. p.967-981, 2000.

VANTOAL, T. T.; PENG, J.; MARTINS, S. S. Using AFLP markers to determinate the contribution of parental genomes during recurrent selection. **Soybean Genetics Newsletter**, v.23, p.214-216, 1996.

VENCOVSKY, R. Análise de variância de freqüências alélicas. **Genetics and Molecular Biology**, v.15, p.53-60, 1992.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRITERS, A.; POT, J.; PALEMAN, J.; KUIPER, M.; ZABEAU, M. AFLP: A new technique for DNA fingerprinting. **Nucleic Acids Research**, v.23, p.4407-4414. 1995.

WEIR, B.S. **Genetics data analysis II**: methods for discrete population genetic data. Suderland, MA: Sinauer Associates, 1996. 455p.

WILLIAMS, J. G.; KUBELIK, A. R.; LIVAK, K. J.; RAFALSKI, L. A.; TINGEY, S. V. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research, v.18, p.6531-6535. 1990.

WRIGHT, S. The interpretation of population structure by F-statistic with special regard to systems of mating. **Evolution**, v.19, p.395-420, 1965.

WRITE, G. & POWELL, W. Cross-species amplification of SSR loci in the Meliaceae family. **Molecular Ecology**, v.6, p.1195-1197, 1997.

XU, M. L.; HUARACHA, E. M.; GASIC, K.; PAUWELS, E.; VAN DEN PUTTE, A.; KEULEMANS, J. W.; KORBAN, S. S. Phenotypic Reaction and Genetic Analysis Using AFLP-derived SCARs for Resistance to Apple Scab. **Journal of Phytopathology**, v.152 (5), p.260, 2004.

ZABEAU, M. Selective restriction fragment amplification: a general method for DNA fingerprinting. **European Patent Application,** nº 0534858 A1. 1993.

ZANE, L.; BARGRLLONI, L.; PATARNELLO, T. Strategies for microssatellites isolation: a review. **Molecular Ecology**, v.11, p.1-6, 2002.

ANEXOS

Anexo 1 – Arquivo de entrada do programa GDA

```
#nexus
begin GDADATA; [Marcadores SSR em macauba]
      dimensions npops=9 nloci=8;
      format missing=? separator=/;
      locusallelelabels
            1 Ac07,
            2 Ac10,
            3 Ac12,
            4 Ac18,
            5 Ac26,
            6 Ac30,
             7 Ac32,
            8 Ac35;
      matrix
Pop_1:
1  169/169  182/182  200/200  173/193  314/314  148/148    ?/?    410/410
2  169/169  182/182  200/200  173/193  278/278  148/148  233/233  410/430
3    ?/?    182/182  198/200  193/193  278/316  148/148  233/233  410/410
4  169/169  182/182  200/200  193/193  278/314  148/158  233/233  410/430
5  153/169  182/182  200/200  173/173  278/278  148/148  233/233  410/410
6  169/169  182/182  200/200  193/193  278/278  150/150  233/233  410/410
7  153/169  182/182  200/200  193/193  300/314  148/148    ?/?    410/410,
Pop_2:
1    ?/?    182/182    ?/?    173/193  314/314  148/148    ?/?    388/418
2  169/169  182/182  200/200  193/193  314/314    ?/?    233/233    ?/?
3  169/169  182/182  200/200  193/193  297/297    ?/?    233/233  410/410
4  161/169  182/182  198/200  173/193  278/314  148/148  233/233  410/410
5  161/169  182/182  194/198  173/193  278/314  148/156    ?/?    410/410,
Pop_3:
1  161/165  182/182    ?/?    173/193  278/278  148/156  233/233  388/388
2  153/169  182/182    ?/?    173/193  278/278  148/148  233/233  410/410
3  153/161  182/186    ?/?    173/193  314/314  148/148    ?/?      ?/?
4  161/169  182/182    ?/?    173/193  278/278  148/156  233/233  430/430
5  161/161  182/182    ?/?    173/193  278/278  148/148  233/233    ?/?  ,
Pop_4:
1    ?/?      ?/?      ?/?    173/193    ?/?    148/148  233/233  388/388
2    ?/?    182/182    ?/?      ?/?      ?/?      ?/?    229/233  410/410
3  161/169  182/186    ?/?    193/193    ?/?    148/158  233/233  410/410
4  169/169  186/186    ?/?    173/193  278/314  148/156  229/233  430/430
5  169/169  182/182  198/200  173/193  278/278  148/148  233/233    ?/? ,
Pop_5:
1  166/169  182/182  200/200  173/193  278/278  148/148  233/233  418/418
2  153/161  182/182  200/200  173/193  297/297  148/148    ?/?      ?/?
3  169/169  182/182  200/200  193/193  278/278    ?/?    233/233  418/418
4  166/166  184/184  200/202  173/193  278/278  148/148  233/233  418/418
5  177/177  184/184  200/202  173/193  281/300    ?/?    233/233    ?/?  ,
Pop_6:
1  165/165  182/182  190/190  173/173  297/297  148/148    ?/?      ?/?
2  165/165  176/176  192/200  173/193  281/297    ?/?      ?/?    388/388
3  153/166  168/182  198/198  173/193  281/297    ?/?    233/233  388/388
4  165/165  182/182  200/200  193/193  281/297  148/148    ?/?    430/430
5  165/165    ?/?      ?/?    193/193  285/285  148/148  233/233    ?/?
6  166/166  182/182  194/200  193/193  285/297    ?/?    233/233    ?/?  ,
Pop_7:
1  166/166  176/176  192/192  193/193  281/297  150/150  233/233  410/418
2  153/153  182/182  192/200  193/193  276/285    ?/?    233/233  410/410
3  153/153  182/182  190/190  173/173  285/300    ?/?    233/233  418/430
4  153/166  182/182    ?/?    193/193  281/281    ?/?    233/233  418/418,
```

```
Pop_8:
1  153/169  186/186    ?/?    193/193  278/278    ?/?    233/233    ?/?
2  166/169  182/186    ?/?    193/193  278/278    ?/?    233/233    ?/?
3  166/166  186/186    ?/?    173/193    ?/?    156/156  233/233    ?/?
4  169/169  186/186    ?/?    173/193  278/278  148/148  233/233    ?/?
5  166/169  186/186    ?/?    193/193  278/278    ?/?    233/233    ?/?
6  153/166  186/186    ?/?    173/193  278/278    ?/?      ?/?      ?/?  ,
Pop_9:
1  169/169  186/186    ?/?    173/193  276/276  148/148  233/233  388/410
2  169/169  186/186    ?/?    173/193  281/281  156/156  233/233  380/388
3  169/169  186/186  212/212  173/193  276/276  156/156  233/233  408/408
4  169/169  186/186  200/200  193/193  276/276  156/156  233/233    ?/?  ;
end;
begin TREE;
        tree
geral=(Pop_1,Pop_2,Pop_3,Pop_4,Pop_5,Pop_6,Pop_7,Pop_8,Pop_9,Pop_10);
end;
```

Anexo 2 – Arquivo de entrada do programa TFPGA

```
1, 55, 33, 55, 12, 77, 11, 00, 44
1, 55, 33, 55, 12, 22, 11, 22, 46
1, 00, 33, 45, 22, 28, 11, 22, 44
1, 55, 33, 55, 22, 27, 14, 22, 44
1, 15, 33, 55, 11, 22, 11, 22, 44
1, 55, 33, 55, 22, 22, 22, 22, 44
1, 15, 33, 55, 22, 67, 11, 00, 44
2, 00, 33, 00, 12, 77, 11, 00, 25
2, 55, 33, 55, 22, 77, 00, 22, 00
2, 55, 33, 55, 22, 55, 00, 22, 44
2, 25, 33, 45, 12, 27, 11, 22, 44
2, 25, 33, 34, 12, 27, 13, 00, 44
3, 23, 33, 00, 12, 22, 13, 22, 22
3, 15, 33, 00, 12, 22, 11, 22, 44
3, 12, 35, 00, 12, 77, 11, 00, 00
3, 25, 33, 00, 12, 22, 13, 22, 55
3, 22, 33, 00, 12, 22, 11, 22, 00
4, 00, 00, 00, 12, 00, 11, 22, 22
4, 00, 33, 00, 00, 00, 00, 12, 44
4, 25, 35, 00, 22, 00, 14, 22, 44
4, 55, 55, 00, 12, 27, 13, 12, 55
4, 55, 33, 45, 12, 22, 11, 22, 00
5, 45, 33, 55, 12, 22, 11, 22, 55
5, 12, 33, 55, 12, 55, 11, 00, 00
5, 55, 33, 55, 22, 22, 00, 22, 55
5, 44, 44, 56, 12, 22, 11, 22, 55
5, 66, 44, 56, 12, 36, 00, 22, 00
6, 33, 33, 11, 11, 55, 11, 00, 00
6, 33, 22, 25, 12, 35, 00, 00, 22
6, 14, 13, 44, 12, 35, 00, 22, 22
6, 33, 33, 55, 22, 35, 11, 00, 55
6, 33, 00, 00, 22, 44, 11, 22, 00
6, 44, 33, 35, 22, 45, 00, 22, 00
7, 44, 22, 22, 22, 35, 22, 22, 45
7, 11, 33, 25, 22, 14, 00, 22, 44
7, 11, 33, 11, 11, 46, 00, 22, 56
7, 14, 33, 00, 22, 33, 00, 22, 55
8, 15, 55, 00, 22, 22, 00, 22, 00
8, 45, 35, 00, 22, 22, 00, 22, 00
8, 44, 55, 00, 12, 00, 33, 22, 00
8, 55, 55, 00, 12, 22, 11, 22, 00
8, 45, 55, 00, 22, 22, 00, 22, 00
8, 14, 55, 00, 12, 22, 00, 00, 00
9, 55, 55, 00, 12, 11, 11, 22, 24
9, 55, 55, 00, 12, 33, 33, 22, 12
9, 55, 55, 77, 12, 11, 33, 22, 33
9, 55, 55, 55, 22, 11, 33, 22, 00
0, 00, 00, 00, 00, 00, 00, 00, 00
```

Anexo 3 – Arquivo de entrada do programa FSTAT

```
9 8 430 3
Ac07
Ac10
Ac12
Ac18
Ac26
Ac30
Ac32
Ac35
1    169169    182182    200200    173193    314314    148148      0       410410
1    169169    182182    200200    173193    278278    148148    233233    410430
1      0       182182    198200    193193    278316    148148    233233    410410
1    169169    182182    200200    193193    278314    148158    233233    410430
1    153169    182182    200200    173173    278278    148148    233233    410410
1    169169    182182    200200    193193    278278    150150    233233    410410
1    153169    182182    200200    193193    300314    148148      0       410410
2      0       182182      0       173193    314314    148148      0       388418
2    169169    182182    200200    193193    314314      0       233233      0
2    169169    182182    200200    193193    297297      0       233233    410410
2    161169    182182    198200    173193    278314    148148    233233    410410
2    161169    182182    194198    173193    278314    148156      0       410410
3    161165    182182      0       173193    278278    148156    233233    388388
3    153169    182182      0       173193    278278    148148    233233    410410
3    153161    182186      0       173193    314314    148148      0         0
3    161169    182182      0       173193    278278    148156    233233    430430
3    161161    182182      0       173193    278278    148148    233233      0
4      0         0         0       173193      0       148148    233233    388388
4      0       182182      0         0         0         0       229233    410410
4    161169    182186      0       193193      0       148158    233233    410410
4    169169    186186      0       173193    278314    148156    229233    430430
4    169169    182182    198200    173193    278278    148148    233233      0
5    166169    182182    200200    173193    278278    148148    233233    418418
5    153161    182182    200200    173193    297297    148148      0         0
5    169169    182182    200200    193193    278278      0       233233    418418
5    166166    184184    200202    173193    278278    148148    233233    418418
5    177177    184184    200202    173193    281300      0       233233      0
6    165165    182182    190190    173173    297297    148148      0         0
6    165165    176176    192200    173193    281297      0         0       388388
6    153166    168182    198198    173193    281297      0       233233    388388
6    165165    182182    200200    193193    281297    148148      0       430430
6    165165      0         0       193193    285285    148148    233233      0
6    166166    182182    194200    193193    285297      0       233233      0
7    166166    176176    192192    193193    281297    150150    233233    410418
7    153153    182182    192200    193193    276285      0       233233    410410
7    153153    182182    190190    173173    285300      0       233233    418430
7    153166    182182      0       193193    281281      0       233233    418418
8    153169    186186      0       193193    278278      0       233233      0
8    166169    182186      0       193193    278278      0       233233      0
8    166166    186186      0       173193      0       156156    233233      0
8    169169    186186      0       173193    278278    148148    233233      0
8    166169    186186      0       193193    278278      0       233233      0
8    153166    186186      0       173193    278278      0         0         0
9    169169    186186      0       173193    276276    148148    233233    388410
9    169169    186186      0       173193    281281    156156    233233    380388
9    169169    186186    212212    173193    276276    156156    233233    408408
9    169169    186186    200200    193193    276276    156156    233233      0
```

Anexo 4 - Arquivo de entrada do programa NTSYS

```
2 9L 9 0
Campinas  SerraNegra   Mococa   CasaBranca   Dourado   Barbosa  SaoPedrodoTurvo IbiturunaG IbiturunaL
 0.0000
 0.0102    0.0000
 0.1088    0.0819    0.0000
-0.0275   -0.0482    0.0042    0.0000
 0.2503    0.2376    0.1084    0.1092    0.0000
 0.3972    0.2661    0.2097    0.2553    0.2195    0.0000
 0.4869    0.5084    0.4428    0.5681    0.4883    0.4233    0.0000
 0.3171    0.4223    0.2440    0.1039    0.2335    0.5537    0.6966    0.0000
 0.5407    0.4891    0.6311    0.2723    0.7093    0.7344    0.9061    0.2438    0.0000
```

Anexo 5 – Estatísticas F obtidas para cada alelo, calculadas pela análise de variância utilizando oito locos microssatélites em nove populações

```
There are 8 active loci
Two-level analysis
  9 active populations
```

Locus	Allele	f	F	Theta-P
Ac07	All	0.247486	0.477826	0.306095
	177	1.000000	1.000000	-0.013019
	166	0.442763	0.502108	0.106499
	165	0.771608	0.884318	0.493494
	161	-0.106805	0.162737	0.243532
	153	0.073780	0.206740	0.143552
	169	0.081119	0.515708	0.472955
Ac10	All	0.662862	0.839416	0.523683
	168	0.000000	0.000000	0.000000
	176	1.000000	1.000000	0.009627
	184	1.000000	1.000000	0.250836
	186	0.292383	0.840984	0.775280
	182	0.655605	0.818469	0.472899
Ac12	All	0.366301	0.447280	0.127788
	212	1.000000	1.000000	0.360341
	192	0.307958	0.495421	0.270883
	190	1.000000	1.000000	-0.034807
	202	-0.104972	-0.006391	0.089217
	194	0.022602	-0.027149	-0.050902
	198	0.261116	0.268698	0.010260
	200	0.268392	0.396731	0.175420
Ac18	All	-0.143959	-0.212720	-0.060108
	193	-0.143959	-0.212720	-0.060108
	173	-0.143959	-0.212720	-0.060108
Ac26	All	0.414494	0.588160	0.296609
	276	0.651729	0.856651	0.588399
	285	0.309645	0.379375	0.101006

	281	0.341277	0.394386	0.080624
	297	0.396575	0.496134	0.164990
	300	0.007334	-0.027982	-0.035577
	316	0.040323	-0.004615	-0.046826
	278	0.572841	0.778735	0.482007
	314	0.441904	0.564736	0.220091
Ac30	All	0.475019	0.597898	0.234064
	156	0.451946	0.612523	0.292995
	150	1.000000	1.000000	0.410904
	158	0.055718	-0.025765	-0.086292
	148	0.457324	0.556262	0.182315
Ac32	All	-0.115538	-0.005035	0.099058
	229	-0.115538	-0.005035	0.099058
	233	-0.115538	-0.005035	0.099058
Ac35	All	0.594604	0.699740	0.259340
	408	1.000000	1.000000	0.112540
	380	-0.053435	0.006210	0.056620
	418	0.105348	0.700974	0.665763
	388	0.628283	0.682588	0.146093
	430	0.650517	0.616158	-0.098313
	410	0.660435	0.754530	0.277104
Overall	---	0.349013	0.513173	0.252171

Bootstrapping over loci
  Number of replicates = 1000
  Nominal confidence interval = 95%
  Random number seed = 1 (specified)