UNIVERSIDADE FEDERAL DO PIAUÍ

Diversidade genética em sapucaia por meio de marcadores ISSR

Raimundo Carvalho Borges

Dissertação apresentada à Universidade Federal do Piauí como parte das exigências do Programa de Pós-graduação em Genética e Melhoramento para obtenção do título de "Mestre".

TERESINA

2015

Raimundo Carvalho Borges

Diversidade genética em sapucaia por meio de marcadores ISSR.

Orientador:
Prof. Dr. Sérgio Emílio dos Santos Valente
Coorientador :
Prof. Dr. Paulo Sarmanho da Costa Lima

Dissertação apresentada à Universidade Federal do Piauí como parte das exigências do Programa de Pós-graduação em Genética e Melhoramento para obtenção do título de "Mestre".

TERESINA
2015

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Biblioteca Setorial do Centro de Ciências Agrárias
Serviço de Processamento Técnico

B732d Borges, Raimundo Carvalho
Diversidade genética em sapucaia por meio de marcadores ISSR. / Raimundo Carvalho Borges - 2015.
67 f. : il.

Dissertação (Mestrado em Genética e Melhoramento)– Universidade Federal do Piauí,Teresina, 2015.
Orientação: Prof. Dr. Sérgio Emílio dos Santos Valente

1 . Banco de germoplasma 2. Castanheira 3. Diversidade genética 4.Marcadores ISSR. I. Título

CDD 583.42

UNIVERSIDADE FEDERAL DO PIAUÍ
PRÓ-REITORIA DE ENSINO DE PÓS-GRADUAÇÃO
COORDENADORIA GERAL DE PÓS-GRADUAÇÃO
NÚCLEO DE PÓS-GRADUAÇÃO EM CIÊNCIAS AGRÁRIAS
COORDENAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E MELHORAMENTO

ATA DE ARGUIÇÃO DA QUADRAGÉSIMA DISSERTAÇÃO DE MESTRADO APRESENTADA POR RAIMUNDO CARVALHO BORGES

Aos vinte e seis dias do mês de junho do ano de dois mil e quinze, às 14:00 horas, no Auditório do Núcleo de Pós-Graduação em Ciências Agrárias da Universidade Federal do Piauí, Campos Ministro Petrônio Portela, Teresina-PI, a Banca Examinadora infranomeada procedeu ao julgamento da Defesa de Dissertação intitulada **"Diversidade genética em sapucaia por meio de marcadores ISSR"** apresentada pelo mestrando **RAIMUNDO CARVALHO BORGES** para obtenção do Título de Mestre em Genética e Melhoramento, do Programa de Pós-Graduação em Genética e Melhoramento, desta Universidade. O Presidente da Banca Examinadora, **Dr. SERGIO EMILIO DOS SANTOS VALENTE** iniciando os trabalhos concedeu a palavra ao candidato **RAIMUNDO CARVALHO BORGES** para uma breve exposição do seu trabalho. Em seguida, o presidente concedeu a palavra, pela ordem e sucessivamente, aos examinadores, os quais passaram a argüir o candidato durante o prazo máximo de 30 (trinta) minutos, assegurando-se ao mesmo igual prazo para responder aos Senhores Examinadores. Ultimada a argüição, que se desenvolveu nos termos regimentais, a Banca, em sessão secreta, expressou seu julgamento, considerando.

Aprovado (AP) / Não aprovado (Nap)	Ap/	Nap
Dr. SERGIO EMILIO DOS SANTOS VALENTE (Presidente) UFPI	(X)	( )
Dr. FABIO BARROS BRITTO (Titular) UFPI	(X)	( )
Dr. PAULO SARMANHO DA COSTA LIMA (Titular) Embrapa Meio Norte	(X)	( )
Dr. DARIO ABEL PALMIERI (Titular) UNESP/Assis	(X)	( )

Em face do resultado obtido, a Banca Examinadora considerou **RAIMUNDO CARVALHO BORGES**...APROVADO.....

Nada mais havendo a tratar, o Presidente lavrou a presente ata que, após lida e achada conforme, foi por todos assinada.

Dr. SERGIO EMILIO DOS SANTOS VALENTE _____

Dr. FABIO BARROS BRITTO _____

Dr. PAULO SARMANHO DA COSTA LIMA _____

Dr. DARIO ABEL PALMIERI _____

À minha família, esposa e filhos

Dedico

AGRADECIMENTOS

Primeiramente, quero agradecer a Deus, por ter me garantido a vida e saúde para chegar ao final de mais um sonho que se tornou realidade.

À minha mãe, Alda Carvalho Borges, mulher forte e sonhadora. Por meio dela eu vim ao mundo, mas não foi só isso, ela acreditou nos seus sonhos e confiou em mim.

Ao meu pai, homem de poucas letras (em memória), que apesar da falta de oportunidades no seu tempo, trabalhou até os últimos dias da sua vida para proporcionar uma vida digna para todos os seus filhos.

Especialmente, à minha esposa Janaina, que durante esta jornada foi compreensiva e amorosa, cuidando de mim e dos nossos filhos. Agradeço, sobretudo, por ela ter acreditado e me dado forças para continuar.

Aos meus irmãos, em particular, minha irmã Francisca e seus filhos Diego (sobrinho e afilhado) e Vanessa, por ter me recebido em sua casa todas as semanas durante todo o curso.

Ao Dr. Sérgio Emílio dos Santos Valente, por ter sido meu orientador e professor. Sempre paciente e tranquilo, é um grande orientador do Programa de Pós-Graduação em Genética e Melhoramento.

Ao Dr. Paulo Sarmanho da Costa Lima, por ter sido meu coorientador. Grande pesquisador da Embrapa Meio-Norte, sempre paciente e compreensivo em todos os momentos que precisei.

A todos os professores do Programa de Pós – Graduação em Genética e Melhoramento, que são verdadeiros mestres na arte de ensinar.

Aos colegas alunos da turma de Pós – Graduação em Genética e Melhoramento (Artemisa Borges, Bruna Lima, Karla Annielle Bernardo, Laíze Raphaelle Lemos, Lívia do Vale Martins, Mário Henrique Torres, Marcones Ferreira Costa) pela compreensão, solidariedade e companheirismo que senti de todos. Peço perdão pela minha falta de interatividade e pelas vezes que não correspondi. Gostaria de destacar a solidariedade que recebi da Artemisa Borges, sobretudo nas atividades no Laboratório onde, por muitas vezes, ela deixou de fazer seu trabalho para me ajudar.

Ao pessoal do Laboratório de Biologia Molecular e Biotecnologia da Embrapa Meio-Norte que tive a oportunidade de conviver durante o período de pesquisa. Com eles aprendi e sem eles nada teria conseguido. Cito, com orgulho: Artemisa Borges, Leonardo Furtado de Oliveira, Jéssica Bárbara Viana e Yeda Gabriela Silva.

À Universidade Federal do Piauí (UFPI) e ao Programa de Pós – Graduação em Genética e Melhoramento, pela oportunidade de realização do curso.

À Embrapa Meio-Norte, pela oportunidade de realização desta pesquisa.

À Prof. Dra. Ângela Celis e à Prof. Dra. Regina Lúcia Ferreira Gomes que são consideradas por todos, e agora por mim, como duas grandes mães que acolhem todos os alunos sempre de forma solidária e amorosa. Agradeço a Deus pela oportunidade de conhecê-las e, de alguma forma, fazer parte da história de suas vidas.

E, por fim, a todos que contribuíram de forma direta ou indireta para a realização desse trabalho.

“Sonho que se sonha só é só um sonho que se sonha só, mas sonho que se sonha junto é realidade.”

Raul Seixas

SUMÁRIO

RESUMO 10
ABSTRACT.....10
1 INTRODUÇÃO 12
2 REVISÃO DE LITERATURA 14
2.1 Classificação botânica 14
2.2 Distribuição geográfica.....15
2.3 Aspecto ecológico.....16
2.4 Aspecto econômico 19
2.5 Aspecto nutricional 19
2.6 Importância medicinal.....20
2.7 Utilizações diversas.....21
2.8 Conservação e caracterização do germoplasma.....22
2.9 Marcadores Moleculares no estudo genético de plantas.....23
2.9.1 Marcadores ISSR.....25
2.10 Diversidade genética.....26
2.11 Estrutura genética de populações.....28
2.12 Fluxo gênico.....32
3 MATERIAL E MÉTODOS 34
3.1 Material genético 34
3.2 Extração do DNA genômico 36
3.3 Reação de amplificação do DNA e eletroforese.....37
3.4 Análise estatística dos dados.....38
3.4.1 Análise de diversidade genética.....38
3.4.2 Agrupamento entre populações.....39
3.4.3 Análise de similaridade genética.....39
3.4.4 Estrutura e diferenciação genética.....40
4 RESULTADOS.....42
4.1 Diversidade genética.....42
4.2 Relações genéticas entre populações.....44
4.3 Similaridade genética entre acessos.....45
4.4 Estrutura e diferenciação genética.....47

5 DISCUSSÃO 51
6 CONCLUSÕES 56
7 REFERÊNCIAS 57

RESUMO

A sapucaia (*Lecythis pisonis* Cambess) é uma espécie arbórea originária da Amazônia brasileira. Suas sementes são ricas em calorias e proteínas com grande potencial para comercialização. O germoplasma da maioria das espécies do gênero *Lecythis* é pouco conhecido, principalmente sobre a variabilidade genética existente, o que é muito importante para sua manutenção. Recentemente, as metodologias mais utilizadas para analisar-se a variabilidade genética têm sido os marcadores moleculares, tais como ISSR (*Inter Simple Sequence Repeat).* Neste trabalho, utilizou-se a técnica de ISSR para estimar a divergência genética entre acessos, os níveis de variação genética e os padrões da estrutura populacional em *L. pisonis.* De 22 iniciadores testados, onze foram utilizados para amplificar amostras de DNA de 17 acessos de sapucaia do Banco Ativo de Germoplasm (BAG) da Embrapa Meio-Norte, correspondentes a três populações (duas do Piauí e uma do Maranhão). Foram analisados 96 locos entre os indivíduos estudados. Alta variação foi encontrada em nível de espécie (porcentagem de bandas polimórficas (PPB) = 94,79%; diversidade genética de Nei (h) = 0,3110 e índice de Shannon (I) = 0,4732). Já nas populações analisadas, a porcentagem de bandas polimórficas variou de 20,83% a 94,79%, a diversidade genética de Nei variou de 0,0863 a 0,2969 e o índice de Shannon variou de 0,1260 a 0,4457. Foi detectada significativa diferenciação genética entre as populações (Φ_{ST} = 10,66%). No entanto, a maior diferenciação genética foi encontrada dentro das populações (89,34%), havendo um elevado fluxo gênico entre elas (N_m = 1,10). Os genótipos BGS 2 e BGS 4 foram os mais divergentes enquanto os genótipos BGS 14 e BGS 15 foram os mais semelhantes. Desta forma, o BAG de sapucaia analisado, apresenta elevada diversidade genética e potencial para uso em trabalho de melhoramento genético.

Palavras–chave: banco de germoplasma; castanheira; diversidade genética; marcadores ISSR; sapucaia.

ABSTRACT

The sapucaia (*Lecythis pisonis* Cambess) is a arboreal species native to the Brazilian Amazon. Its seeds are rich in calories and protein with great potential for commercialization. The germplasm from most Lecythis species has little known, especially about the genetic variability existing what is very important for its maintenance. Recently, the methodologies most used for analyze the genetic variability has been the molecular markers, such as ISSR (Inter Simple Sequence Repeat). In this work, it was used the ISSR technique for estimating the genetic divergence among accessions, levels of genetic variation and the patterns of populational structure in *L. pisonis*. From 22 *primers* tested, eleven were used to amplify DNA samples from 17 sapucaia accesses the Active Bank Germoplasm (BAG) of Embrapa Mid-North, corresponding to three populations (two of Piauí and a of the Maranhão). Were analyzed 96 loci among the access studied. High variation has been found in species level (percentage of polymorphic bands (PPB) = 94.79%; genetic diversity of Nei (h) = 0.3110 and Shannon index (I) = 0.4732). Already in the populations studied, the percentage of polymorphic bands ranged from 20.83% to 94.79%, the genetic diversity of Nei ranged from 0.0863 to 0.2969 and the Shannon index ranged from 0.1260 to 0.4457. Significant genetic differentiation was detected among populations (ΦST = 10.66%). However, the largest genetic differentiation was found within populations (89.34%), with a high gene flow between them (Nm = 1.10). The BGS 2 and BGS 4 genotypes were the most divergent, genotypes while the BGS 14 and BGS 15 were the most similar. Thus, the sapucaia BAG analysis has high genetic diversity, as well as potential use in genetic improvement.

Keywords: germplasm bank; chestnut; genetic diversity; ISSR markers; sapucaia.

1 INTRODUÇÃO

A sapucaia (*L. pisonis*) é uma espécie popularmente conhecida no Brasil e pertencente às Lecythidaceae, uma família com distribuição pantropical e uma das mais abundantes nas paisagens brasileiras (MATTA; SCUDELLER, 2012; MORI, 1988). A sapucaia é originária da Amazônia, com ampla distribuição pelo território brasileiro e em especial na Região Meio-Norte do Brasil, constituída pelos estados do Piauí e Maranhão, onde ainda é pouco conhecida, mas que apresenta potencialidades para o mercado de amêndoas (SOUZA, V. et al. 2008).

A sapucaia é caracterizada como uma árvore vistosa, podendo atingir até 50 metros de altura. O tronco pode alcançar 90 centímetros de diâmetro (LORENZI, 2002). As folhas são decíduas, suas flores têm pétalas normalmente púrpuras e seus frutos são arredondados com casca rígida e coloração castanha. Quando o fruto amadurece, as sementes são liberadas por uma tampa que se solta na parte inferior. As amêndoas constituem uma importante fonte de calorias, minerais e macro nutrientes, podendo ser consumidas cozidas, cruas ou torradas (TEIXEIRA, 2014).

As diversas partes da planta têm sido utilizadas na medicina popular e estudos recentes demonstraram a presença de compostos com moderada atividade citotóxica (SILVA, L. et al., 2012). Dentre as diversas utilidades e importâncias da sapucaia, podemos destacar seu uso na indústria madeireira, na produção de cosméticos, no cultivo ornamental e na produção de objetos artesanais. Outro destaque é o potencial econômico das suas amêndoas, consideradas como a mais saborosa da Amazônia e com perspectiva de potencial comercialização (REVILLA, 2002; SAMPAIO, 2000; VILLACHICA, 1996).

A Castanha-do-Brasil, da mesma família da sapucaia, é atualmente uma das castanhas mais comercializadas no mundo e o principal produto extrativo para exportação da Região Norte do Brasil. Desta forma, a exploração de castanhas tornou-se a principal atividade econômica na região amazônica e esta atividade vem sustentando milhares de famílias (HOMMA, 2004). Acredita-se que este extrativismo logo irá se estender para outras castanheiras, como a sapucaia. Assim, é importante ressaltar que o desconhecimento do uso sustentável dos recursos naturais e a intensidade de exploração têm provocado prejuízos irreparáveis à flora brasileira, colocando muitas espécies de frutíferas nativas em risco de extinção (RIBEIRO, R.;

RODRIGUES, 2006). Esses fatos, cada vez mais, têm chamado a atenção para a necessidade de conservação das espécies.

As espécies podem ser conservadas, tanto no seu local de ocorrência (*in situ*) como em Bancos Ativos de Germoplasmas (*ex situ*). O germoplasma conservado serve como um reservatório de genes e constitui uma fonte de variabilidade genética disponível (RAMALHO et al., 2012). A variabilidade genética dentro de uma espécie é fundamental para garantir seu potencial adaptativo frente às adversidades ambientais e à pressão antrópica. Quantificar essa variabilidade dentro das populações é necessário para avaliar como as espécies enfrentam essas adversidades e se mantém vivas e reprodutivas ao longo dos tempos (RIBEIRO; RODRIGUES, 2006). Para a definição de estratégias de conservação e uso sustentável dos recursos genéticos das espécies nativas, são de fundamental importância a compreensão dos padrões de distribuição da diversidade genética e o nível de diferenciação intra-específico.

Os marcadores moleculares de DNA têm sido muito utilizados nos estudos de diversidade genética e melhoramento de plantas (BORÉM; CAIXETA, 2009). O seu uso aumenta consideravelmente a eficiência de um programa de melhoramento, pois permite que a seleção e novos cruzamentos sejam realizados em uma mesma geração (FERREIRA; GRATTAPAGLIA, 1998). Em Coleções de Germoplasma, os marcadores moleculares são úteis na caracterização dos genótipos, fornecendo parâmetros para a identificação de indivíduos divergentes e auxiliando o melhorista na seleção de combinações mais promissoras e favoráveis aos cruzamentos (RAMALHO et al., 2012). Marcadores moleculares de natureza dominante, como o ISSR (*Inter Simple Sequence Repeat)*, têm se popularizado nos estudos genéticos de plantas, por apresentar vantagens como: baixo custo, alta transferibilidade e reprodutibilidade, obtenção de elevado polimorfismo, dentre outras (VIJAYAN, 2005).

A Embrapa Meio-Norte, em Teresina-Piauí, possui um Banco Ativo de Germoplasma de sapucaia e apesar da importância desta espécie, poucos estudos genéticos foram realizados. Até o momento, nenhum estudo foi publicado sobre o uso de marcadores moleculares de DNA para estimar a diversidade genética e a estrutura populacional desta espécie. Diante disto, este trabalho teve os seguintes objetivos:

Geral:

- Avaliar a variabilidade genética por meio de marcadores moleculares ISSR entre acessos de Sapucaia pertencentes ao BAG da Embrapa Meio-Norte, provenientes de populações do Piauí e Maranhão.

Específicos:

- Estimar os parâmetros genéticos, por meio de marcadores ISSR, para que se possa caracterizar a diversidade genética dos acessos;
- Verificar se há divergência genética entre os acessos e compreender como ela está distribuída molecularmente;
- Analisar estrutura genética dentro e entre as populações.

.

2 REVISÃO DE LITERATURA

2.1 Classificação botânica

A espécie *L. pisonis* pertence à divisão Magnoliophyta (Angiospermae); classe Magnoliopsida (Dicotyledonae); ordem Lecythidales e família Lecythidaceae. Tem como sinônimos científicos: *L. ollaria* Spruce, *L. paraensis* Huber, *L. paraensis* Hub. Ex Spruce, *L. pisonis* subsp. *usitata* S.A. Mori & Prance, *L. usitata* Miers. (ANGIOSPERMAS PHYLOGENY GROUP - APG III, 2009). A família Lecythidaceae conta com cerca de 300 espécies pertencentes a 17 gêneros. O gênero Lecythis é neotropical e conta com 27 espécies descritas, sendo o segundo mais abundante da família (SMITH; MORI; PRANCE, 2010).

Segundo Lorenzi (2002), *L. pisonis* é popularmente conhecida no Brasil como sapucaia, castanha de sapucaia, sapucainha, pau de cachimbo, papo de anjo, fruta de cotia, fruta da lepra, dentre outros nomes. Em outros países pode ser conhecida como castaña de monte, nuez del paraíso, olla-de-momo, sapucaya (Espanha); quatelé, marmite de singe, noix de sapucaia (França); marmite de singe, quatelé (Guiana Francesa); cream nut, paradise nut, monkey nut, monkey pot, sapucaia nut (Inglaterra) (MORI,1990; MORI; PRANCE, 1990).

A palavra sapucaia tem origem tupi (sa + puca + ia) e significa “olho que se abre”, isso porque quando o opérculo do fruto se abre, parece que se vê um olho.

Outra explicação para o nome sapucaia é que tenha derivado da palavra galinha (gûyrasapukaîa = ave que grita) na linguagem tupi. Isso porque era comum, entre índios e portugueses, a troca de sementes de sapucaia por galinhas. Além de *L. pisonis*, outras árvores que pertencem à família botânica das *Lecythydaceae,* também podem ser conhecidas com o nome de sapucaia. Portanto, as sapucaias e seus frutos, já eram bastante conhecidos e aproveitados pelas populações que habitavam o Brasil na época da chegada dos primeiros europeus, no século XVI (PANTANO, 2014).

2.2 Distribuição geográfica

A família Lecythidaceae, à qual pertence à sapucaia, é pantropical, sendo que cento e vinte e duas espécies pertencentes a nove gêneros estão distribuídas por todo o Brasil, tendo a sua maior diversidade na floresta amazônica, onde Lecythidaceae é também uma das famílias mais abundantes. As 27 espécies descritas do gênero *Lecythis* estão distribuídas da Nicarágua para os vales da Cordilheira dos Andes; na costa do Pacífico do Equador e da Colômbia; em toda a região da Amazônia, das Guianas e na Mata Atlântica brasileira (MATTA, L.; SCUDELLER, 2012).

No Brasil, onde ocorrem 54% das espécies de Lecythidaceae das Américas, há mais espécies desta família que qualquer outro país no mundo (MORI, 1988). *L. pisonis* é encontrada no Amazonas, Bahia, Espírito Santo, Maranhão, Minas Gerais, Pará, Pernambuco, Piauí (ESCORCIO et al.,1999), Rio de Janeiro, Rondônia (*UNITED STATES DEPARTMENT OF AGRICULTURE* - USDA, 2014) e São Paulo (MARTINS et al., 2002). Por ser bastante frequente na Amazônia Central e Oriental, acredita-se que seja originária dessas regiões (SOUZA, A. et al., 1996). Conforme Soares (1990), distribui-se do Ceará até o Rio de Janeiro, sendo particularmente frequente no sul da Bahia e norte do Espírito Santo. Ocorre, também, em áreas de cerrado e de florestas pouco densas (REDE DE SEMENTES CERRADO, 2014). A sapucaia é uma espécie existente na flora nativa brasileira e, embora seja pouco conhecida no estado do Piauí, apresenta potencialidades para o mercado de amêndoas (CARVALHO; SOUZA; ALVES, 2008).

2.3 Aspecto ecológico

A sapucaia apresenta altura média de 20 a 30 metros podendo atingir até 50 metros de altura, com o tronco podendo alcançar 90 cm de diâmetro (LORENZI, 2002). Apresenta casca acinzentada a marrom escuro, com fissuras verticais profundas (Figura 1a). As folhas são decíduas e, quando novas, são vermelhas, creme ou verde claro, depois se tornando verde escuro, brotando justamente antes da floração (Figura 1b). As inflorescências são de racemos, nascendo em ramos abaixo das folhas, com 6-20 flores. As flores possuem 3-7 cm de diâmetro e têm pétalas normalmente púrpuras, às vezes brancas, sempre descorando para branco após a queda (Figura 1c) (MORI; PRANCE, 1990).

Os frutos são arredondados e achatados, globosos, oblongos e utilizados como recipientes (MORI; PRANCE, 1990), possuem um pericarpo com até 3 cm de espessura, muito duro e uma tampa que se solta quando maduros, liberando as sementes (PANTANO, 2014) (Figura 1d). Souza, V. et al. (2008) classificaram os frutos, do BAG de sapucaia da Embrapa Meio Norte, de acordo com o peso médio em: pequenos (<1.100g), médios (1.300g a 1.500g), grandes (1.600g a 1.900g) e extragrandes (> 2000g).

As sementes são fusiformes, com 8 – 12 sulcos, e possuem funículo basal cercado pelo arilo branco e carnoso (MORI; PRANCE, 1990). Sousa, V. et al. (2008) encontraram sementes com diâmetro médio de 1,4 cm, comprimento médio de 2,35 cm e peso médio de 2,27g. Braga et al. (2007), constataram que as sementes de sapucaia apresentaram grande variação em todas as dimensões avaliadas e encontraram valores médios de 4,31 cm para o comprimento, 2,58 cm para largura e 8,65 g para o peso médio (Figura 1e).

A sapucaia é considerada uma planta decídua, heliófita ou esciófita, seletiva higrófita, ocorrendo principalmente no interior da mata primária densa (LORENZI, 1992). Está adaptada a regiões com altitude de até 800 m, com precipitação média entre 1800 a 5000 mm. O crescimento da sapucaia é lento e a regeneração natural pode ser considerada como regular, requerendo bastante luz para se desenvolver (SAMPAIO, 2000). Pode ser encontrada com densidades variadas de 0,2 a 11 árvores por hectare, como por exemplo, na região de Manaus (SOUZA, A. et al.,1996).

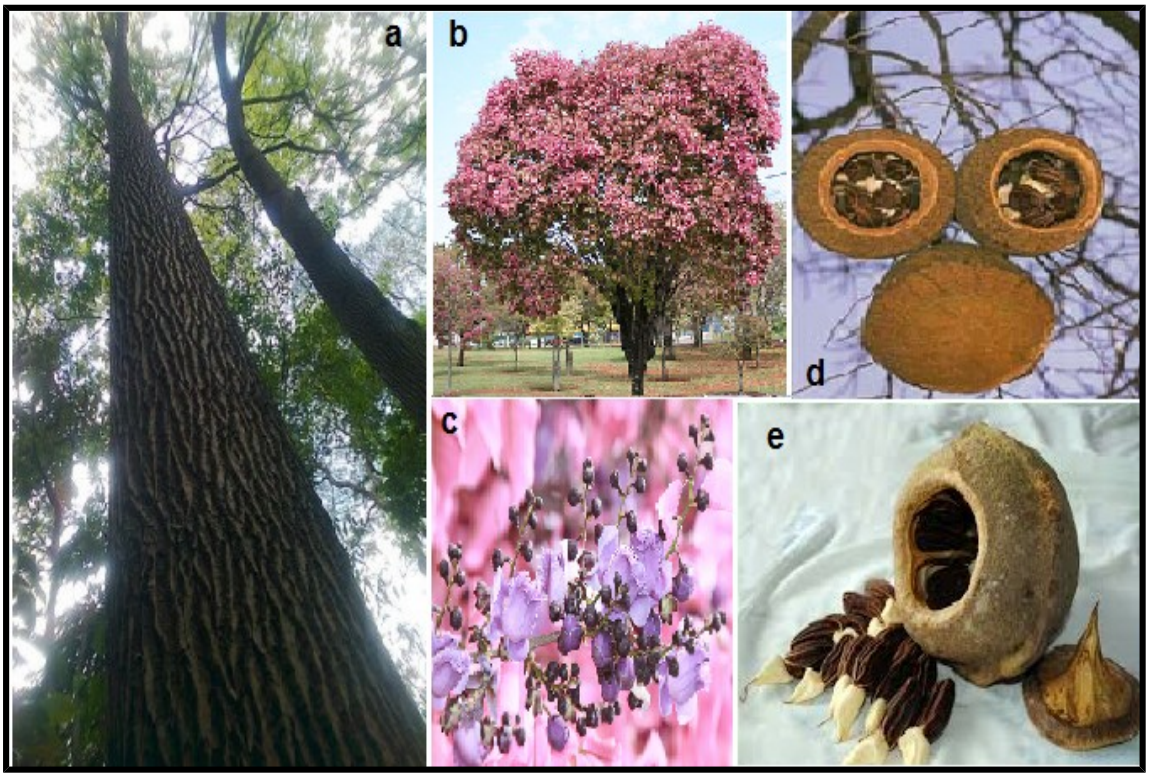

Figura. 1- Sapucaia. **a**: arvore adulta; **b**: folhas na primavera; **c**: inflorescência; **d**: frutos com e sem a tampa; **e**: frutos e sementes. Fonte: Pantano, 2014.

Segundo Mori (1988), as espécies de Lecythidaceae quando em áreas de estações secas definidas, florescem principalmente na estação seca e frutificam no início da estação chuvosa. Onde não há estação seca definida, outros fatores podem servir como estímulo para floração. Por exemplo, no sul da Bahia, onde não ocorrem meses do ano com menos chuva, a sapucaia floresce ano após ano ao final do inverno até o final da primavera. Este fato sugere que a sapucaia inicia a floração em resposta ao aumento no comprimento do dia e/ou temperatura, fatores que ocorrem na primavera. No hemisfério sul, as flores de *L. pisonis* aparecem durante a primavera. Na costa leste do Brasil as flores aparecem de setembro a novembro e na Amazônia, a floração é mais proeminente de julho a janeiro (MORI; PRANCE, 1990). A floração pode ocorrer junto com o surgimento das novas folhas de cor rosa ou lilás, e como consequência, toda copa adquire uma bela coloração lilás. Dependendo da região pode haver desfolha parcial ou total (LORENZI, 1992; SAMPAIO, 2000).

Na família Lecythidaceae a evolução floral ocorre principalmente no androceu que pode ser do tipo actinomorfo ou do tipo zigomorfo. As abelhas são os principais polinizadores, sendo que somente em *Lecythis poiteaui* ocorre polinização por morcegos. As flores de Lecythidaceae oferecem três recompensas aos seus

polinizadores: pólen não diferenciado, pólen diferenciado, e néctar (MORI; MATTOS; SANTOS, 1980). A sapucaia é zigomorfa e as abelhas (*Xylocopa frontalis)* retiram da flor um tipo de pólen que não germina para alimentar suas larvas, enquanto outro tipo de pólen, o que germina, fica na cabeça e nas costas da abelha, na posição certa para polinizar a próxima flor visitada (MORI, 1988).

As sementes são dispersas por macacos, araras, papagaios e morcegos que comem as sementes quando os frutos se abrem. Alguns roedores apreciam as sementes com alguns dias de maturidade após terem caído no chão (FAO, 1986). O processo de germinação ocorre facilmente, mas pode variar quanto à porcentagem e o tempo de germinação. Segundo Sampaio (2000), quando semeadas logo após a coleta, 80% das sementes germinam. De acordo com Ataíde et al. (2011), a germinação teve início 15 dias após a semeadura em substrato de papel toalha ou areia. Sob condições experimentais (disponibilidade de água e temperatura de 26°C), Cardoso e Nakau (2014) observaram ocorrência de germinação entre 15 e 31 dias (média de 19,5 ± 1,5 dias) após a semeadura. No entanto, a germinação aumentou proporcionalmente com o aumento da densidade da semente, de forma que estes autores recomendam o uso de sementes com densidades >0,9 $g.mL^{-1}$, para garantir uma porcentagem elevada de plântulas.

Para a produção de mudas, os frutos podem ser colhidos diretamente da árvore, ao iniciarem sua abertura espontânea, e levados ao sol para completar a abertura e liberação das sementes. Também é possível recolher as sementes do chão após a abertura do fruto e antes de serem consumidas por roedores. O desenvolvimento das mudas é lento, ficando prontas para plantio em 8 – 10 meses (REDE DE SEMENTES DO CERRADO, 2014).

Para Lorenzi (1992), o desenvolvimento das plantas é moderado (no campo), podendo atingir 2,5m aos dois anos. Siqueira e Ribeiro (2001) observaram que a sapucaia apresentou 1,5 m após 24 meses. De acordo com Sampaio (2000), as plantas podem alcançar 60 cm após o primeiro ano de plantio. Com um incremento médio anual (IMA) de 1 m, as árvores com dez anos de idade alcançam 10 a 12 m. No entanto, o uso de propagação vegetativa reduz o porte das plantas, facilitando assim, o corte do fruto maduro e o aproveitamento das castanhas para a comercialização no mercado.

2.4 Aspecto econômico

Atualmente as castanhas têm recebido atenção especial, pois são fontes naturais de vitaminas, minerais, proteínas e ácidos graxos essenciais, podendo assim contribuir para a dieta humana e de animais (CARVALHO, I. et al., 2012). Dentre as castanhas mais comercializadas mundialmente, estão a castanha de caju e a castanha-do-pará (SANTOS, J.; SENA; ROCHA, 2010).

Mesmo sendo geralmente considerada a melhor castanha da Amazônia, a castanha de *L. pisonis* é pouco comercializada, sendo raramente encontrada no mercado (SAMPAIO, 2000). Isso pode estar relacionado ao fato de seus frutos não caírem da árvore quando maduros. Carvalho M. (1996) mostrou que a sapucaia está entre as frutíferas nativas que apresentam potencial econômico.

Villachica (1996) relatou que a sapucaia, pelo seu sabor adocicado, pode ocupar parte do mercado de outras amêndoas tropicais e de clima temperado, assim como desenvolver um mercado próprio. Para Ferrão (2001) o interesse do mercado internacional pelas sementes de sapucaia levou à instalação de pomares entre a Bahia e o Espírito Santo. Segundo Revilla (2002) há um alto potencial extrativista e de futuro, devido às grandes populações desta espécie, sendo possível uma alta produção, desde que haja demanda.

Devido à qualidade da castanha de sapucaia, o seu comércio poderia ser lucrativo, uma vez que Sampaio (2000) informou que se a espécie fosse enxertada, usando-se os mesmos métodos para a castanha-do-brasil, e plantada a 100 árvores/ha, a produtividade seria de 5.000 frutos/ha, com cerca de 200.000 castanhas/ha. Segundo Revilla (2002) a espécie chega a produzir, no seu ambiente natural, de 8 a 10 ton/ha/ano de peso fresco.

2.5 Aspecto nutricional

A semente de sapucaia pode ser consumida no estado natural, tostada, assada ou cozida (MATTA, A., 2003), constituindo um excelente alimento e podendo substituir as nozes, amêndoas ou castanhas europeias (ESCORCIO et al., 1999). Analisando as características químico-nutricionais de 16 acessos de sapucaia, Souza, V. et al. (2008) observaram que a quantidade de energia variou de 718,98 a 754,32 kcal 100 g^{-1} com média de 736,87 kcal 100 g^{-1} apresentando em média 65,99 % de gordura, 19,62 % de proteína e 3,43% de cinza.

Nesta mesma análise, foi revelado quantidades consideráveis de minerais, indicando que a sapucaia é bastante rica em cálcio (294,40 mg, 100 g^{-1}), fósforo (748,22 mg, 100 g^{-1}), potássio (674,29 mg, 100 g^{-1}), magnésio (449,89 mg, 100 g^{-1}), cobre (3,04 mg, 100 g^{-1}) e manganês (6,86 mg, 100 g^{-1}). Vale ressaltar que estes valores são superiores ou equivalentes aos relatados na literatura para as sementes comestíveis mais comercializadas no Brasil, como castanha-de-caju, castanha-do-pará, avelã, noz, amendoim e macadâmia (FREITAS; NAVES, 2010).

Portanto, as sementes de sapucaia são uma importante fonte alternativa de calorias, minerais e macronutrientes, tais como lipídeos e proteínas, podendo ser utilizada por humanos e animais. Segundo a literatura, as sementes de sapucaia são consumidas por vários povos indígenas da Bacia Amazônica sendo muito apreciadas na Europa e América do Norte (DURATEX, 1989; GOMES, 1977).

2.6 Importância medicinal

Diversas partes da sapucaia têm sido usadas na medicina popular. A casca contém taninos e saponinas não alcalóides, sendo empregada no combate à tosse (MILLIKEN et al., 1986). Sua decocção é tida como tônica e possui propriedades diuréticas e cicatrizantes, sendo utilizada para o tratamento de icterícia, hepatite, estados de febre (SAMPAIO, 2000; REVILLA, 2002) e complicações do fígado (WICKENS, 1995). O pó ou o óleo da casca do caule possui uso anti-sifilítico (VIEIRA; MARTINS, 2000).

O óleo das sementes é tido como diurético, sendo eficaz contra a diabete, no tratamento da gota e da sífilis, podendo ser usado como um emoliente na redução da dor muscular (AGRA; FREITAS; BARBOSA, 2007). A emulsão feita com as amêndoas é anticatarral e antipirética (DURATEX, 1989; FERRÃO, 2001). A água de maceração dos frutos é utilizada no combate da albuminúria, diabetes, catarro vesical e micoses superficiais da pele (SAMPAIO, 2000; REVILLA, 2002). As folhas são popularmente utilizadas sob a forma de banhos no tratamento de prurido do corpo (FRANCO; BARROS, 2006).

Silva, L. et al. (2012) descreveram, pela primeira vez, a presença de compostos que mostraram uma moderada atividade citotóxica em extratos de folhas de *L. pisonis*.

2.7 Utilizações diversas

A madeira da sapucaia possui várias utilidades que vão desde a construção civil e naval, bem como na construção de móveis, artigos domésticos decorativos, brinquedos e instrumentos musicais (SOUZA, M. et al., 1997). O líber fibroso pode ser usado na produção de cordas grosseiras, confecção de tecidos (MILLIKEN et al., 1986) ou pode ser utilizado substituindo o papel para enrolar cigarros (MATTA, A., 2003).

Da casca é possível extrair a estopa usada em calafetagem, como também, pode ser utilizada como isca para peixes, uma vez que são ricas em uma substância venenosa chamada de saponinas (MILLIKEN et al., 1986). O fruto é utilizado como adorno (LORENZI, 1992), como recipiente para uso doméstico (WICKENS, 1995) e também podem ser feitas obras torneadas como compoteiras, vasos, cofres e outros (CORRÊA, 1984).

O óleo de sapucaia é descrito como um líquido amarelo claro, com odor agradável e gosto adocicado (SAMPAIO, 2000). Devido ao seu alto teor em ácidos linoléico e oléico, pode ser empregado na formulação de produtos hidratantes e para o cuidado da pele e dos cabelos. Além disso, pode ser usado na confecção de sabões, cremes, xampu, condicionadores e como óleo para o corpo, com propriedades emolientes ou para massagens (ARAÚJO et al., 2007).

Dentre as diversas utilidades da sapucaia, podemos destacar seu cultivo ornamental em alguns estados brasileiros como em São Paulo e Rio de Janeiro. Por exemplo, na época do império, o paisagista francês Auguste François Glaziou escolheu a sapucaia para compor o eixo monumental no jardim do palácio do Imperador Dom Pedro II, no Rio de Janeiro (TERRA, 1993).

Culturalmente, alguns fatos curiosos estão ligados à sapucaia como o das crianças do povo Ka'apor serem proibidas de comer as castanhas da sapucaia pois tornariam-se preguiçosas (BALÉE,1994). Além disso, tudo indica que o provérbio "macaco velho não mete a mão em cumbuca" está relacionado ao fato de macacos jovens, que, por apreciarem suas sementes, enfiam a mão na urna do fruto, enchem-na com sementes e não conseguem retirá-las depois (SOARES, 1990).

Diante da sua importância, é essencial a conservação da sapucaia em bancos de germoplasma para resguardar esse patrimônio genético. Alem da conservação, estudos que envolvam a caracterização do germoplasma são fundamentais para o

conhecimento da diversidade genética a qual poderá ser aproveitada em programas de melhoramento genético.

2.8 Conservação e caracterização do germoplasma

Podemos definir germoplasma como um conjunto de genótipos de uma espécie, considerada como um todo. A variação encontrada em uma determinada espécie pode ser devido ao ambiente ou as diferenças genéticas. A existência de variação genética é importante para a evolução das espécies e um pré-requisito para o melhoramento genético de plantas. Assim, o germoplasma é a fonte de variabilidade genética disponível para o melhoramento de plantas (BESPALHOK, GUERRA; OLIVEIRA, 2014).

A conservação dos recursos genéticos é considerada uma das questões mais importantes para a sobrevivência da humanidade (RAMALHO et al., 2012). O germoplasma conservado serve como um reservatório de genes que os melhoristas podem acessar quando precisam resolver problemas específicos (BESPALHOK, GUERRA; OLIVEIRA, 2014). A conservação das espécies pode ser feita *in situ* ou *ex situ*. No primeiro caso, a conservação da variabilidade genética é feita no local de origem e, no segundo caso, a conservação é feita em Bancos Ativos de Germoplasma (BAGs), ou seja, o local onde os genótipos de interesse são conservados (BORÉM; CAIXETA, 2009).

Uma das vantagens da conservação *in situ* é a possibilidade das espécies continuarem seus processos evolutivos, o que lhes permitiria coevoluírem com os seus patógenos e pragas. Já a conservação *ex situ*, tem a vantagem de armazenar material genético de espécies com um valor potencial como fonte de variação para programas de melhoramento genético; conservar material genético de espécies ancestrais de cultivares modernas e de espécies sob uma eminente ameaça de extinção (RAMALHO et al., 2012). Segundo Costa et al. (2011), os BAGs são uma alternativa para a conservação dos recursos genéticos e a avaliação da diversidade genética entre os acessos de um BAG resulta em informações sobre potenciais genitores a serem usados em programas de melhoramento, possibilitando a identificação de duplicatas e o intercâmbio de germoplasma entre pesquisadores.

No Brasil, a conservação da variabilidade genética tem sido coordenada pela Empresa Brasileira de Pesquisas Agropecuária (EMBRAPA). A conservação *ex situ* é

feita em diversos bancos de germoplasma distribuídos nas várias Unidades da EMBRAPA. Até o momento, só existe uma coleção com genótipos de *L. pisonis*, a coleção da Embrapa Meio-Norte/Teresina-PI, criada em 2006 e que conta com 23 genótipos estabelecidos em campo por propagação de sementes provenientes de diversas áreas de ocorrência da espécie nos Estados do Maranhão e Piauí. No entanto, apenas dois estudos foram realizados utilizando este banco de germoplasma: o primeiro, realizado por Souza, V. et al. (2008), avaliou a variabilidade fenotípica entre os acessos para características físicas de frutos e amêndoas e características químico-nutricionais de amêndoas. No segundo, realizado por Santos, F. (2010), foi realizada a caracterização molecular de 16 acessos usando marcadores moleculares do tipo RAPD.

Portanto, é importante o desenvolvimento de mais pesquisas que visem ampliar o conhecimento dessa espécie contribuindo à manutenção e multiplicação deste germoplasma para usos agronômicos futuros. A utilização de marcadores moleculares de DNA pode auxiliar na quantificação da diversidade e na identificação de genótipos desejáveis para programas de melhoramento.

2.9 Marcadores moleculares no estudo genético de plantas

Os marcadores moleculares são capazes de viabilizar a caracterização de grande número de genótipos através de procedimentos relativamente simples e rápidos. Nas últimas décadas, tem aumentado o número de marcadores moleculares juntamente com o crescimento de estudos de melhoramento genético, genética de populações e bioinformática. Em consequência, isso tem gerado um imenso progresso no conhecimento aplicável à conservação e manipulação da biodiversidade (SILVA, C. et al., 2014).

Os marcadores moleculares possuem diversas vantagens com relação aos marcadores morfológicos: são fenotipicamente neutros, geram um número praticamente ilimitado de polimorfismo genético, são aplicáveis a qualquer tipo de material vegetal e em qualquer estágio de desenvolvimento, permitem a identificação correta de variedade sem necessitar de muitos caracteres e estão livres dos efeitos epistáticos (FALEIRO, 2007).

Para Ferreira e Grattapaglia (1998), um marcador molecular é definido como todo e qualquer fenótipo molecular oriundo de um gene expresso, como no caso de

isoenzimas, ou de um segmento específico de DNA correspondente a regiões expressas ou não do genoma. Os marcadores de DNA se baseiam fundamentalmente nas análises das diferenças em pequenas sequências do DNA entre indivíduos. Diversas técnicas de biologia molecular têm sido empregadas com sucesso para revelar classes distintas de marcadores moleculares. Geralmente, são estas técnicas que dão o nome aos diferentes tipos de marcadores, os quais podem ser de natureza dominante ou codominante (DELGADO, 2006).

Os marcadores de DNA são muito úteis no melhoramento genético de plantas, pois têm sido usado em estudos de diversidade genética, melhoramento assistido, mapeamento genético, *fingerprinting*, reconhecimento, isolamento de genes, caracterização e identificação de variações genéticas, etc. (BORÉM; CAIXETA, 2009).

O desenvolvimento da amplificação de segmentos de DNA via PCR ("*Polymerase Chain Reaction*", ou reação da polimerase em cadeia), possibilitou o desenvolvimento de diversos marcadores baseados em PCR. Entre eles, podemos citar: o RAPD (*Random Amplified Polymorphic* DNA) que significa polimorfismo do DNA amplificado ao acaso, o AFLP (*Amplified Fragment Length Polymorphism*), ou seja, polimorfismo de comprimento de fragmentos amplificados, o SSR (*simple sequence repeats microsatellites*) que se baseia em sequências simples repetitivas ou microssatélites e o ISSR (*Inter Simple Sequence Repeats*) que são sequências simples repetidas entre dois microssatélites. Todos estes, representam a principal classe de marcadores em estudos genéticos de plantas, devido à sua grande utilidade e exploração (MALIK et al., 2014).

Os marcadores moleculares de natureza codominante apresentam certa vantagem em relação aos marcadores de natureza dominante. Marcadores codominantes como os microssatélites são mais informativos, especialmente porque podem diferenciar os genótipos homozigóticos dos heterozigóticos. No entanto, tais marcadores precisam ser desenvolvidos separadamente para cada espécie, o que pode levar tempo e custar muito caro (KREMER et al., 2005). Vários exemplos mostraram que a transferibilidade de SSR continua limitada às espécies filogeneticamente relacionadas (BARRENECHE et al., 2004; COLLEVATTI et al., 1999).

Marcadores dominantes, como RAPDs e ISSR, tornaram-se populares em estudos genéticos de plantas nos últimos anos. Esses marcadores são uma

alternativa atraente, devido a sua natureza hipervariável, o vasto número de locos que podem ser examinados, o baixo custo, a alta transferibilidade e a pequena quantidade de material usado por amostra (KREMER et al., 2005). No entanto, o uso de um ou de outro marcador vai depender, entre outros fatores, do objetivo do trabalho, dos recursos financeiros disponíveis, da disponibilidade de recursos humanos e da espécie a ser estudada (FALEIRO, 2007).

2.9.1 Marcadores ISSR

As limitações do uso de marcadores codominantes têm levado ao uso de marcadores de natureza dominante para avaliar a diversidade genética em plantas tropicais, principalmente, para espécies que não possuem marcadores codominantes ou quando o sequenciamento do DNA não está disponível (KREMER et al., 2005). Os marcadores ISSR, são de natureza dominante e utilizados com sucesso para estimar a diversidade genética, inter e intra específica, em uma ampla gama de espécies de plantas (REDDY; SARLA; SIDDIQ, 2002)

Estudos têm mostrado que dados de diversidade genética obtidos com marcadores dominantes ISSR são compatíveis com pesquisas realizadas com marcadores codominantes (MADESIS et al., 2014; WANG et al., 2013). Estes marcadores são baseados na técnica de PCR, que envolve a amplificação de segmentos de DNA entre duas regiões de microssatélites orientadas na direção oposta.

A técnica utiliza um único *primer* de microssatélites, geralmente de 16 – 25 pb de comprimento, como iniciador para amplificação de sequências Inter - SSR de diferentes tamanhos. As repetições microssatélites utilizadas como iniciadores podem ser di-nucleotídica, tri-nucleotídica, tetra-nucleotídica ou penta-nucleotídica. Os *primers* podem ser desancorados ou, mais usualmente, ancorados na extremidade 5’ ou 3’ com 1 a 4 bases degeneradas (GUPTA et al., 1994; ZIETKIEWICZ; RAFALSKI; LABUDA,1994).

Os marcadores ISSR tem alta reprodutibilidade e combinam a maior parte dos benefícios das análises de AFLP e microssatélites com a universalidade do RAPD (REDDY; SARLA; SIDDIQ, 2002). A variabilidade nos ISSR pode ocorrer sempre que em um genoma esteja faltando a sequência repetida ou se há uma deleção ou uma inserção que modifica a distância entre as repetições. Além disso, o grau de

polimorfismo também pode variar de acordo com a natureza do *primer*. Assim, quando são utilizados *primer* di-nucleotídeos e ancorados na posição 5' ou 3' é possível obter um maior número de bandas e um elevado polimorfismo. No entanto, os iniciadores ancorados na extremidade 3' apresentam um padrão de bandas mais nítidos que aqueles ancorados na extremidade 5' (VIJAYAN, 2005).

Dependendo do método de detecção é possível obter níveis mais elevados de polimorfismo. Assim, um maior número de bandas é obtido quando é usada a eletroforese em gel de poliacrilamida. No entanto, níveis elevados de polimorfismo são detectados quando os produtos de amplificação de ISSR são obtidos em gel de agarose (REDDY; SARLA; SIDDIQ, 2002).

Os marcadores ISSR possuem potencial para integração em programas de melhoramento de plantas. Têm sido utilizados na caracterização e manutenção de germoplasma, para estimar a extensão da diversidade genética de diversas espécies, na resolução de problemas relacionados com a filogenia, no mapeamento genômico e em estudos de biologia evolutiva (SOUZA, I.; SOUZA, V.; LIMA, 2012; VIJAYAN, 2005; YU; JING; CHENG, 2014; WONGSAWAD; PEERAPORNPISAL, 2014).

2.10 Diversidade genética

Ao se iniciar um programa de melhoramento genético de determinada espécie é necessário o conhecimento da diversidade genética existente para melhor explorar essas diferenças. Segundo Torres et al. (2000), a diversidade genética pode ser entendida como a variação hereditária devido a constituição genética dos indivíduos de uma população, sendo portanto, responsável por parte de suas diferenciações fenotípicas. Do ponto de vista molecular, a diversidade genética é o polimorfismo entre indivíduos de uma população, em um ou mais fragmentos de DNA, identificados na forma de presença de bandas de tamanhos diferentes (alelos), derivadas de um mesmo loco, no caso de um marcador codominante, ou na presença e ausência de bandas, no caso de marcadores dominantes (RAMALHO et al., 2012).

O estudo da diversidade genética é o processo pelo qual a variação entre populações, entre grupos ou entre indivíduos de um grupo é analisada por meio de um método específico ou de uma combinação de métodos (BORÉM; CAIXETA,

2009). A diversidade genética pode ser quantificada de diversas maneiras: número de alelos por locos, porcentagem de locos polimórficos, heterozigosidade esperada (índices de diversidade de Nei), que é a proporção de heterozigotos esperada para uma população que esteja em equilíbrio de Hardy – Weinberg (FREELAND, 2005) e o índice de Shannon. Essa quantificação pode ser feita tanto por técnicas multivariadas como por marcadores moleculares que permitem diferenciar os genótipos.

O estudo da diversidade entre e dentro de populações fornece as bases para a identificação de indivíduos divergentes, auxiliando o melhorista na seleção de combinações mais promissoras e favoráveis aos cruzamentos (FALEIRO et al., 2011; MELCHINGER, 1999). Por meio da análise da diversidade genética pode-se também gerar conhecimento sobre a dispersão de uma espécie, selecionar genótipos com certa resistência a doenças e determinar locais para futuras coletas de germoplasma (BORÉM; CAIXETA, 2009). Assim, há a necessidade de ampliarmos as informações genéticas para a conservação de espécies florestais como *L. pisonis*.

O uso de marcadores moleculares é extremamente útil em quase todas as etapas de um programa de conservação de germoplasma, pois dão maior credibilidade aos resultados obtidos (BORÉM; CAIXETA, 2009). Um procedimento amplamente utilizado no estudo de diversidade, usando marcadores moleculares, consiste em obter um grande número de bandas polimórficas entre os indivíduos de uma população, sem se preocupar se a variabilidade é alélica ou gênica. Assim, por meio de uma matriz de dados de 0 e 1, onde a ausência de bandas em um indivíduo é representada por 0 e a presença por 1, estima-se a similaridade genética entre os indivíduos (RAMALHO et al., 2012).

Uma das estimativas de similaridade genética entre indivíduos (sg_{ij}) mais utilizada é o coeficiente de Jaccard calculado por meio da expressão:

$$\mathbf{sg_{ij} = a/(a+b+c) \quad (1)}$$

Nessa expressão, **a,** significa o número de combinações com a presença de uma determinada banda nos indivíduos i e j; **b** as combinações com a presença da banda no indivíduo i e ausência no j; e **c**, as combinações com a ausência da banda no indivíduo i e presença em j (ROHLF, 1992).

A partir da estimativa de similaridade pode-se obter a dissimilaridade genética:

$$(dg_{ij}) = 1\text{-}sg_{ij} \quad (2)$$

Como as estimativas de sg_{ij} variam de 0 até 1, estimativas próximas de 0 indicam que dois indivíduos são geneticamente muito diferentes e próximas de 1, são muito semelhantes. No caso de se usar a dissimilaridade, interpreta-se de forma inversa (RAMALHO et al., 2012).

O conhecimento do parentesco dos acessos de um BAG é importante, pois permite eliminar acessos duplicados ou orientar na obtenção de outros não representados no banco. A similaridade genética corresponde ao parentesco entre dois acessos e normalmente pode estar correlacionada com a variação de diversos caracteres. Assim, os marcadores moleculares têm sido considerados mais úteis para esse fim, pois fornecem o relacionamento com maior precisão (RAMALHO et al., 2012). Estimativas de similaridade e dissimilaridade genética têm sido muito utilizadas em estudos de diversidade entre acessos em bancos de germoplasma (LING et al. 2015; SANTOS, M. et al., 2008; SOUZA, I. et al. 2011; SOUZA, I.; SOUZA, V.; LIMA, 2012).

2.11 Estruturas genética de populações

Segundo Futuyma (2002), população é um grupo de organismos da mesma espécie que ocupam uma região geográfica mais ou menos definida e que exibem uma continuidade reprodutiva ao longo das gerações. No entanto, a maioria das populações pode agrupar em subpopulações menores, nas quais geralmente ocorrem os cruzamentos. Esse agrupamento é chamado de estrutura populacional ou subdivisão populacional. Desta forma, quando a população está estruturada, inevitavelmente existe diferenciação genética entre as subpopulações (HARTL; CLARK, 2010). Estas diferenças nas frequências alélicas entre as subpopulações podem ser resultantes de forças evolutivas tais como: mutação, migração, seleção, deriva genética, por efeito do acaso ou por características reprodutivas da espécie (HAMRICK; MURAWSKI, 2009).

A estrutura das populações pode ser analisada usando diversas abordagens: estatística F de Wright (WRIGHT, 1965), análise da diversidade gênica em populações subdivididas (NEI, 1977), coeficiente de ancestralidade de Cockerham (COCKERHAM, 1969; VENCOVSKY, 1992), análise de variância molecular – AMOVA (EXCOFFIER; SMOUSE; QUATTRO, 1992) e os métodos estatísticos Bayesianos. Estas abordagens são muito eficientes, principalmente quando baseadas em estimativas de frequências alélicas obtidas através de marcadores moleculares.

A estatística F de Wright tem sido muito utilizada para quantificar o efeito do endocruzamento da subdivisão populacional. O índice de fixação de Wright (F) é um útil quantificador de diferenciação genética entre e dentro dos diferentes níveis hierárquico populacional a serem comparados. Assim, a estatística F permite a caracterização da variabilidade entre as populações (F_{ST}), assim como dos níveis médios de endogamia ao nível subpopulacional (F_{IS}) e no nível da população total (F_{IT}). O índice F_{ST} é a medida mais inclusiva da subdivisão populacional, pois pode-se comparar os níveis menos inclusivos com os mais inclusivos da hierarquia populacional e mede todos os efeitos da estrutura populacional combinada. Desta forma, F_{ST} é a estatística informativa para examinar o nível total de divergência genética entre subpopulações (HARTL; CLARK, 2010).

Teoricamente, os valores de F_{ST} podem variar de 0 (nenhuma divergência genética) a 1 (fixação de alelos). No entanto, o valor observado é muito menor do que 1. Wright (1965) sugeriu os seguintes valores como orientação para a interpretação de F_{ST}:

- De 0 a 0,05 (pequena diferenciação genética)
- De 0,05 a 0,15 (moderada diferenciação genética)
- De 0,15 a 0,25 (grande diferenciação genética)
- Acima de 0,25 (diferenciação genética muito grande)

A diversidade gênica em populações subdivididas de Nei (1977, 1987) permite a obtenção de uma estimativa de divergência (G_{ST}) e uma comparação dos níveis de heterozigosidade entre e dentro das populações a partir de uma base diferente daquela que fundamenta as estimativas de F_{ST}. Esta abordagem é derivada da fórmula de diversidade gênica média total (dada por $H_T = H_S + D_{ST}$), onde H_S é a diversidade gênica média dentro das populações, e D_{ST} é a diversidade gênica

média entre as populações. Assim, a diferenciação genética relativa entre populações é obtida por:

$$G_{ST} = D_{ST}/H_T \quad (3)$$

Os coeficientes de coancestralidade (θ) permitem a avaliação da divergência em diferentes níveis hierárquicos e possibilitam a obtenção de estimativa de endogamia a partir de análise não viesada. São obtidos a partir da decomposição dos componentes de variação das frequências alélicas, conforme Cockerham (1969). Os valores de θ tem significado análogos ao do F_{ST} de Wright e medem a diferenciação genética entre as populações.

A abordagem de variância molecular (AMOVA) foi inicialmente desenvolvida para estimar a estrutura genética de populações com base em dados moleculares de haplótipos em organismos haplóides, mas é facilmente estendida para produzir uma análise coerente, utilizando dados de marcadores moleculares dominantes, para organismos diplóides (MICHALAKIS e EXCOFFIER, 1995).

A AMOVA produz estimativas de componentes de variância análogos a estatística – F, conhecidos como estatística - ϕ, que expressa a correlação da diversidade haplotípica em diferentes níveis da subdivisão populacional. Os índices de estrutura genética são estimados com base na informação sobre o conteúdo de haplótipos alélicos, assim como das suas frequências (EXCOFFIER; SMOUSE; QUATTRO, 1992).

Na montagem da AMOVA as informações sobre as divergências alélicas entre haplótipos é inserida como uma matriz do quadrado das distâncias entre todos os pares de haplótipos. O significado dos componentes de covariância associados aos diferentes níveis possíveis de estrutura genética (dentro dos indivíduos, dentro de populações, dentro de grupos de populações, entre os grupos) é testado através de uma abordagem permutacional (EXCOFFIER; SMOUSE; QUATTRO, 1992). Assim, a AMOVA proporciona um quadro geral para a análise da estrutura genética populacional. O número de níveis hierárquicos no quadro da AMOVA depende da estrutura genética em que o conjunto de dados foi organizado. Exemplificando, podemos considerar os níveis de hierarquia, onde P = número total de populações e N = número total de dados genotípicos ou número total de genes para dado haplótipo (Tabela 1).

Tabela 1- Esquema de Análise de Variância Molecular (AMOVA). GL: grau de liberdade; SQ: soma dos quadrados; QM: quadrado médio; E (QM): esperança do quadrado médio.

FONTE DE VARIAÇÃO	GL	SQ	QM	E (QM)
Entre populações	P - 1	SQ_a	QM_a	$n\sigma_a^2 + \sigma_b^2$
Dentro de populações	N - P	SQ_b	QM_b	σ_b^2
TOTAL	N - 1			σ^2

Os componentes de variâncias de cada nível de hierarquia são extraídos das esperanças dos quadrados médios. Segundo Michalakis e Excoffier (1995) temos:

$$\Phi_{ST} = \sigma_a^2 / \sigma^2 \quad (4)$$

Onde Φ_{ST} é igual à proporção da variabilidade molecular de haplótipos entre as populações.

A inferência da estrutura populacional também pode ser realizada usando uma abordagem Bayesiana. Neste caso, Pritchard, Stephens e Donnelly (2000) descreveram um modelo baseado em método de agrupamento utilizando dados de genótipos multilocos. Neste modelo, os autores assumem que há K populações (onde K pode ser desconhecido), onde cada um dos K é caracterizado por um conjunto de frequências alélicas em todos os locos. Assim, os indivíduos da amostra são atribuídos para as populações, ou em conjunto para duas ou mais populações, se houver mistura de genótipos. Esse modelo pode ser aplicado para a maioria dos marcadores genéticos, pois não assume um processo de mutação em particular. Com este método e suas extensões é possível demonstrar a presença de estrutura populacional, atribuir indivíduos às populações, estudar zonas híbridas e identificar indivíduos migrantes e miscigenados (FALUSH; STEPHENS; PRITCHARD, 2003; FALUSH; STEPHENS; PRITCHARD, 2007).

Com base em dados genéticos, também é possível estabelecer a similaridade genética para várias populações. Para isso, utilizam-se medidas de distâncias genéticas que expressam a diferença genética entre duas populações. Assim, o

índice de distância indica um valor igual a zero, quando não há diferença entre duas populações e o valor igual a um quando a diferença é máxima. Com as distâncias é possível verificar estrutura de agrupamento, ou a reconstrução da história evolutiva das populações. Um dos índices de distância genética, muito utilizado, na avaliação de populações foi estabelecido por Nei (1972,1987). No entanto, novas medidas de distância genética têm sido utilizadas, principalmente para ser usada com dados moleculares (GOLDSTEIN et al., 1995).

2.12 Fluxo gênico

De acordo com Slatkin (1987), o fluxo gênico pode ser entendido como o intercâmbio de genes entre as populações, incluindo o movimento dos gametas, zigotos, indivíduos ou grupos de indivíduos de um lugar para outro. Em plantas, o fluxo de pólen ou sementes, ou ambos, têm sido utilizados como índices de fluxo gênico (SMYTH; HAMRICK, 1987). O estudo do fluxo gênico tem demonstrado importância relativa na diferenciação de populações. No entanto, existe uma escassez de informações sobre os níveis de fluxo gênico em um grande número de plantas (GOVINDARAJU, 1989).

Em virtude da dificuldade em se obter uma estimativa direta do fluxo gênico em populações naturais, foram propostos métodos indiretos para obter níveis aproximados de fluxo gênico nestas populações (WRIGHT, 1951). As medidas diretas referem-se ao fluxo gênico contemporâneo, enquanto que as indiretas são baseadas na estrutura de populações e referem-se ao fluxo gênico histórico, ou passado.

Métodos indiretos para estimar o fluxo gênico são baseados na análise das frequências alélicas em populações. Relativamente poucos métodos indiretos estão disponíveis para inferir a magnitude e direção de fluxo gênico. O primeiro relato foi via F_{ST} de Wright (1951), que é baseado na análise da estrutura genética de populações, sendo rotineiramente usado para estimar o número absoluto de organismos migrantes (N_m) que entram em cada população ou subpopulação a cada geração (MCDERMOTT; MCDONALD, 1993). O F_{ST} é uma função efetiva do número de migrantes por geração, N_m, onde N é o tamanho populacional e m é a proporção de migrantes por geração. Desta forma tem-se que:

$$N_m = 1 / 4 (1/F_{ST} - 1) \quad (5)$$

Outro método utilizado para determinar o N_m é através da estimativa de divergência genética (Gst) de Nei (1987). O uso desta estimativa demonstrou-se bastante atrativa por diversas razões citadas em Govindaraju (1989). De acordo com Crow; Aoki, 1984 o fluxo gênico aparente (N_m) entre populações pode ser estimado indiretamente conforme a equação apresentada abaixo que corrige a análise para pequeno número de populações.

$$N_m = 1/4\alpha (1/Gst - 1) \quad (6)$$

Onde:

Nm: número de migrantes por geração

Gst: divergência genética entre as populações

α: correção para o número finito de populações no modelo de ilhas (n), conforme a equação (7):

$$\alpha = [n/ (n-1)]^2 \quad (7)$$

Segundo Slatkin (1987), os valores de N_m foram agrupados em três categorias: alta ($N_m > 1$), intermediários ($0,25 < N_m < 0,99$) e baixo ($N_m < 0,25$). Em Hartl e Clark (2010), para os seguintes valores de N_m teremos:

- $N_m = 0,25$ (um migrante a cada quatro gerações);
- $N_m = 0,5$ (um migrante a cada duas gerações);
- $N_m = 1$ (um migrante a cada geração);
- $N_m = 2$ (dois migrantes a cada geração);

A migração é uma força atuante contra a divergência genética entre as subpopulações. Assim, o movimento de um ou mais indivíduo por geração entre populações é suficiente para prevenir a diferenciação substancial entre as populações. Geralmente, se $N_m < 1$, uma considerável diferenciação entre as populações pode ser esperada, e se $N_m > 1$, haverá pouca diferenciação entre as populações (WRIGHT, 1951).

3 MATERIAL E MÉTODOS

3.1 Material genético

O estudo foi realizado com 17 genótipos de sapucaia que fazem parte da coleção de Germoplasma da Embrapa Meio – Norte (Tabela 2), estabelecido em Teresina – PI (05º05'S de latitude e 42º 49'W de longitude e 72m de altitude). Esses genótipos foram coletados em quatro áreas de ocorrência da espécie nos estados do Piauí e Maranhão (Figura 2).

Tabela 2 – Genótipos de sapucaia analisados por ISSR e seus respectivos locais de coleta, Teresina, PI, 2015.

Áreas /Populações	Código dos acessos	Município de coleta	Unidade Federativa
1	BGS1	José de Freitas	Piauí
	BGS2	José de Freitas	Piauí
	BGS3	José de Freitas	Piauí
	BGS4	José de Freitas	Piauí
	BGS6	José de Freitas	Piauí
	BGS7	José de Freitas	Piauí
	BGS8	José de Freitas	Piauí
	BGS10	José de Freitas	Piauí
	BGS11	José de Freitas	Piauí
	BGS12	José de Freitas	Piauí
2	BGS14	Altos	Piauí
	BGS15	Altos	Piauí
	BGS16	Altos	Piauí
	BGS17	Altos	Piauí
3	BGS18	Timon	Maranhão
	BGS19	Timon	Maranhão
4*	BGS13	Teresina	Piauí

* Não representa uma população

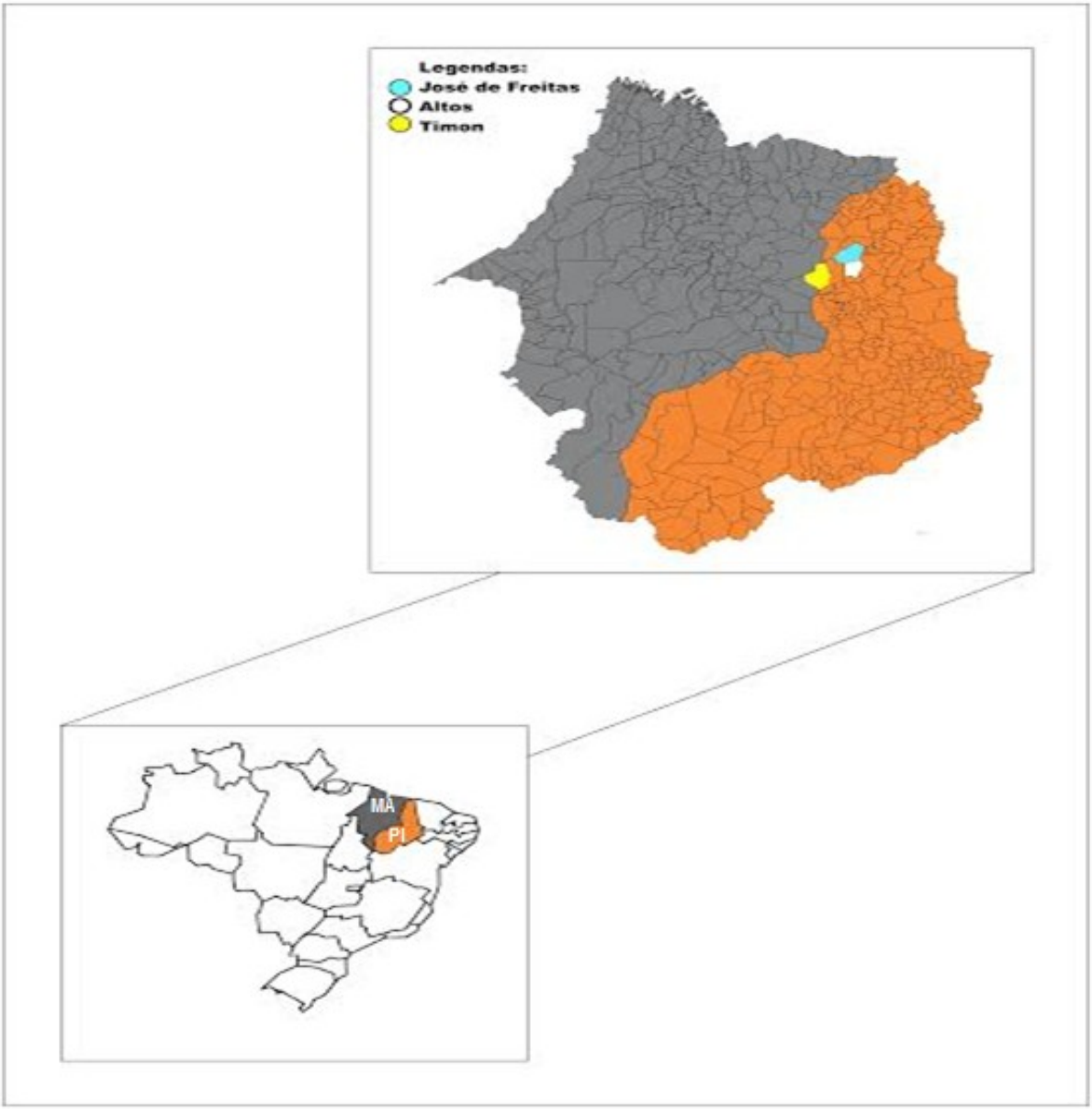

Figura 2- Mapa de localização indicando os estados e municípios onde as populações estão localizadas.

A população 1 está representada pelo município de José de Freitas (04º45’21" de latitude sul e 42º34’33" de longitude oeste) que dista 45 Km em linha reta de Teresina. Os genótipos desta população foram coletados em três pontos diferentes: Rodovia José de Freitas – Estaca Zero (BGS 1); Fazenda Santa Rita (BGS 2, BGS 3, BGS 4, BGS 6, BGS 7, BGS 8, BGS 10) e Fazenda Montes Claros (BGS 11, BGS 12).

A população 2 está representada pelo município de Altos – PI (05º02’17" de latitude sul e 42º27’36" de longitude oeste) localizado a 41 Km de Teresina. Esta população está representada pelos seguintes genótipos e seus locais de coletas:

BGS 14 (margem da BR – 343), BGS 15 (margem da BR – 344), BGS 16 (margem da BR – 345) e BGS 17 (margem da BR – 346). Os municípios de José de Freitas e Altos pertencem à macrorregião Meio-Norte e à microrregião de Teresina.

Os dois genótipos BGS 18 e BGS 19 pertencem à população 3. A mesma está representada pelo município de Timon – MA, localizado a 5,8 Km de Teresina. Ambos os genótipos desta população foram coletados à margem da BR -316.

O genótipo (BGS 13) foi coletado no Campus da Universidade Federal do Piauí (05º05’13" de latitude sul e 42º48’41" de longitude oeste) e por ser único, não representa uma população do município de Teresina. Portanto, o mesmo foi utilizado somente na análise de agrupamento entre os acessos.

3.2 Extração do DNA genômico

Foi extraído material genético de 17 acessos de Sapucaia pertencentes à Coleção de Sapucaia da Embrapa Meio-Norte. Para a extração do material genético, foram utilizadas folhas jovens e sadias. Todas as amostras coletadas foram devidamente identificadas, envolvidas em papel toalha, embaladas individualmente, armazenadas em gelo e conduzidas ao Laboratório de Biotecnologia e Biologia Molecular da Embrapa Meio-Norte, Teresina-PI.

A extração do DNA foi realizada obedecendo-se as recomendações do kit de purificação Qiagen (DNEASY, 2006). Macerou-se 100 mg de tecido vegetal fresco em tubos de 2,0 ml contendo cinco esferas de vidro (3 mm) e tampão do kit de purificação. A maceração foi feita com o auxílio do homogeneizador e Extrator Precellys 24.

A quantificação de DNA foi realizada em gel de agarose a 0,8% com TBE 0,5x (Tris-Borato-EDTA). Os padrões de bandas foram visualizados corando-se os géis com GelRed-Biotium ® 1x e utilizando-se luz ultravioleta para comparar o DNA das amostras com o DNA-λ na concentração de 100 ng/ul e estimar a concentração de DNA por µl e, consequentemente, a quantidade total de DNA extraído.

A qualidade do DNA obtido foi mensurada com o auxílio do equipamento Nanodrop 2000 Spectrophotometer (Thermo Scientific). As amostras de DNA foram diluídas em tampão TE na concentração de 7 $ng.\mu L^{-1,}$ para serem usadas nas reações de PCR. O DNA foi armazenado a -20°C.

3.3 Reação de amplificação do DNA e eletroforese

Foram testados 22 *primers* quanto à resolução e o nível de polimorfismo gerado. Estes *primers* foram desenvolvidos pela *University of Britsh Columbia* (UBC), Vancouver, Canadá. Os *primers* foram amplificados em amostras de DNA de três indivíduos, um em cada população. Desta forma, foram selecionados os que amplificaram o maior número de marcadores polimórficos e de boa qualidade. Dos 22 *primers* testados, 11 foram selecionados para a caracterização molecular dos 17 genótipos estudados (Tabela 3).

Tabela 3 – *Primers* de ISSR usados na caracterização molecular dos acessos de sapucaia. Tm: Temperatura de *Melting*, Ta: Temperatura de anelamento. Teresina, PI, 2015.

Primer	**Tm (°C)**	**Ta (°C)**	**Sequência (5'-3')***
UBC 807	47,0	52,0	$(AG)_8T$
UBC 810	45,4	52,0	$(GA)_8T$
UBC 825	51,4	55,0	$(AC)_7A$
UBC 826	52,8	56,0	$(AC)_8C$
UBC 836	48,9	54,0	$(GA)_8YA$
UBC 840	47,4	52,0	$(GA)_8YT$
UBC 841	48,5	50,0	$(GA)_8YC$
UBC 842	48,8	50,0	$(GA)_8YG$
UBC 845	48,1	46,0	$(CT)_8RG$
UBC 856	52,8	59,0	$(AC)_8YA$
UBC 864	43,6	43,0	$(ATG)_5$

* Y- (C,T); R- (A,G);

As reações de amplificação foram realizadas de acordo com o protocolo de Williams et al. (1990) modificado. As concentrações utilizadas para a reação final de 10 µL foram de: tampão 1,0x [20 mM Tris HCl, pH 8,0; 0,1 mM EDTA; 1 mM DTT; 50% (v/v) glicerol], $MgCl_2$ a 2,0 mM, 0,8 mM de dNTP (dATP,dCTP, dGTP e dTTP), 0,8 µM do *primer*, 1 U de Taq DNA polimerase (Invitrogen) e 1 µL de DNA genômico (7 ng/µl). Foi usado um termociclador Veriti 96 Well Thermal Cycler (Applied Biosystems®) com capacidade para 96 amostras.

As reações envolveram repetidos ciclos de desnaturação, anelamento e extensão seguindo os seguintes parâmetros: uma fase inicial de desnaturação a 94ºC por 4 minutos; seguida de 40 ciclos de desnaturação a 94ºC por 1 minuto, anelamento a uma temperatura dependente do *primer* durante 1 minuto, extensão a 72ºC por 1 minuto e uma etapa de extensão final a 72ºC por 5 minutos.

Os produtos de cada amplificação foram separados em gel de agarose a 1,5% com tampão 0,5X TBE, por meio de eletroforese submetida a 110 volts durante 4 horas. Os padrões de bandas foram visualizados corando-se os géis com GelRed™ e utilizando-se luz ultravioleta para visualização do padrão de bandas. Esses padrões foram fotografados.

As fotografias dos géis foram analisadas para a determinação do número de bandas, com base na resolução e no grau de amplificação por meio de avaliação visual, comparando-as com o marcador molecular DNA Ladder 100 e 50 pb (Invitrogen).

3.4 Análise estatística dos dados

Cada banda ISSR foi considerada como um loco bi-alélico independente (WILLIAMS et al., 1990) e os indivíduos foram genotipados quanto a ausência (0) ou presença (1) da banda. A partir dos dados foram geradas duas matrizes binárias: uma com 16 indivíduos que serviu para as análises de diversidade e estruturação genética e outra com 17 indivíduos que foi utilizada na análise de similaridade genética entre os acessos. Em todas as análises foi assumido que os locos eram dominantes e estavam em equilíbrio de Hardy – Weinberg.

3.4.1 Análise de diversidade genética

Com a finalidade de avaliar os níveis de diversidade genética, foram estimados os seguintes parâmetros: porcentagem de locos polimórficos (% PLP), sendo que o loco foi considerado polimórfico quando a sua freqüência não ultrapassou 0,95; o índice de diversidade de Shannon (*I)* calculado conforme Lewontin, 1972 e o coeficiente de diversidade genética de Nei (*h*), calculado conforme Nei, 1987. Nestas análises foi utilizado o *software* POPGENE versão 1.31 (YEH; BOYLE; XIYAN, 1999). Outra estimativa de diversidade avaliada foi a

heterozigosidade imparcial H (imparcial) de Nei (1978) considerando a expansão de Taylor (LYNCH; MILLIGAN, 1994) que foi obtida com o programa TFPGA (*Tools for Population Genetic Analyses*) versão 1.3 (MILLER,1997).

3.4.2 Agrupamento entre populações

As relações genéticas entre populações foram determinadas com base nas estimativas das distâncias genéticas de Nei (1972). Estas foram utilizadas para construção de um dendrograma para o qual foi empregado o método de agrupamento UPGMA (*Unweighted Pair Group Method with Arithmetic Mean* - Média aritmética não ponderada para agrupamento aos pares) (NEI, 1987).

A consistência de cada nó gerado pelo conjunto original de dados foi calculada pelo método de reamostragem com 1000 *bootstraps* sobre os locos. Todos os cálculos foram obtidos com o *software* TFPGA (MILLER, 1997) utilizando a expansão de Taylor (LYNCH e MILLIGAN, 1994).

3.4.3 Análise de similaridade genética

A similaridade genética foi estimada entre os 17 acessos e calculada a partir do coeficiente de similaridade de Jaccard conforme a expressão de Rohlf (1992). A matriz de similaridade genética foi utilizada para a análise de agrupamento dos acessos, pelo método UPGMA. Foi calculado o coeficiente de correlação cofenética (r) entre a matriz de similaridade genética e a matriz dos valores cofenéticos, para se verificar o ajuste entre a matriz de similaridade e o dendrograma (CRUZ; CARNEIRO, 2003). Valores de r acima de 0,8 indicam boa representatividade (BUSSAD; MIAZAKI; ANDRADE, 1990). Realizou-se a análise de reamostragem (*bootstrap)* com 1000 repetições para se verificar a estabilidade dos agrupamentos. Todos os cálculos foram feitos com o auxílio do programa PAST 1.34 (HAMMER; HARPER; RYAN, 2001).

3.4.4 Estrutura e diferenciação genética

Os 16 indivíduos foram hierarquizados formando dois grupos (Piauí e Maranhão) subdivididos em três populações, duas do Piauí (José de Freitas e Altos) e uma do Maranhão (Timon). A diferenciação genética entre as populações e os grupos foi avaliada através de índices de diferenciação genética entre populações, análogos ao F_{ST} de Wright. Para isso, foram estimados dois coeficientes de diversidade: o coeficiente G_{ST}, com base na análise de diversidade genética em populações subdivididas de Nei (1987), e os coeficientes theta S e theta P com base no método de Weir e Cockerham (1984). O primeiro foi obtido com o programa POPGENE, enquanto o segundo por meio do *software* TFPGA que calcula as frequências alélicas com base na expansão de Taylor delineada por Lynch e Milligan (1994). Intervalo de confiança (IC) de 95% e 1000 repetições foram utilizados. O fluxo gênico, entre as populações, foi estimado indiretamente pelo índice Gst em função de N_m conforme a equação 5 ($N_m = 0.25\ (1 - G_{ST})\ /\ G_{ST}$) (MCDERMOTT; MCDONALD, 1993).

A estruturação e a distribuição da variabilidade genética, também foram avaliadas através da análise da variância molecular (AMOVA) com a determinação de Φ_{st} (análogo ao F_{st} de WRIGHT, 1965). Foram obtidas as variâncias dentro e entre as populações. A AMOVA foi realizada utilizando o *software* Arlequin versão 3.01 (EXCOFFIER; LAVAL; SCHNEIDER, 2006). As significâncias dos componentes de variâncias foram obtidas por meio de 10100 permutações.

Para análise da estruturação da diversidade sem hierarquização *a priori*, foi empregada a abordagem bayesiana com base no método de Markov Chain Monte Carlo (MCMC) que identifica a presença de estrutura na população; o número mais provável de grupos existentes; a atribuição dos indivíduos amostrados a cada grupo e a proporção de cada genótipo advindo de outros grupos. Esta análise foi realizada no programa *Structure* versão 2.3.3 (PRITCHARD; STEPHENS; DONNELLY, 2000).

Uma primeira análise foi realizada para averiguar a correlação entre os indivíduos e suas populações geográficas. Para isso os genótipos foram agrupados sem informar a qual população cada indivíduo pertence (FALUSH; STEPHENS; PRITCHARD, 2007). Nesta análise utilizou-se do modelo de mistura fixando o número de populações para K=3. O período de execução do programa foi de 1.000.000 interações com 30.000 *burnin.*

Para averiguar o número de grupos mais provável, foram testados os valores de K variando de 1 a 6 e o melhor valor de K foi escolhido baseado nos valores de ΔK (EVANNO; REGNAUT; GOUDET, 2005). O programa foi executado com um período 30.000 *burnins* e um número de 1.000.000 interações sendo 10 repetições independentes para cada valor de K. Em seguida, os dados foram submetidos ao teste de atribuição considerando o melhor valor de K escolhido.

Para identificar a existência de migrantes, foi realizada uma análise com o programa *Structure,* usando a opção USEPOPINFO (FALUSH; STEPHENS; PRITCHARD, 2007). Este modelo parte do princípio que cada indivíduo se originou, com alta probabilidade, na região geográfica em que foi amostrado, mas com uma pequena probabilidade de que ele seja um imigrante ou descendente de imigrantes recentes. Utilizou-se também o parâmetro GENSBACK = 3, de modo que cada indivíduo poderia não pertencer à população de origem ou ter um pai, avô ou bisavô em outra população. A função MIGRPRIOR = 0,05 foi utilizada para garantir uma probabilidade de 95% de que os indivíduos pertencem à população informada antes de haver mistura. O programa foi executado com 1.000.000 de interações depois de 30.000 *burnin*.

4 Resultados

4.1 Diversidade genética

Para os 17 genótipos de *L. pisonis*, os 11 *primers* de ISSR selecionados geraram 96 locos (bandas). A Figura 3 mostra o perfil de bandas do *primer* UBC 841.

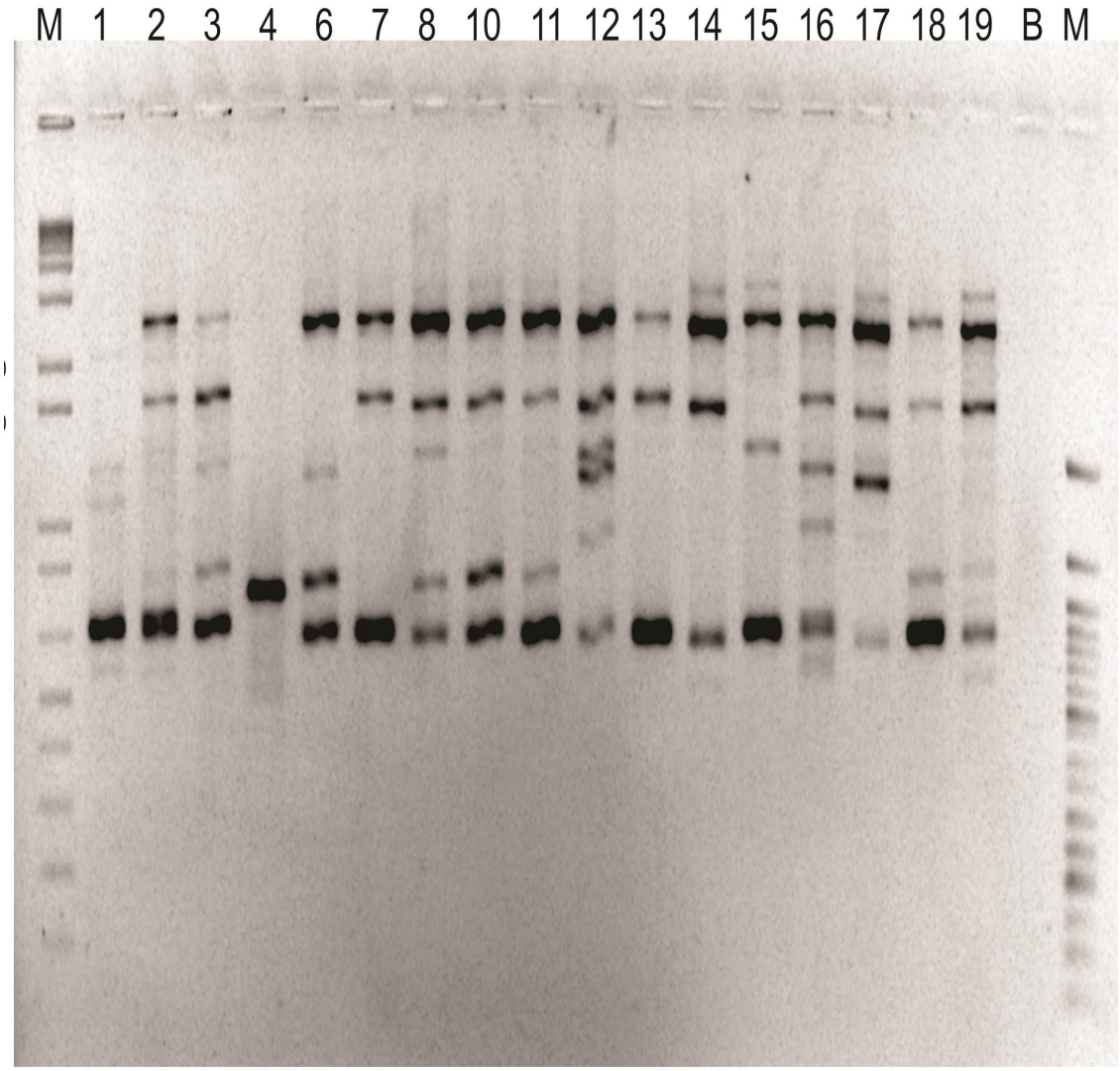

Figura 3 – Padrões ISSR de amplificação do DNA genômico de 17 genótipos de sapucaia obtidos por meio do *primer* UBC 841. M (direita): marcador de 100kb; M (esquerda): marcador de 50kb; B: Controle negativo.

Dos 96 locos analisados, 91 foram polimórficos, gerando assim, uma porcentagem de locos polimórficos (PLP) de 94,79%. Cada *primer* originou entre 5 e 11 bandas (Tabela 4), que variaram de 200 a 2500 pb.

Tabela 4 – *Primers* analisados, total de locos amplificados, total de locos polimórficos e porcentagem de polimorfismo gerado pelos *primers* ISSR. Teresina, PI, 2015.

Primer	**Total de locos**	**Locos polimórficos**	**Polimorfismo (%)**
UBC 807	11	11	100,00
UBC 810	11	11	100,00
UBC 825	9	9	100,00
UBC 826	11	11	100,00
UBC 836	5	4	80,00
UBC 840	11	9	81,20
UBC 841	8	8	100,00
UBC 842	8	7	87,50
UBC 845	6	6	100,00
UBC 856	11	11	100,00
UBC 864	5	4	80,00
Total	**96**	**91**	**94,79**
Média	**8,7**	**8,3**	

O polimorfismo total para a espécie foi de 94,79%. Nas populações, a porcentagem de locos polimórficos variou de 87,50% (população de José de Freitas) a 20,83% (população de Timon) com uma média de 54,51% (Tabela 5).

No nível de espécies, as médias dos índices de diversidade genética foram *h*= 0,3110, *I* = 0,4732 e H(imparcial) = 0,3227. A maior diversidade genética foi encontrada na população de José de Freitas enquanto que a menor está na população de Timon. Os índices de diversidade genética nas populações do Piauí foram muito maiores do que na população do Maranhão (Tabela 5).

Tabela 5 – Índices de diversidade genética entre populações de *L. pisonis*. *h*: diversidade genética (Nei, 1987); *I*: Indice de Shannon; H(imparcial): heterozigosidade; PLP: porcentagem de locos polimórficos.

Populações	***h***	***I***	**H(imparcial)**	**PLP (%)**
Nível de espécie	**0,3110**	**0,4732**	**0,3227**	**94,79**
José de Freitas	0,2969	0,4457	0,3115	87,50
Altos	0,2158	0,3166	0,2391	55,21

Timon	0,0863	0,1260	0,1030	20,83
Piauí	0,3120	0,4734	0,3266	94,79
Maranhão	0,0863	0,1260	0,1030	20,83

4.2 Relações genéticas entre populações

Na Figura 4, onde são apresentados os agrupamentos pelo método UPGMA, a partir da distância de Nei (1972), observa-se a formação de dois nós, onde as populações de José de Freitas e Altos formam um grupo sendo as mais similares geneticamente.

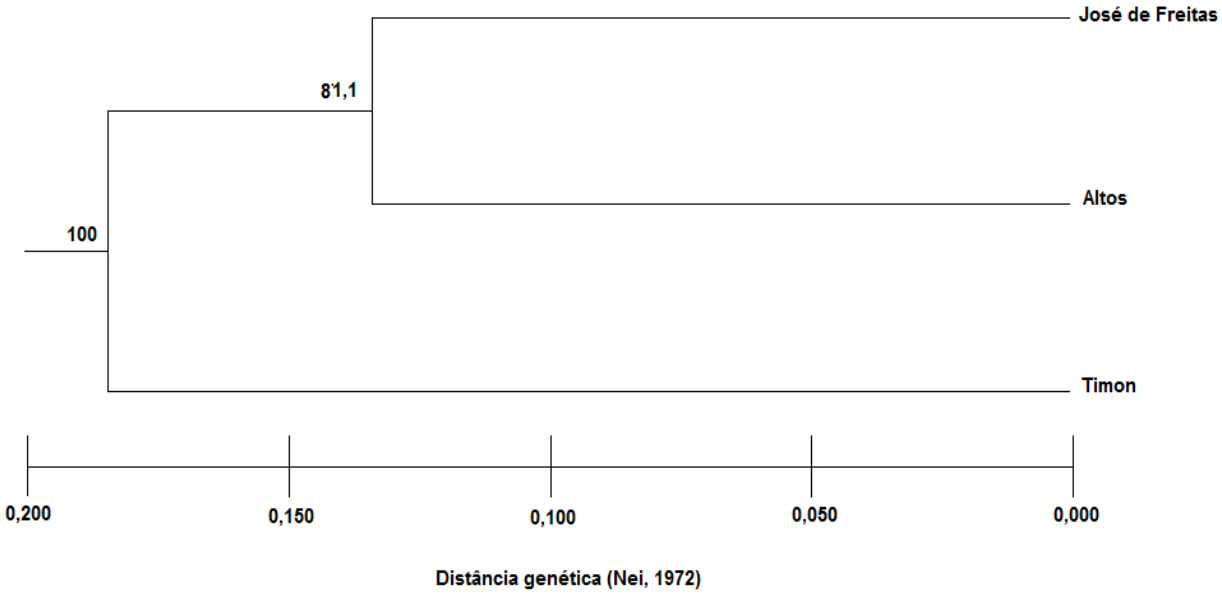

Figura 4: Análise de agrupamento (UPGMA- *Unweighted Pair Group Method with Arithmetic Mean)* a partir da Distância Genética de Nei (1972), para populações de *L. pisonis*. Números próximos aos nós indicam as proporções de replicações similares a partir de 1.000 bootstraps.

A maior distância genética (0,1951) está entre as populações de Timon e José de Freitas (Tabela 6). As proporções de replicações foram altas (81,10% e 100%) para as 1000 permutações conduzidas pelos *bootstraps.*

Tabela 6 – Distância genética de Nei (1972) entre as populações de *L.pisonis*.

Populações	Distância genética	Similaridade genética

José de Freitas / Altos	0,1336	0,8749
José de Freitas / Timon	0,1951	0,8227
Altos / Timon	0,1733	0,8409

4.3 Similaridade genética entre acessos

A análise de similaridade genética avaliada com base no coeficiente de Jaccard para todos os 17 acessos de sapucaia mostrou valores que variaram de 0,28 a 0,82. A média de similaridade genética entre os pares foi de 0,48. A similaridade máxima (0,8276) foi observada entre os acessos BGS 14 e BGS 15 enquanto que a menor (0,2836) ocorre entre os acessos BGS 2 e BGS 4 (Anexo A).

Por meio do dendrograma gerado pelo método UPGMA (Figura 5) e tomando-se como ponto de corte a similaridade média de 0,48, foi possível observar a formação de quatro grupos. O valor cofenético encontrado foi alto (r = 0,8833; 10000 permutações) demonstrando alta associação entre as similaridades obtidas pelo coeficiente de Jaccard (matriz de similaridade) e as representadas no dendrograma (matriz cofenética).

Observou-se que 14 dos 17 acessos (82,35%) foram reunidos em um único grupo (G1). Este, por sua vez, é composto por indivíduos de todas as populações: os dois do município de Timon (BGS 18 e BGS 19), três (BGS 14, BGS 15, BGS 16) dos quatro indivíduos coletados no município de Altos e a maioria dos indivíduos de José de Freitas (BGS 3, BGS 4, BGS 6, BGS 7, BGS 8, BGS 10, BGS 12, BGS 11). O valor mais baixo (0,38) de similaridade, neste grupo, foi encontrado entre os acessos BGS 4 e BGS 7, ambos da população de José de Freitas. Os demais grupos (G2, G3 e G4) foram formados pela inclusão de um único acesso.

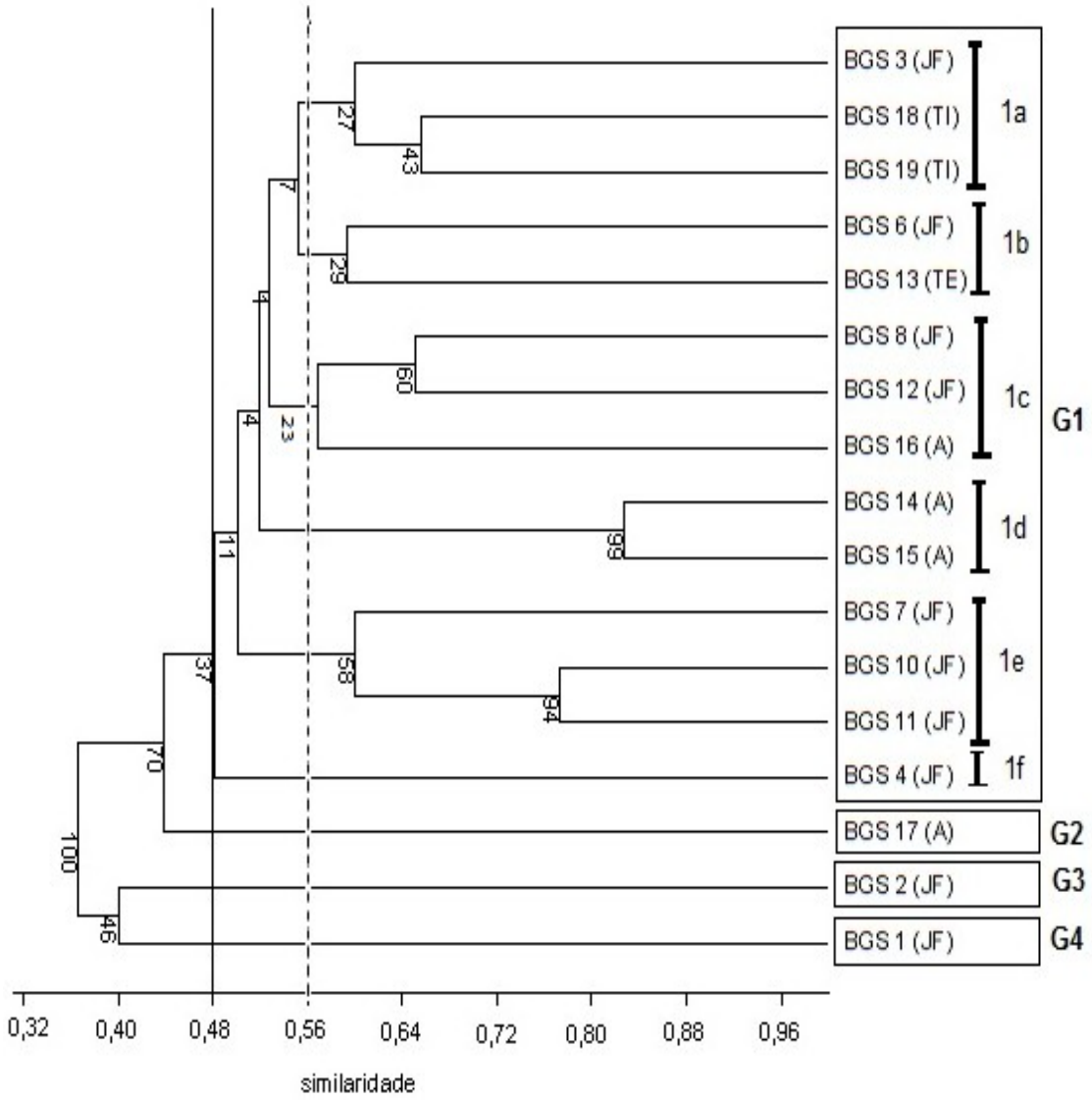

Figura 5. Dendrograma gerado pelas similaridades genéticas entre os 17 acessos de sapucaia, a partir do índice de similaridade de Jaccard e o método UPGMA, obtido a partir de 91 locos polimórficos produzidos por meio de 11 *primers* ISSR. JF: José de Freitas; A: Altos; T: Timon; TE: Teresina.

Com um coeficiente de similaridade de 0,56, é possível dividir o G1 em seis subgrupos: 1a, 1b, 1c, 1d,1e, 1f. Observa-se que os três primeiros subgrupos são heterogêneos, ou seja, formados por indivíduos oriundos de José de Freitas (JF) e de outra população.

Assim, os dois acessos de Timon estão incluídos no subgrupo 1a juntamente com o acesso BGS 3 (JF). Da mesma forma, o único acesso de Teresina (BGS 13) está incluído no subgrupo 1b juntamente com o acesso BGS 3 (JF). Já o subgrupo 1c inclui um dos acessos de Altos (BGS 16) e dois acessos de José de Freitas. Os demais subgrupos são homogêneos, ou seja, o subgrupo 1d é formado apenas por

indivíduos de Altos e os subgrupos 1e e 1f são formados por indivíduos de José de Freitas.

Pela análise de *bootstraps* os agrupamentos, entre pares de acessos, que apresentaram maior consistência foram BGS 14 e BGS 15; BGS 10 e BGS 11; BGS 8 e BGS 12 com 99%, 94% e 60% respectivamente.

4.4 Estrutura e diferenciação genética

O valor de G_{st} estimado foi de 0,1850 e o de *theta S* foi de 0,2746. De outro modo, pode-se dizer que existem 18,50% da variação genética entre as populações pelo valor de G_{st} e 27,46% de variação entre as populações calculada pelo valor de *theta S*. Assim, ambas as estimativas, indicam maior variação dentro das populações. O coeficiente *theta P* mostrou que a diferenciação genética entre os grupos (Piauí e Maranhão) não foi significativa. O fluxo gênico (N_m), estimado com base na estimativa G_{st}, resultou no valor de 1,10 indivíduos, indicando uma alta taxa de migrantes entre as populações (Tabela 7).

Tabela 7 - Estrutura hierárquica, diferenciação genética e fluxo gênico em três populações naturais de *L. pisonis*.

	H_T	H_S	H_R	G_{st}	Theta S	Theta P	N_m
Estimativas	0,2833	0,2309	0,1997	0,1850	0,2746	0,0892*	1,10
Superior (IC 95%)	-	-	-	-	0,3570	0,1898	-
Inferior (IC 95%)	-	-	-	-	0,1903	-0,0310	-
Desvio padrão	0,0212	0,0130	0,0122	-	0,0415	0,0557	-

H_T = heterozigosidade média total; H_S = heterozigosidade média dentro das populações; H_R = heterozigosidade média dentro das regiões; Gst = índice de diferenciação genética entre as populações (Nei, 1987); Theta S = diferenciação genética entre as populações (Cockerham, 1969); Theta P = diferenciação genética entre as regiões (Cockerham, 1969); N_m = fluxo gênico; IC = intervalo de confiança; * = Não significativo.

Assim como os índices G_{st} e theta S, a AMOVA (Tabela 8) também revelou que a maior porcentagem da variabilidade genética de *L. pisonis* foi observada dentro das populações. Quando todas as populações são hierarquizadas como um único grupo, observa-se que 10,66% da variação genética ocorreu entre as populações.

Tabela 8 - Resultado da AMOVA para as três populações de *L. pisonis* com base em marcadores ISSR.

Fonte de variação	Soma dos quadrados	Variação	Variação (%)	Índices de fixação	p
Entre populações	37,92	1,71	10,66	Φ_{ST} = 0.1066	0.011*
Dentro de populações	86,30	14,38	89.33		

p = Teste de significância (10100 permutações); * = significativo.

Os resultados da análise Bayesiana executada no programa *Structure* indicam que os agrupamentos não estão fortemente correlacionados com as populações geográficas, isso porque indivíduos de uma mesma população geográfica apresentaram grande probabilidade de pertencerem a mais de um grupo.

Entre os valores de K testados o *software Structure* indicou o valor mais provável de K=4 (Figura 6). Esse valor geralmente é o que melhor explica a subdivisão populacional.

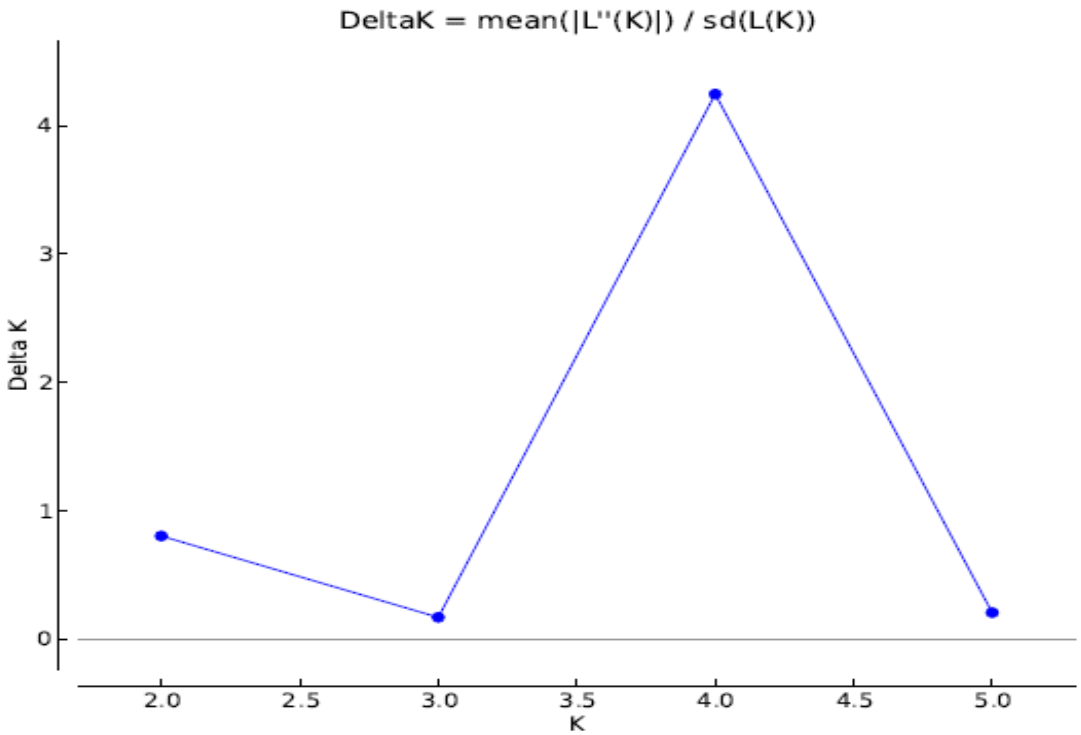

Figura 6. Valores de ΔK, calculado de acordo com o proposto por Evanno et al. (2005). O maior valor de ΔK corresponde ao K ótimo.

No teste de atribuição considerando o melhor valor de K (K= 4), os indivíduos estruturam-se formando quatro grupos independentemente da localização geográfica (Figura 7).

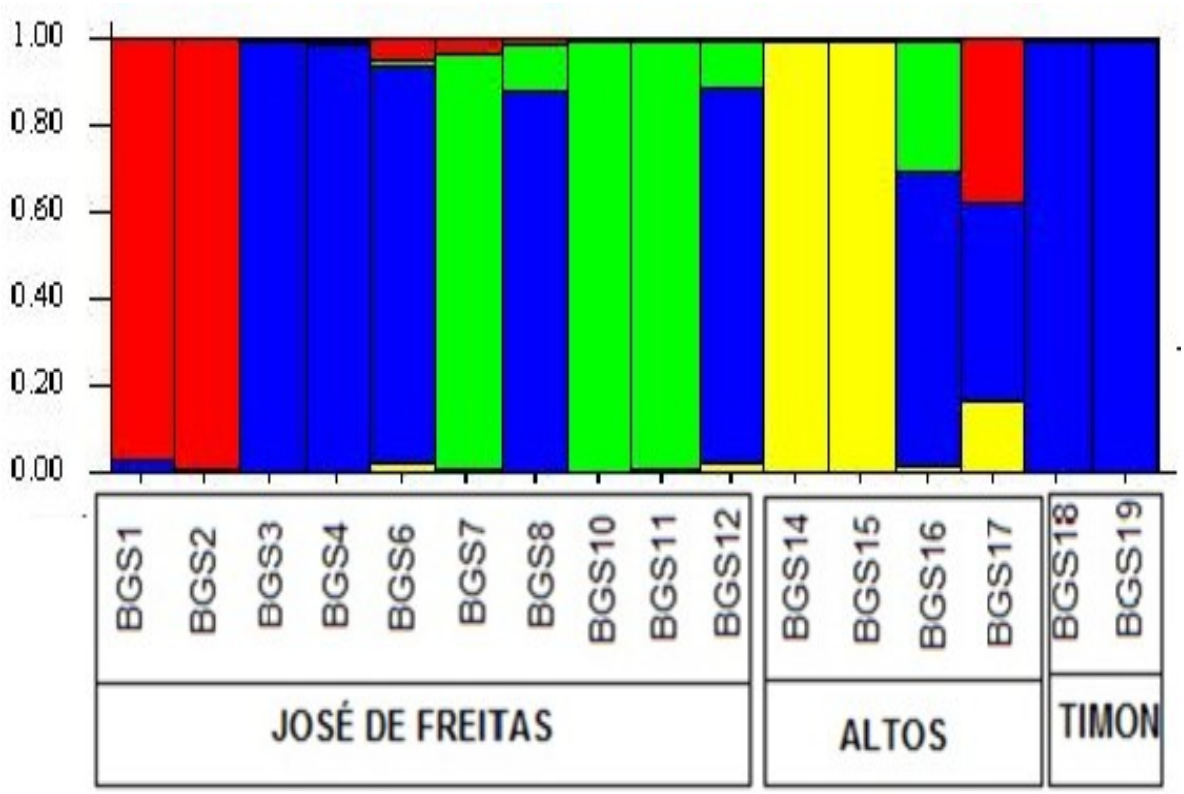

Figura 7. Teste de atribuição para os acessos de sapucaia avaliados com K=4. As barras verticais coloridas representam os acessos amostrados. A mesma cor em diferentes populações indica que pertencem ao mesmo grupo. Cores diferentes na mesma barra indicam a porcentagem do genoma compartilhado com cada grupo. No eixo Y estão os valores de Q: coeficiente de participação de um indivíduo em um determinado grupo, e no eixo X, as populações.

Desta forma, os indivíduos da população de José de Freitas tendem a formar três grupos geneticamente distintos, onde os indivíduos BGS 1 e BGS 2 formam o grupo 1 (cor vermelha); os indivíduos BGS 7, BGS 10, BGS 11 compõem o grupo dois (cor verde); o grupo 3 (cor azul), agrupando metade dos indivíduos de José de Freitas juntamente com os de Timon. Dois indivíduos de Altos (BGS 14, e BGS 15) se diferenciaram dos demais para formar um quarto grupo (cor amarela). O acesso BGS 17 pertencente a Altos, mostrou indicação de ocorrência de mistura de genomas, com até 83,3% do genoma composto por alelos em frequências parecidas a pertencerem aos grupos 1 (37,6%) e 3 (45,7%). Da mesma forma, o indivíduo BGS

16, apresentou ascendência mista com 67,9% de probabilidade de pertencer ao grupo três.

Pelo teste de migrante, realizado no programa *Structure*, todos os indivíduos mostraram alta probabilidade de possuírem ascendência pura na população previamente informada (Figura 8).

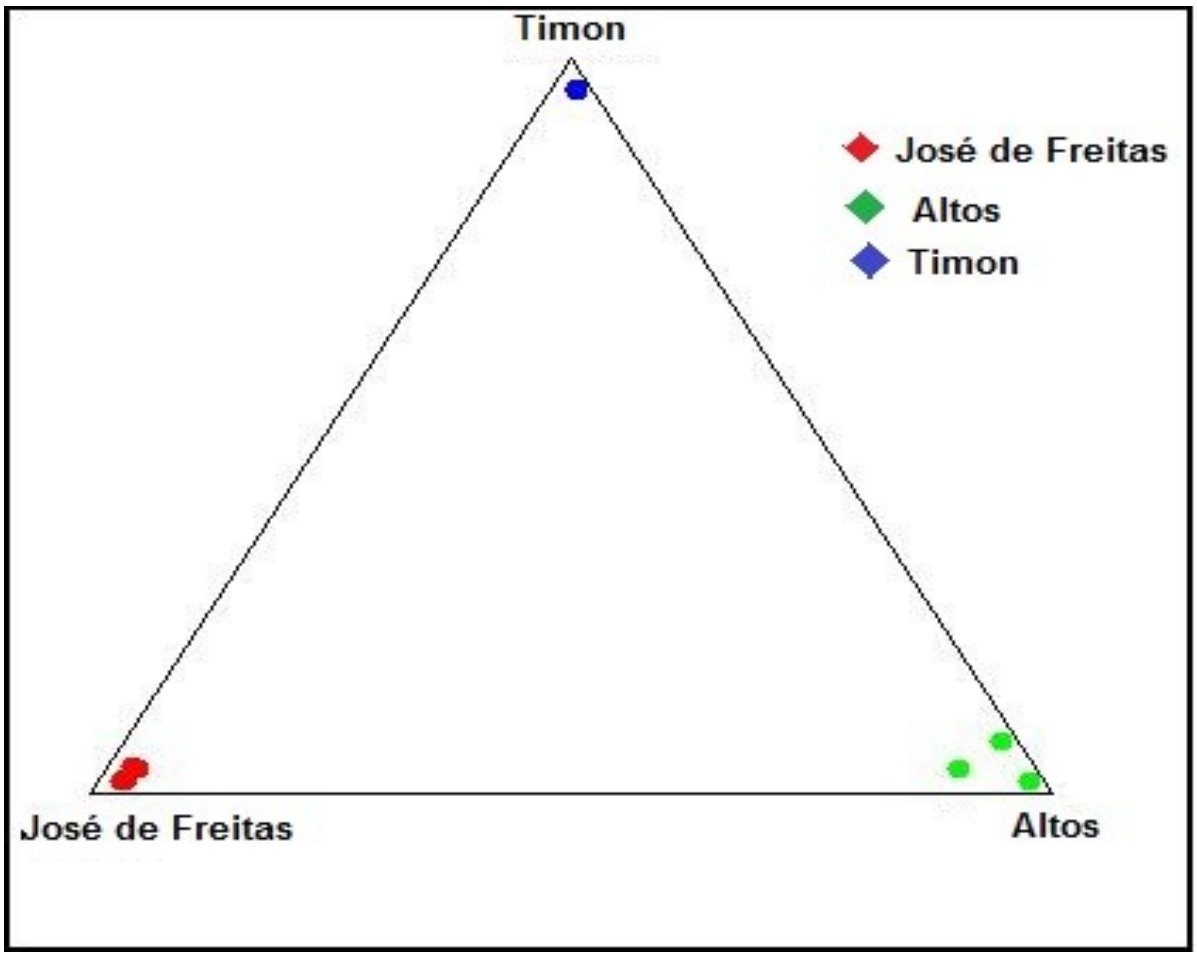

Figura 8 – Teste de agrupamento pelo *software Structure* usando a função USEPOPINFO; MIGRPRIOR = 0,05 e GENSBACK = 3. As figuras coloridas indicam a ascendência média estimada para os indivíduos de cada população.

5 DISCUSSÃO

A Coleção de Germoplasma de sapucaia mantida na Embrapa Meio Norte, apresentou alto polimorfismo (94,79%), com base em marcadores ISSR. Santos, F. (2010) também encontrou alto polimorfismo (82,0%) para este mesmo BAG, utilizando marcadores RAPD. Isso mostra o elevado grau de diversidade genética desta espécie, já que a proporção de locos polimórficos é um fator muito usado para quantificar a diversidade genética em populações de plantas (BORÉM; CAIXETA, 2009). Certamente, o baixo polimorfismo encontrado na população de Timon - MA (20,83%) pode estar relacionado com o tamanho da amostra, uma vez que só foram amostrados dois acessos representantes desta localidade.

Os resultados encontrados pelos índices de diversidade de Nei (h = 0,31; H = 0,32) e pelo índice de *Shannon* (I = 0,47) confirmaram alta diversidade genética no germoplasma de sapucaia. Isso sugere que o BAG de sapucaia da Embrapa Meio-Norte pode contribuir para a manutenção da variabilidade genética desta espécie. De outra forma, esta variabilidade existente pode ser usada em programas de melhoramento genético. Em espécies arbóreas perenes, a média de diversidade genética é de h = 0,140 (HAMRICK; GODT, 1989). No germoplasma de sapucaia, os menores índices de diversidade foram encontrados na população de Timon - Maranhão (h = 0,08; H = 0,10; I = 0,12). No entanto, essa não é a verdadeira variabilidade genética desta população, devido ao pequeno número de genótipos amostrados. Este fato sugere a necessidade de se realizarem novas coletas para que esta população esteja bem representada no BAG de sapucaia.

Estudos genéticos em sobreviventes de *C. pumila*, uma espécie de castanha Americana devastada por uma doença fúngica severa, indicaram um nível mais baixo de diversidade genética (He = 0,151) do que os registrados para castanheiras da China (He = 0,328) e da Europa (He = 0,317) (DANE; HAWKINS; HUANG, 1999). Mesmo assim, a variabilidade em *C. pumila* está acima da média estabelecida por Hamrick e Godt (1989). Isso reforça a explicação de que os baixos índices de diversidade e a baixa porcentagem de polimorfismo na população de Timon, estejam relacionados com o pequeno número de genótipos amostrados.

O agrupamento pela distância genética de Nei (UPGMA) mostrou uma maior similaridade genética (87,49%) entre as populações de José de Freitas e Altos que foram agrupadas com 81,1% de consistência (Tabela 6; Figura 4). Esse fato pode

estar relacionado com a maior proximidade geográfica entre as duas populações. Embora a população de Timon, esteja geograficamente mais distante das outras duas e tenha o rio Parnaíba como barreira geográfica, as distâncias genéticas encontradas entre a população de Timon e as demais não foi elevada (Tabela 6). Assim, existe muita variabilidade genética no germoplasma de sapucaia e as populações estudadas mostraram-se bastante semelhantes geneticamente.

O isolamento por meio da distância geográfica é um mecanismo que pode ser utilizado para explicar a subdivisão populacional em plantas (Wright, 1951). Neste estudo, não foi possível fazer a correlação entre a similaridade genética e a distância geográfica para as populações de sapucaia. No entanto, o que se sabe é que a sapucaia é uma espécie arbórea de vida longa, preferencialmente alógama, com dispersão de pólen por abelhas e sementes dispersas por animais e pelo homem (MORI; PRANCE, 1990; SAMPAIO, 2000).

Bandeira; Deimling e Kraemer (2010), por meio de marcadores ISSR, avaliaram a variabilidade genética de 25 progênies em cinco populações naturais de jaborandi (*Pilocarpus pennatifolius* Lemaire), um arbusto perene da flora nativa brasileira. Os dados de distância genética mostraram elevada semelhança genética e as maiores identidades não ocorreram necessariamente entre as populações mais próximas geograficamente. Rivas et al. (2013) avaliaram por meio de ISSR populações de cupuí (*Theobroma subincanum* Mart.), uma espécie com elevado potencial econômico e que ainda não foi domesticada e relataram que as populações com maior similaridade genética foram as mais próximas geograficamente, sugerindo que a estrutura genética desta espécie é determinada pela distância geográfica. Esses resultados mostram que nem sempre, a distância genética está correlacionada com a distância geográfica.

O dendrograma de agrupamento entre os acessos sugere que os indivíduos foram agrupados independentemente de sua origem geográfica (Figura 5). Isso pode ser percebido pelo fato de que, dos quatro grupos formados, mais de 80% dos acessos pertenceram a um único grupo contendo indivíduos de todas as populações. Resultados semelhantes foram observados por Santos, F. (2010). Este fato pode estar relacionado com as semelhanças genéticas entre as populações. No entanto, com 56% de similaridade genética é possível definir vários subgrupos (Figura 5). Isso evidencia a alta diferenciação genética encontrada dentro das populações. Alguns agrupamentos observados neste trabalho, estão de acordo com

um estudo de caracteres fenotípicos, com este mesmo BAG, realizado por Souza, V. et al. (2008), que verificaram a formação de quatro grupos e elevada variabilidade fenotípica entre os acessos de sapucaia para a maioria das características analisadas.

É interessante observar que em comparação com os resultados de Souza, V. et al. (2008), alguns pares de acessos mais divergentes, apresentaram características contrastantes de importância para o melhoramento genético. Por exemplo, os acessos BGS2 e BGS4; BGS1 e BGS7 são contrastantes para o número de amêndoas por frutos enquanto que BGS 2 e BGS 14 são contrastantes para o peso total da amêndoa.

A variabilidade encontrada no Germoplasma de sapucaia pode servir de base inicial para futuros programas de melhoramento genético e como os genótipos mais divergentes foram o BGS 2 e BGS 4, ambos da população de José de Freitas, estes podem constituir bons genitores, uma vez que, Bered (1999), destaca que através da variabilidade existente, o melhorista pode escolher como genitores de um cruzamento, indivíduos distantes geneticamente, a fim de formar combinações híbridas e maximizar a expressão de características desejáveis o que contribui para a ampliação da variância genética em populações segregantes.

Os índices de diferenciação genética juntamente com a AMOVA ao revelarem que a maior diferenciação ocorre dentro das populações estão de acordo com o descrito por Hamrick e Godt (1996). Estes sugeriram que o sistema reprodutivo é importante para a formação da estrutura genética de uma população e, em geral, as espécies preferencialmente alógamas têm alta diferenciação dentro das populações e menor diferenciação entre as populações. Isso pode explicar os resultados encontrados para as populações de sapucaia uma vez que essa espécie é preferencialmente alógama e tem como principal polinizador as abelhas (MORI; PRANCE, 1990).

No geral, as populações podem estar subdivididas em unidade menores. No entanto, na natureza, as subpopulações não estão necessariamente isoladas geneticamente umas das outras em virtude da ocorrência de migração ou troca de genes entre elas (HARTL; CLARK, 2010). Com isso, a maior diferenciação genética intrapopulacional observada para sapucaia está dentro do esperado, por serem plantas preferencialmente alógamas e pelo índice de fluxo gênico encontrado (N_m = 1,10).

Segundo Hamrick (2012), o fluxo gênico entre indivíduos dentro das populações e entre as populações podem ter grandes impactos sobre a distribuição da variação genética. Isso porque o fluxo gênico, ao introduzir novas variações, reduz a diferenciação genética entre as populações. Na verdade, um indivíduo migrante ou mais entre as populações é suficiente para prevenir a diferenciação substancial entre elas. Geralmente, se $N_m > 1$, então haverá pouca diferenciação entre as populações (MCDERMOTT; MCDONALD, 1993; WRIGHT, 1951). Assim, é necessária pequena migração para impedir a diferenciação significativa entre populações, quando medida pelo índice de fixação F_{ST}. No entanto, entre as populações, a diferenciação não é, de modo algum, insignificante, mesmo se F_{ST} for muito pequeno (WRIGHT, 1965). Além do mais, o fluxo gênico em função de F_{ST} é considerado um fluxo histórico, ou seja, não indica se está havendo obrigatoriamente fluxo gênico em determinado evento reprodutivo (LEVIN, 1984).

Desta forma, a diferenciação genética significativa de 10,66% (AMOVA, Tabela 8) encontrada entre as populações de sapucaia, mesmo tendo sido encontrado elevado fluxo gênico ($N_m = 1,10$), pode estar relacionada com processos aleatórios na transmissão dos alelos de uma geração para outra ou por diferenças casuais na frequência alélica entre os fundadores iniciais destas populações (HARTL e CLARK, 2010). Segundo Levin (1984), a colonização de habitats vazios pela dispersão de sementes tem como consequência alterações da estrutura espacial da espécie, conduzindo, muitas vezes, a alguma diferenciação genética entre as populações. Isso seria possível, já que as sementes de sapucaia costumam ser dispersas por ação antrópica.

Silva, V. et al. (2012) analisaram populações de Castanha do Brasil (*Bertholletia excelsa)* com marcadores ISSR e também observaram uma grande diferenciação genética (74,29%) dentro das populações. Da mesma forma, Buckley (1988) encontrou apenas 3,75% de diferenciação entre duas populações dessa castanheira. Isso mostra que os dados de diferenciação genética entre as populações de sapucaia são semelhantes ao esperado para castanheiras da mesma família. Hamrick (1991) relatou que em espécies arbóreas, preferenciamente alógamas, e com padrões de dispersão de pólen e sementes pelo vento ou por animais, em que a grande quantidade de variação genética é mantida por extenso fluxo gênico sobre grandes áreas geográficas, o esperado é que a diferenciação genética seja maior dentro das populações, como foi observado neste trabalho.

Em concordância com o agrupamento UPGMA, na análise de agrupamento feita pelo programa *Structure,* os três grupos geográficos de sapucaia se estruturam em quatro grupos independentemente das populações de origem em decorrência do fato em que alguns indivíduos apresentaram alta probabilidade de serem originários de um local diferente daquele em que foram coletados (Figura 7). Isso poderia estar relacionado com a possibilidade de alguns indivíduos serem imigrantes ou descendentes de imigrantes recentes. No entanto, o teste de migrante agrupou fortemente todos os indivíduos de acordo com sua população informada e a probabilidade de indivíduos de uma população possuir ancestrais em outra população foram baixas (Figura 8). Desta forma, a diferenciação genética entre indivíduos dentro das populações podem ser explicadas pelo sucesso no fluxo de pólen.

A observação de que não existe diferenciação genética significativa entre as populações dos dois estados (Tabela 7), está de acordo com o gráfico de agrupamento pelo coeficiente de Jaccard e com a análise Bayesiana realizada por meio do programa *Structure* (Figura 7), os acessos do Maranhão agruparam juntamente com acessos do Piauí. Assim, o elevado fluxo gênico encontrado poderia estar limitando a divergência genética neste nível hierárquico. No entanto, o número de indivíduos amostrados na população de Timon é muito pequeno e, portanto, pode não expressar a verdadeira diferenciação genética entre os dois grupos.

6 CONCLUSÕES

A Coleção de Germoplasma de sapucaia da Embrapa Meio-Norte apresenta potencial para uso em trabalho de melhoramento genético, diante da elevada diversidade genética, detectada por marcadores ISSR.

Os genótipo BGS 2 e BGS 4, que são os mais divergentes, podem constituir uma boa combinação para programas de melhoramento genético, já que estes genótipos, foram contrastantes para o número de amêndoas por frutos em análise fenotípica.

A maior diferenciação genética foi encontrada dentro das populações e existe diferenciação genética significativa entre elas. No entanto, devido ao pequeno número de acessos, é necessária a realização de novas coletas para que se mantenha, de forma representativa, a variabilidade genética das regiões de procedência dos acessos.

7 REFERÊNCIAS

AGRA, M. F.; FREITAS, P. F.; BARBOSA, F. J. M. Sinopse das plantas conhecidas como medicinais e venenosas no Nordeste do Brasil. **Revista Brasileira de Farmacognosia,** v.17, p.116-155, 2007.

APG III. Angiosperm Phylogeny Group III. An update of the Angiosperm Phylogeny Group Classification for the orders and families of flowering plants: APG III. **Botanical Journal of the Linnean Society**, v.161, p.105-121, 2009.

ARAÚJO, V. F.; PETRY, A. C.; ECHEVERRIA, R. M.; FERNANDES, E. C.; JÚNIOR, F. P. **Plantas da Amazônia para produção cosmética**: **uma abordagem química - 60 espécies do extrativismo florestal não-madeireiro da Amazônia**. Brasília, 2005. 244 p.

ATAÍDE, W. L. S.; SHIMIZU, E. S. C.; LEAO, N. V. M.; FELIPE, S. H. S. Biometria de frutos e sementes e germinação em diferentes temperaturas e substratos de *Lecythis pisonis* Cambess. In: SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA EMBRAPA, 2011, Belém-PA. **Anais...** Belém: EMBRAPA AMAZÔNIA ORIENTAL.

BALÉE, W. **Footprints of the Forest – Ka'apor ethnobotany – the historical ecology of Plant Utilization by an Amazonian people.** New York: Columbia University Press, 1994. 369 p.

BANDEIRA, A. J.; DEIMLING, L. I.; KRAEMER, J. E. G. Variabilidade genética do jaborandi (*Pilocarpus pennatifolius* Lemaire; Rutaceae) em populações naturais da região noroeste do Rio Grande do Sul. **Revista de Iniciação Científica da ULBRA,** 2009/2010.

BARRENECHE, T.; CASASOLI, M.; RUSSELL, K.; AKKAK, A.; MEDDOUR, H.; PLOMION, C.; VILLANI, F.; KREMER, A. Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). **Theoretical and Applied Genetics,** v.108, p.558–566, 2004.

BERED, F. Variabilidade genética: ponto de partida para o melhoramento de plantas. In: SACCHET, A. M. O. F. **Genética, para que te quero?** Porto Alegre: Ed. UFRGS, 1999. p.99-104.

BESPALHOK, F. J. C.; GUERRA, E. P.; OLIVEIRA, R. **Introdução ao Melhoramento de Plantas**. In: BESPALHOK, F. J. C.; GUERRA, E. P.; OLIVEIRA, R. Melhoramento de Plantas. Disponível em <www.bespa.agrarias.ufpr.br. > acesso em dez. de 2014.

BORÉM, A.; CAIXETA, E.T. **Marcadores Moleculares**. 2ª Ed. Viçosa, MG: Editora UFV, 2009. 532 p.

BRAGA L. F.; SOUSA, M. P.; GILBERT, S.; CARVALHO, M. A. C. Caracterização Morfométrica de Sementes de Castanha de Sapucaia (*Lecythis pisonis* Cambess - Lecythidaceae) **Revista de Ciências Agro-Ambientais**, Alta Floresta, v.5, n.1, p.111 - 116, 2007.

BUCKLEY, D. P.; O'MALLEY, D. M.; APSIT, V. Genetics of Brazil nut (*Bertholletia excelsa* HUMB. & BONPL.: Lecythidaceae). Genetic variation in natural populations. **Theoretical and Applied Genetics,** v.76, p.923–928, 1988.

BUSSAD, W. O.; MIAZAKI, E. S.; ANDRADE, D. F. **Introdução à análise de agrupamentos**. São Paulo: Associação Brasileira de Estatística, 1990, 105 p.

CARDOSO, V. J. M.; NAKAO, E. A. *Lecythis pisonis* Cambess (sapucaia): seed morphometric traits and germinative response Biotemas, n.27, p.23-28, 2014.

CARVALHO, I. M. M.; QUEIRÓS, L. D.; BRITO, L. F.; SANTOS, F. A.; BANDEIRA, A. V. M.; SOUZA A. L.; QUEIROZ, J. H. Caracterização química da castanha de sapucaia (*Lecythis pisonis* Cambess) da região da zona da mata mineira. **Bioscience Journal,** v.28, n.6, p.971-977, 2012.

CARVALHO, M. G.; SOUZA, V. A. B; ALVES, R. A. **Caracterização nutricional de nozes de chihá, castanhado - gurguéia e sapucaia, in natura e torradas.** Revista Ciências Agronômicas, v.39, n.4, p.517-523, 2008.

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

COLLEVATTI, R. G; BRONDNAI, R. V; GRATTAPAGLIA, D. Development and characterization of microsatellite markers for genetic analysis of a Brazilian endangered tree species. **Heredity,** v.83, p.748–756, 1999.

CORRÊA, M. P. **Dicionário das plantas úteis do Brasil e das exóticas cultivadas.** Colaboração de Leonan Azevedo Penna. Rio de Janeiro: Imprensa Nacional, 1984.

COSTA, T. S.; SILVA, A. V. C.; LÉDO, A. S.; SANTOS, A. R. F.; JÚNIOR, J. F. S. Diversidade genética de acessos do banco de germoplasma de mangaba em Sergipe. **Pesquisa Agropecuária Brasileira, Brasília**, v.46, n.5, p.499-508, 2011.

CROW, J. F.; AOKI, K. Group selection for a polygenic behavionral trait: estimating the degree of population subdivision. **Proc. Natl. Acad. Sci**. **USA** v.81, p.6073-6077, 1984.

CRUZ, C. D.; CARNEIRO P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: Ed. UFV, v.2, 2003, 585p.

DANE, F.; HAWKINS, L. K.; HUANG, H. Genetic Variation and Population Structure of *Castanea pumila* var. ozarkensis. **Journal of the American Society for Horticultural Science**, v.124, n.6, p.666–670, 1999.

DELGADO, A. A. Uso de marcadores moleculares en plantas; aplicaciones en frutales del trópico. **Agronomía Mesoamericana**, v.17, n.2, p.221-242, 2006.

DNEASY plant mini kit for miniprep purification of total cellular DNA from plant cells and tissues, or fungi. [São Paulo]: QIAGEN, 2006.55p. (Dneasy Plant Handbook, 7).

DURATEX S/A (São Paulo, BR). **Árvores do Brasil**. São Paulo: Prêmio, 1989. 118p.

ESCORCIO, S. P.; BARBOSA, A. S.; CHAVES, M. H.; MOITA NETO, J. M. Caracterização e composição química da fração lipídica de amêndoas de *Lecythis pisonis* Camb. In: CONGRESSO BRASILEIRO DE QUÍMICA, 39, 1999. Goiânia. **Química para o desenvolvimento sustentado (Resumos...)**. Goiânia: ABQ – Regional de Goiás. 1999. p.104.

EVANNO, G.; REGNAUT, S.; GOUDET, J. Detecting the number of clusters of individuals using the *software Structure*: a simulation study. **Molecular Ecology**, v.14, n.8, p.2611-20. 2005.

EXCOFFIER, L.; LAVAL, G.; SCHNEIDER. S. **Arlequin ver. 3.01**. An Integrated *Software* Package for Population Genetics Data Analysis. Computational and Molecular Population Genetics Lab (CMPG). Institute of Zoology, University of Bern, Bern, 2006.

EXCOFFIER, L.; SMOUSE, P. E.; QUATTRO, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. **Genetics**, v.131, p.479-491, 1992.

FALEIRO, F. G. **Marcadores genético-moleculares aplicados a programas de conservação e uso de recursos genéticos.** 1.ed. Planaltina: Embrapa Cerrados, 2007. 99 p.

FALUSH, D.; STEPHENS, M.; PRITCHARD, J. K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. **Genetics**, v.164, p.1567–1587, 2003.

FALUSH, D.; STEPHENS, M.; PRITCHARD, J. K. Inference of population structure using multilocus genotype data: dominant markers and null alleles. **Molecular Ecology Notes**, v.7, p.574–578, 2007.

FAO. **Food and fruit-bearing.** Forest species 3: examples from Latin America. Rome,1986.

FERRÃO, J. E. M. **Fruticultura tropical**: **espécies com frutos comestíveis**. 2 v. Lisboa: Instituto de Investigação Científica Tropical, 2001. 580 p.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3.ed. Brasília: Embrapa-Cenargen, 1998. 220p. (Embrapa-Cenargen. Documentos, 1998).

FRANCO, E. A. P.; BARROS, R. F. M. Uso e Diversidade de plantas Medicinais no Quilombo Olho D'água dos Pires, Esperantina, Piauí. **Revista Brasileira de Medicina,** n.8, p.78-88, 2006.

FREELAND, JR. **Molecular ecology**. John Wiley & Sons, Southern Gate, 2005. 388p.

FREITAS, J. B.; NAVES, M. M. V. Composição química de nozes e sementes comestíveis e sua relação com a nutrição e saúde. **Revista de Nutrição**, Campinas, v.23, n.2, p.268-279, 2010.

FUTUYMA, D. J. Biologia Evolutiva. 2 ed. Ribeirão Preto: FUNPEC-RP, 2002, 631p.

GOLDSTEIN, D. B.; LINARES, A. R.; CAVALLI-SFORZA, L. L.; FELDMAN, M. W. An evaluation of genetic distances for use with microsatellite loci. **Genetics**, v.139, p. 71-463, 1995.

GOMES, R. P. **Fruticultura brasileira**. São Paulo: Nobel, 1977. 448 p.

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

GUPTA, M., CHYI, Y-S. ROMERO-SEVERSON J.; OWEN, J. L. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. **Theoretical and Applied Genetics**, v.89, p.998–1006, 1994.

HAMMER, O.; HARPER, D. A. T.; RYAN, P. D. PAST: Paleontological Statistics *Software* Package for Education and Data Analysis. **Palaeontologia Electronica**, v. 4, 2001. 9p.

HAMRICK, J. L. Gene movement in tropical tree populations tropical breeding systems: one and done? **Heredity**, v.109, p.330–331, 2012.

HAMRICK, J. L.; GODT, M. J. W. Allozyme diversity in plant species. p. 43-63 in Brown, A. H. D., Clcgg, M. T., Kahler, A. L. & Weir, B. S. (eds). **Plant population genetics, breeding and germplasm resources**. Sinauer, Sunderland, Mass. 1989. 449p.

HAMRICK, J. L.; GODT, M. J. W. Conservation Genetics of Endemic Plant Species. In: Avise, J. C.; Hamrick, J. L. (Eds). **Conservation Genetics, Case Histories From Nature**. New York: Chapman and Hall, 1996. p.281- 304, 1996.

HAMRICK, J. L.; MURAWSKI, D. A. Levels of allozyme diversity in populations of uncommon Neotropical tree species. **Journal of Tropical Ecology**, v.7, p.395-399, 1991. Published online: 10 Jul., 2009.

HARTL, D. L.; CLARK, A. G. **Princípios de genética de populações.** 4. ed. Artmed, 2010. 659p.

HOMMA, A. K. O. **Cemitério das Castanheiras**. Ciência Hoje. São Paulo, v.34, nº 202. Março: 2004.

KREMER, A. A. U.; CARON, H. A. U.; CAVERS, S. A. U; COLPAERT, N. A. U; GHEYSEN, G. A. U.; GRIBEL, R. A. U; LEMES, M. A. U.; LOWE, A. J. A. U.; MARGIS, R. A. U.; NAVARRO, C. A. U.; SALGUEIRO, F. T. I.; Monitoring genetic diversity in tropical trees with multilocus dominant markers. **Heredity**, v. 95, p. 274–280, 2005.

LEVIN, D. A. Imigration in plants: an exercise in the subjunctive. In: DIRZO, R. & SARVKHAN, J. **Perspectives on plant population ecology**. Sunderland, Sinauer, 1984, p.242-6.

LEWONTIN, R. C. The apportionment of human diversity. **Evolutionary biology**, v. 6, p.381-398, 1972.

LING, Y.; ZHANG, X.Q.; MA, X.; CHEN, S.Y.; CHEN, T.T.; LIU, W. Assessment of genetic diversity of bermudagrass germplasm from southwest China and Africa by using AFLP markers. **Genetics and Molecular Research,** v.14, n.1, p.1748-1756, 2015.

LORENZI, H. Árvores Brasileiras: **Manual de Identificação e cultivos de plantas arbóreas do Brasil**. 2ª Ed. São Paulo: Nova Odessa. 2002.

LYNCH, M.; MILLIGAN, B. G. Analysis of population genetic structure with RAPD markers. **Molecular Ecology**, v.3, p.91–9, 1994.

MADESIS, P.; ABRAHAM, E. M.; KALIVAS, A.; GANOPOULOS, I.; TSAFTARIS, A. Genetic diversity and structure of natural *Dactylis glomerata* L. populations revealed by morphological and microsatellite-based (SSR/ISSR) markers. **Genetics and Molecular Research**, v.13, n. 2, p.4226-4240, 2014.

MALIK, W.; ASHRAF, J.; IQBAL, M. Z.; KHAN, A. A.; QAYYUM, A.; ABID, M. A.; NOOR, E.; AHMAD, M. Q.; ABBASI, G. H. Molecular Markers and Cotton Genetic Improvement: Current Status and Future Prospects. **Scientific World Journal, 2014.**

MARTINS, L.; COUTINHO, E. L.; PANZANI, C. R.; XAVIER, N. J. D. **Frutas nativas do Brasil e exóticas.** Campinas: CATI, 2002. 112p.

MATTA, A. A. **Flora Médica Brasiliense.** 3. ed. Revista – Manaus: Editora Valer e Governo do Estado do Amazonas, 2003. 356 p. (Série Poranduba, 3).

MATTA, L. B. V.; SCUDELLER, V. V. Lecythidaceae Poit. in the Tupé Sustainable Development Reserve, Manaus, Brazil. **Brazilian Journal of Botany**, São Paulo, v.35, n.2, 2012.

MCDERMOTT, J. M.; MCDONALD, B. A. Gene flow in plant pathosystems. **Annual Review of Phytopathology,** v.31, p.353-73, 1993.

MELCHINGER, A. E. Genetic diversity and heterosis. In: COORS, J. G.; PANDEY, S. (Ed.). The genetics and exploitation of heterosis in crops. **Crop Science of America**. p.99-118, 1999.

MICHALAKIS, Y.; EXCOFFIER, L. A generic estimation of population subdivision using distances between alleles with special reference to microsatellite loci. **Genetics** v.142, p.1061-1064, 1996.

MILLER, M. P. **Tools for Population Genetic Analysis (TEPGA) Version 1.3**. Department of Biological Sciences Northern Arizona University, Flagstaff, 1997.

MILLIKEN, W. MILLER, R. P.; POLLARD, S. R.; WANDELLI, E. V. **The ethnobotany of the Waimiri Atroari indians of Brazil.** Kew: Royal Botanic Garden, 1986. 146p.

MORI, S.A. Biologia da polinização em Lecythidaceae. **Acta Botânica Brasílica**, v.1, n.2, p.121-124, 1988.

MORI, S. A. Diversificação e conservação das Lecythidaceae neotropicais. **Acta Botânica Brasílica**, v.4, n.1, p.45-68, 1990.

MORI, S. A.; MATTOS S. L. A.; SANTOS, T. S. Observações sobre a fenologia e biologia floral de *Lecythis pisonis* Cambess. (Lecythidaceae). **Revista Theobroma (Brasil),** v.10, n.3, p. 11-103, 1980.

MORI, S. A.; PRANCE, G. T. Lecythidaceae part II: the zygomorphic-flowered new world genera (*Couroupita, Corythophora, Bertholletia*, *Couratari, Eschweilera, &*

Lecythis), with a study of secondary of neotropical lecythidaceae. New York: The NewYork Botanical Garden, 1990. 375 p. (Flora Neotropica**.** Monograph, 21 II).

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

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

NEI, M. Genetic distance between populations. **American Naturalist**,v.106, p.283-292, 1972.

NEI, M. **Molecular Evolutionary Genetics**. Columbia University Press, New York, 1987.

PANTANO, A. P. Sapucaia: (*Lecithis ollaria ou L. pisonis*)**.** Disponível em: <http://www.jardimdeflores.com.br/sapucaia.html>. Acesso em 28 out. 2014.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, Washington, v.155, n.2, p.945-959, 2000.

RAMALHO, M. A. P.; SANTOS, J. B.; PINTO, C. A. B. P. **Genética na agropecuária**. 5. Ed. Lavras: UFLA, 2012. 566p.

REDDY, M. P.; SARLA, N.; SIDDIQ, E. A. Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. **Euphytica,** v.128, p.9–17, 2002.

REDE DE SEMENTES DO CERRADO. Disponível em: <http://www.rededesementesdocerrado.com.br/Espécies/Lecythidaceae/pisonis_2708/index.html> . Acesso em: 02 jan. 2014.

REVILLA, J. **Plantas úteis da Bacia Amazônica.** Manaus: INPA/ SEBRAE, 2002.
RIBEIRO, R. A.; RODRIGUES, F. M. Genética da conservação em espécies vegetais do cerrado. **Revista de Ciências médicas e biológica**, Salvador, v.5, n.3, p.253-260, set./dez. 2006.

RIVAS, L. H.; GIUSTINA, L. D.; LUZ, L. N.; KARSBURG, I. V.; PEREIRA, T. N. S.; ROSSI, A. A. B. Genetic diversity in natural populations of *Theobroma subincanum*

Mart. in the Brazilian Amazon**. Genetics and Molecular Research,** v.12, n.4, p. 4998-5006, 2013.

ROHLF, F. J. Numerical taxonomy and multivariate analysis system. New York, 1992. 470p.

SAMPAIO, P. T. B. Castanha-sapucaia (*Lecythis pisonis)*. In: CLAY, J. W.; SAMPAIO, P. T. B.; CLEMENT, C. R. Biodiversidade amazônica: exemplos e estratégias de utilização. Manaus: Programa de Desenvolvimento Empresarial e tecnológico, 2000. p.141-147.

SANTOS, F. M. G. Caracterização molecular de germoplasma de sapucaia (*Lecythis pisonis* Cambess) por meio de marcadores RAPD [Monografia]. Teresina: Faculdade NOVAFAPI, curso de Biomedicina, 2010.

SANTOS, J. C.; SENA, A. S.; ROCHA, C. I. L. DA. Competitividade brasileira no comércio internacional de castanha-do-Brasil. In: CONGRESSO SOBER, 48., 2010, Campo Grande. **Anais**... Campo Grande, 2010. p.1-14.

SANTOS, M. F.; SOUSA, C. C.; CARVALHAES, M. A.; SILVA, K. J. D.; LIMA, P. S. C. Similaridade entre genótipos de babaçu (*Orbignya phalerata* mart.) estimada por marcadores RAPD. In: I Congresso de Genética do Centro-Oeste. Brasília, 2008.

SILVA, C. B. M. C**.;** SANTOS, E. S. L.; AMORIM, J. S. E. S. L.; CORRÊA, R. X.; SOUZA, A.P. Genetic Breeding and Diversity of the Genus *Passiflora*: Progress and Perspectives in Molecular and Genetic Studies. **International Journal of Molecular Sciences**, v.15, n.8, p.14122–14152, 2014.

SILVA, L. L.; OLIVEIRA, J. P. C.; GOMES, B. S.; SOUSA-NETOA, B. P.; FERREIRA, E. L. F.; CHAVES, M. H.; OLIVEIRA, F. A. Chemical constituents of *Lecythis pisonis* and cytotoxic activity. **Revista brasileira de farmacognosia**. v.22, n.5 , p.1140-1144, 2012.

SILVA, V. S.; MARTINS, K.; CAMPOS, T.; WADT, L. H. O. Diversidade genética de populações naturais de castanheira (*Bertholletia excelsa*) com marcadores ISSR. In: **CONGRESSO BRASILEIRO DE RECURSOS GENÉTICOS**, 2., 2012, Belém, PA. Anais... Brasília, DF: Sociedade Brasileira de Recursos Genéticos, 2012. 1 CD-ROM.

SIQUEIRA, E. R. de; RIBEIRO, F. E. (editores). **Mata Atlântica de Sergipe.** Aracaju: Embrapa Tabuleiros Costeiros, 2001, 132 p.

SLATKIN, M. Gene flow and the geographic structure of natural population. **Science**, v.236, p.787-792, 1987.

SMITH, N. P.; MORI, S. A.; PRANCE, G. T. Lecythidaceae. In: **Lista de Espécies da Flora do Brasil**. Jardim Botânico do Rio de Janeiro, 2010.

SMYTH, C. A.; HAMRICK, J. L. Realized gene flow via pollen in artificial populations of musk thistle, *Carduus nutans* L. **Evolution**, v. 41, p. 613- 619, 1987.

SOARES, C.B.L.V. **Árvores nativas do Brasil.** Rio de janeiro: Salamanca, 1990.

SOUZA, A. G. C.; SOUZA, N. R.; SILVA, S. E. L.; NUNES, C. D. M.; CANTO, A. C.; CRUZ, L. A. A. **Fruteiras da Amazônia.** Brasília: EMBRAPA-SPI/Manaus: EMBRAPA-CPAA, 1996, 204 p. (Coleção Biblioteca Botânica Brasileira, 1).

SOUZA, I. G. B. ; VALENTE, S. E. S. ; BRITTO, F. B. ; SOUSA, V. A. B. ; LIMA, P.S.C. RAPD analysis of the genetic diversity of mango (*Mangifera indica*) germplasm in Brazil. **Genetics and Molecular Research,** v.10, n.4, p. 3080-3089, 2011.

SOUZA, I. G. B.; SOUZA, V. A. B.; LIMA, P. S. C. Molecular characterization of *Platonia insignis* Mart. (''Bacurizeiro'') using inter simple sequence repeat (ISSR) markers. **Molecular Biology Reports,** 2012. DOI 10.1007/s11033-012-2462-6.

SOUZA, M. H., MAGLIANO, M. M., CAMARGOS, J. A. A., SOUZA, M. R. **Madeiras Tropicais Brasileiras**. MMA/LPF, Brasília, 1997.

SOUZA, V. A. B.; CARVALHO, M. G.; SANTOS, K. S.; FERREIRA, C. S. Características físicas de frutos e amêndoas e características químico-nutricionais de amêndoas de acessos de sapucaia. **Revista Brasileira de Fruticultura**, Jaboticabal, v.30, n.4, p.946-952, 2008.

TEIXEIRA, E. **Frutas do Brasil:** sapucaia. Disponível em: <http:/ /www.bibvirt.futuro.usp.br/especiais/frutasnobrasil/sapucaia.htm>. Acesso em: 20 de out. 2014.

TERRA, C. G. **O Jardim no Brasil do século XIX***: Glaziou revisitado.* Rio de Janeiro: UFRJ/EBA, 1993.

TORRES, A. C.; FERREIRA, A. T.; SÁ, F. G.; BUSO, J. A.; CALDAS, L. S.; NASCIMENTO, A. S.; BRÍGIDO, M. M.; ROMANO, E. **Glosário de Biotecnologia Vegetal**. Brasília: EMBRAPA/CNPH, 2000. 128p.

UNITED STATES DEPARTMENT OF AGRICULTURE - USDA. Agricultural Research Service – ARS, National Genetic Resources Program. Germplasm Resources Information Network - (GRIN) [Base de Dados Disponível na Internet]. National Germplasm Resources Laboratory, Beltsville, Maryland. Disponível em:< http://www.ars-grin.gov2/cgibin/npgs/html/index.pl> . Acesso em: 9 dez. 2014.

VENCOVSKY, R. Análise de variância de frequências alélicas. **Revista Brasileira de Genética**, v.15, p.53-60, 1992.

VIEIRA, R. F.; MARTINS, M. V. M. Recursos genéticos de plantas medicinais do cerrado: uma compilação de dados. **Revista Brasileira de Plantas Medicinais,** Botucatu, v.3, n.1, p.13-36, 2000.

VIJAYAN, K. Inter Simple Sequence Repeat (ISSR) Polymorphism and Its Application in Mulberry Genome Analysis. **International Journal of Industrial Entomology,** v.10, n. 2, p.79-86, 2005.

VILLACHICA, H. **Frutales y hortalizas promisorios de la Amazonia.** Lima: TCA/FAO/DGIS/PAUD/ICRAF/FIDA/PNUMAIICA/GTZ/UNAMAZ, 1996. 367 p.

WANG, Z.; LIAO, L.; YUAN, X.; GUO, H.; GUO, A.; LIU, J. Genetic diversity analysis of *Cynodon dactylon* (bermudagrass) accessions and cultivars from different countries based on ISSR and SSR markers. **Biochemical Systematics and Ecology**. v.46, p.108-115, 2013.

WEIR, B. S.; COCKERHAM, C. C. Estimating F-statistics for the analysis of population structure. **Evolution,** v.38, n.6, p.1358-1370, 1984.

WICKENS, G. E. **Edible nuts.** Roma: FAO, 1995. 198 p. (Non Wood Forest Products,. 5).

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

WONGSAWAD, P.; PEERAPORNPISAL, Y. Molecular identification and phylogenetic relationship of green algae, *Spirogyra ellipsospora* (Chlorophyta) using ISSR and rbcL markers. **Saudi Journal of Biological Sciences** (2014), http://dx.doi.org/10.1016/j.sjbs.2014.01.003.

WRIGHT, S. The genetical structure of populations. **Annals of Eugenics**, v.15, p.323- 353, 1951.

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

YEH, F. C.; BOYLE, T. Y. Z.; XIYAN, J. M. **POPGENE version 1.31**: Microsoft Windows based freeware for populations genetic analysis. University of Alberta and Center for International Forestry Reseach, 1999. 29 p.

YU, J.; JING, Z. B.; CHENG, J. M. Genetic diversity and population structure of *Stipa bungeana*, an endemic species in Loess Plateau of China, revealed using combined ISSR and SRAP markers. **Genetics and Molecular Research**, v.13, n.1, p.1097-1108, 2014.

ZIETKIEWICZ, E.; RAFALSKI, A.; LABUDA, D. Genome fingerprinting by Simple Sequence Repeat (SSR) - anchored polymerase chain reaction amplification. **Genomics,** v.20, p.176-183, 1994.

ANEXOS

ANEXO A. Matriz de similaridade genética obtida pelo coeficiente de similaridade de Jaccard para os 17 genótipos de sapucaia

Genótipos	BGS1	BGS2	BGS3	BGS4	BGS6	BGS7	BGS8	BGS10	BGS11	BGS12	BGS13	BGS14	BGS15	BGS16	BGS17	BGS18	BGS19
BGS1	1	0,4	0,4	0,3088	0,3968	0,3279	0,4	0,3833	0,4	0,4308	0,3824	0,3582	0,3971	0,3881	0,4	0,3438	0,3636
BGS2	0,4	1	0,3279	0,2836	0,3934	0,3684	0,375	0,3793	0,4194	0,3433	0,3582	0,3333	0,3939	0,3433	0,3279	0,3387	0,3385
BGS3	0,4	0,3279	1	0,55	0,5862	0,4655	0,5079	0,5	0,5079	0,5397	0,5077	0,4844	0,5	0,4697	0,4426	0,5517	0,6491
BGS4	0,3088	0,2836	0,55	1	0,5397	0,3846	0,4925	0,4603	0,4925	0,4571	0,4928	0,4085	0,4857	0,4783	0,3286	0,4844	0,5231
BGS6	0,3968	0,3934	0,5862	0,5397	1	0,459	0,5469	0,5424	0,5469	0,5075	0,5938	0,5	0,5606	0,5075	0,4603	0,5667	0,5313
BGS7	0,3279	0,3684	0,4655	0,3846	0,4590	1	0,4603	0,6154	0,5862	0,4030	0,4844	0,3939	0,4769	0,4921	0,3710	0,4032	0,4
BGS8	0,4	0,375	0,5079	0,4925	0,5469	0,4603	1	0,5161	0,5938	0,6508	0,5441	0,4783	0,5362	0,6	0,4394	0,5397	0,5781
BGS10	0,3833	0,3793	0,5	0,4603	0,5424	0,6154	0,5161	1	0,7736	0,5238	0,5397	0,4462	0,5077	0,5484	0,4032	0,5345	0,5
BGS11	0,4	0,4194	0,5079	0,4925	0,5469	0,5862	0,5938	0,7736	1	0,625	0,5441	0,5	0,5588	0,5522	0,3768	0,4923	0,5075
BGS12	0,4308	0,3433	0,5397	0,4571	0,5075	0,4030	0,6508	0,5238	0,625	1	0,5286	0,5073	0,5429	0,5362	0,4697	0,4776	0,5846
BGS13	0,3824	0,3582	0,5077	0,4928	0,5938	0,4844	0,5441	0,5397	0,5441	0,5286	1	0,5441	0,5797	0,5286	0,3803	0,5873	0,5294
BGS14	0,3582	0,3333	0,4844	0,4085	0,5	0,3939	0,4783	0,4462	0,5	0,5073	0,5441	1	0,8276	0,4857	0,5079	0,4923	0,5075
BGS15	0,3971	0,3939	0,5	0,4857	0,5606	0,4769	0,5362	0,5077	0,5588	0,5429	0,5797	0,8276	1	0,5652	0,4776	0,5075	0,5217
BGS16	0,3881	0,3433	0,4697	0,4783	0,5075	0,4921	0,6	0,5484	0,5522	0,5362	0,5286	0,4857	0,5652	1	0,4265	0,5231	0,5373
BGS17	0,4	0,3279	0,4426	0,3286	0,4603	0,3710	0,4394	0,4032	0,3768	0,4697	0,3803	0,5079	0,4776	0,4265	1	0,5517	0,5161
BGS18	0,3438	0,3387	0,5517	0,4844	0,5667	0,4032	0,5397	0,5345	0,4923	0,4776	0,5873	0,4923	0,5075	0,5231	0,5517	1	0,6552
BGS19	0,3636	0,3385	0,6491	0,5231	0,5313	0,4	0,5781	0,5	0,5075	0,5846	0,5294	0,5075	0,5217	0,5373	0,5161	0,6552	1