

**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

Genômica populacional do bacurizeiro (*Platonia insignis* Mart.)

Caroline Bertocco Garcia

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Genética e
Melhoramento de Plantas

**Piracicaba
2021**

Caroline Bertocco Garcia
Licenciada em Ciências da Natureza

Genômica populacional do bacurizeiro (*Platonia insignis* Mart.)
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientadora:
Profa. Dra. **ELIZABETH ANN VEASEY**

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Genética e
Melhoramento de Plantas

Piracicaba
2021

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Garcia, Caroline Bertocco

Genômica populacional do bacurizeiro (*Platonia insignis* Mart.) / Caroline Bertocco Garcia. -- versão revisada de acordo com a resolução CoPGr 6018 de 2011 -- Piracicaba, 2021.

56 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Genética da conservação 2. Bacurizeiro 3. Marcadores SNPs 4. Diversidade genética I. Título

DEDICATÓRIA

Aos apaixonados pela biodiversidade brasileira.

.

AGRADECIMENTOS

Agradeço primeiramente à minha família, em especial à minha mãe, Marisa Bertocco Garcia, que me ajudou no primeiro e mais difícil passo da minha carreira acadêmica, o qual foi a tarefa árdua de ler e escrever. Dona Marisa, professora com magistério aceitou essa tarefa de me ensinar como ninguém, após a tentativa de tantos professores que falharam nessa missão devida a minha dislexia.

À professora, orientadora e minha mãe acadêmica Elizabeth Ann Veasey, que me aceitou e me ajudou tanto academicamente quanto pessoalmente e principalmente, acreditou no meu potencial. A minha permanência no programa num primeiro momento somente foi possível com seu apoio, o qual serei eternamente grata.

Agradeço ao professor e amigo Wellington Ferreira do Nascimento, que foi meu grande parceiro, que me ajudou a amadurecer como cientista e me deu todo apoio durante o mestrado.

À Universidade de São Paulo (USP), que me acolheu durante a graduação, mestrado e agora nos próximos quatro anos no doutorado.

À todos os pesquisadores que auxiliaram durante a coleta do material vegetal, pesquisador Dario Dantas Amaral, Prof. Dr. Wellington Ferreira do Nascimento, Prof. Dr. Edson Ferreira da Silva, Igor Araújo, Prof. Dr. Santiago Linorio Ferreyra Ramos e Allison Vieira da Silva. Agradeço a todos os pesquisadores que colaboraram com o trabalho laboratorial e de análises de dados, Dra. Maria Imaculada Zucchi, Profa. Dra. Doriane Picanço Rodrigues, Prof. Dr. Wellington Ferreira do Nascimento, Allison Vieira da Silva, Igor Araújo, Dra. Flaviane Malaquias Costa e Dr. Carlos Eduardo De Araújo Batista, Dr. Alessandro Alves Pereira e Dra. Caroline Grando e ao técnico Marcos Cella.

A todos os amigos antigos, Rayssa, Wayner, Thiago, Robson, Leandro e aos novos amigos, Igor, Cristiane, Wellington, Leticia, Flaviane, Doriane que contribuíram imensamente para o desenvolvimento da minha dissertação e me deram suporte emocional durante todo o processo.

Agradeço especialmente ao Victor e Allison, que tornaram os momentos que passei nos últimos dois anos memoráveis e que levarei para sempre no meu coração.

Agradeço à ESALQ-USP e ao Departamento de Genética, pela estrutura e riqueza de conhecimento passada através de excelentes professores integrantes do corpo docente. À FAPESP pelo financiamento com o qual foi possível realizar o projeto e à CAPES pela bolsa de mestrado que garantiu a minha permanência na cidade de Piracicaba durante o mestrado.

“Se as portas da percepção estivessem limpas, tudo apareceria para o homem tal como é: infinito.”

(William Blake)

SUMÁRIO

RESUMO.....	7
ABSTRACT.....	8
1. INTRODUÇÃO	9
2. OBJETIVOS	11
2.1. Objetivo Geral.....	11
2.2. Objetivos específicos	11
3. REVISÃO DE LITERATURA.....	13
3.1. Aspectos taxonômicos e distribuição geográfica do bacurizeiro	13
3.2. Importância econômica e conservação	15
3.3. Genômica populacional e conservação de bacurizeiro	17
3.4. Marcadores moleculares utilizados para estudos de estrutura e diversidade genética de <i>P. insignis</i>	19
3.5. Marcadores SNPs e a técnica de “genotipagem por sequenciamento” (GBS)	20
4. MATERIAIS E MÉTODOS	23
4.1. Coleta do material vegetal.....	23
4.2. Extração de DNA genômica	24
4.3. Normalização das amostras.....	25
4.4. Montagem da biblioteca genômica pela técnica de GBS.....	25
4.5. Filtragem dos dados sequenciados e identificação de SNPs.....	26
4.6. Análises estatísticas.....	26
5. RESULTADOS.....	29
5.1. Diversidade e estrutura genética e genômica do bacurizeiro	29
6. DISCUSSÃO	39
6.1. Estrutura e diversidade genética do bacurizeiro em seus principais biomas e regiões de ocorrência no Brasil	39
6.2. A fragilidade ambiental das populações do bacurizeiro e as práticas conservacionistas sustentáveis	43
7. CONCLUSÕES	47
REFERÊNCIAS.....	47

RESUMO

Genômica populacional do bacurizeiro (*Platonia insignis* Mart.)

A espécie *Platonia insignis* Mart., conhecida como bacurizeiro, é uma espécie frutífera nativa do Brasil e corresponde a única representante do gênero monotípico *Platonia* Mart., pertencente à família Clusiaceae. Considerada uma espécie semi-domesticada, ocorre principalmente nas regiões Norte e Nordeste nos biomas de cerrado e amazônico com centro de diversidade no Estado do Pará. O bacurizeiro é uma espécie de importância econômica em ascensão com sua polpa comercializada com o maior valor de mercado dentre as frutíferas brasileiras. O crescente interesse pelos seus frutos, denominados bacuri, exige o conhecimento a respeito da estrutura e diversidade das populações dessa espécie assim como o Estado de conservação, para garantir futuros programas de melhoramento genético e ações voltadas à conservação da espécie. Nesse contexto, o seguinte estudo teve como objetivo avaliar a variabilidade genética e genômica do bacurizeiro ao longo da sua distribuição geográfica, assim como as condições de conservação a qual essa frutífera se encontra. Para tanto, a partir de 2.031 marcadores SNPs obtidos pela técnica de genotipagem por sequenciamento (GBS) com o sequenciamento de nova geração (NGS) foi possível determinar a estrutura e diversidade genética de seis populações dos Estados de Roraima (RR), Amazonas (AM), Pará (PA) (bioma Amazônia) e Maranhão (MA) (bioma Cerrado). As populações avaliadas apresentaram elevada estruturação genética, observada nos agrupamentos formados pela análise de DAPC e dendrogramas, bem como na análise par-a-par baseada nos valores de diferenciação genética (F_{ST}) entre os grupos definidos pela análise de DAPC e entre as populações avaliadas. Foi também detectada correlação positiva e significativa ($r = 0,85$; $p < 0,005$) entre a diferenciação genética e a distância geográfica entre as populações, indicando a ocorrência de isolamento por distância, pelo teste de Mantel. Em contraste com todos os trabalhos anteriores de caracterização molecular já publicados para a espécie, a maior parte da variabilidade genética (59%) se concentra entre as populações de bacurizeiro, principalmente, entre as regiões (Norte e Nordeste) e biomas (Amazônia e Cerrado) analisados. Os maiores índices de diversidade genética foram observados para a população de Chapadinha, MA, que ocorre no Bioma Cerrado ($H_E = 0,1746$; $H_O = 0,2078$), próxima a uma área de proteção ambiental, mostrando excesso de heterozigotos. As demais populações da região Norte apresentaram níveis bem mais baixos de diversidade genética, com valores de H_E entre 0,0424 e 0,0752, e valores positivos de fixação de Wright, indicando níveis expressivos de endogamia. O avanço das áreas urbanas, eventos de queimada para a pastagem, e implementação de cultivares de espécies como a soja (*Glycine max* (L.) Merr.), são os principais fatores para a redução das populações de *P. insignis*, refletindo a intensa ação antrópica que se observa hoje na Amazônia. Dessa forma, as populações do bacurizeiro são encontradas fragmentadas, reduzidas, com baixa variabilidade e baixo fluxo gênico entre as populações, o que caracteriza o processo evolutivo de deriva genética com efeitos de gargalo genético pronunciados. Com exceção da população de Chapadinha, MA, a qual vem mantendo níveis moderados de diversidade genética, os resultados aqui obtidos expõem a fragilidade a qual essa espécie se encontra, extremamente reduzida, com a presença pronunciada de endogamia e perda drástica de variabilidade genética.

Palavras-chave: Genômica da conservação, Bacurizeiro, Marcadores SNPs, Diversidade genética, Estrutura genética populacional

ABSTRACT

Population genomics of bacurizeiro (*Platonia insignis* Mart.)

Platonia insignis Mart., also known as *bacurizeiro*, is a fruit species native to Brazil and corresponds to the only representative of the monotypic genus *Platonia* Mart., belonging to the Clusiaceae family. It is considered a semi-domesticated species and can be found mainly in the Brazilian North and Northeast, in the Cerrado and Amazon biomes, with a center of diversity in Pará state. The *bacurizeiro* is a species of increasing economic importance due to the trade of its pulp, which has the highest market value among Brazilian fruits. The growing interest in its fruits, called *bacuri*, requires knowledge about the structure and diversity of the populations of this species as well as its conservation status, to guarantee future plant breeding programs and to promote actions towards the conservation of this species. In this context, the following study aimed to assess the genetic and genomic variability of *bacurizeiro* throughout its geographic distribution, as well as the conservation conditions in which this fruit tree is found. Therefore, from 2,031 SNP markers obtained by the technique of genotyping by sequencing (GBS) with new generation sequencing (NGS), the genetic structure and diversity of six populations in the states of Roraima (RR), Amazonas (AM), Pará (PA) (Amazon biome) and Maranhão (MA) (Cerrado biome) was determined. The assessed populations showed strong genetic structure, observed by the clusters formed by DAPC and dendrograms, and the pairwise analysis based on fixation index (F_{ST}) values between groups defined by DAPC and among populations. We also detected a positive and significant correlation ($r = 0,85$; $p < 0,005$) between the genetic differentiation and the geographic distances among populations, indicating the occurrence of isolation by distance using the Mantel test. In contrast to all previously published works on molecular characterization for the species, most of the genetic variability (59%) is concentrated among *bacurizeiro* populations, mainly between the regions (North and Northeast) and biomes (Amazon and Cerrado) analyzed. The highest levels of genetic diversity were observed for the population of Chapadinha, MA, located in the Cerrado biome ($H_E = 0.1746$; $H_O = 0.2078$), close to the environmental protection area, showing an excess of heterozygotes. The other populations in the North region showed much lower genetic diversity levels, with H_E values between 0.0424 and 0.0752, and positive Wright's fixation values, indicating expressive levels of inbreeding. The advance of urban areas, burning events for pasture, and the implementation of crop species such as soybeans (*Glycine max* (L.) Merr.), are the main factors for the decline of *P. insignis* populations, which reflects the intense anthropic action currently observed in the Amazon Rainforest. Thus, the *bacurizeiro* populations are found fragmented, reduced, and with low indexes of variability and gene flow between populations, which characterizes the evolutionary process of genetic drift with pronounced genetic bottleneck effects. Except for the population of Chapadinha, MA, which has maintained moderate levels of genetic diversity, the results obtained here expose the fragility in which this species is found, extremely reduced with the steady presence of inbreeding and drastic loss of genetic variability.

Keywords: Conservation genomics, Bacurizeiro, SNP markers, Genetic diversity, Populational genetic structure

1. INTRODUÇÃO

O bacurizeiro (*Platonia insignis* Mart.) é uma árvore frutífera semi-domesticada (Clement, 1999), que pertence à família Clusiaceae, subfamília Clusioideae e representa a única espécie inserida no gênero monotípico *Platonia* Mart. (Cronquist, 1981). A espécie *P. insignis* possui origem Amazônica, com seu centro de diversidade no Estado do Pará, destacando-se as populações situadas nas mesorregiões do Nordeste do Pará e Ilha do Marajó (Silva & Donato, 1993; Clement, 1999; Cavalcante, 1996). Sua dispersão alcançou os Estados do Mato Grosso, Tocantins e Roraima, apresentando-se em vegetação secundária diferentemente dos Estados Acre, Amapá e Amazonas onde ocorre em vegetação primária. No nordeste, o Estado do Maranhão e norte do Piauí apresentam populações do bacurizeiro em Floresta de transição entre Floresta Amazônica e a Caatinga do semiárido nordestino (Carvalho & Nascimento, 2017).

Atualmente, as frutas tropicais tem se tornado frequentemente um produto de preferência na escolha dos consumidores, devido às qualidades nutricionais e aromas atrativos. Esse fato tem despertado grande interesse para a exportação de frutos e produtos derivados, o que torna a fruticultura brasileira uma atividade rentável em vista da ampla diversidade de frutos nativos do país (Souza et al., 2013). Nesse contexto, as espécies frutíferas Amazônicas do Brasil apresentam potencial agroindustrial ainda pouco explorado (Clement & Venturieri, 1990).

Dentre essas espécies destaca-se o bacurizeiro, que produz o fruto denominado “bacuri”, rico em aminoácidos, vitaminas e sais minerais. A sua poupa, como seu principal produto, é consumido em forma de sucos, geleias, sorvetes ou *in natura* e a casca pode ser aproveitada na preparação de doces e cremes (Souza et al., 2000; Shanley et al., 2017). O óleo extraído de suas sementes é comumente utilizado como anti-inflamatório e da casca pode ser extraído o azeite com altas concentrações de ácido palmítico e ácido oleico (Cavalcante, 1996; Villachica et al., 1996). É considerada uma espécie de dupla aptidão, pois além do aproveitamento máximo de cada parte do fruto, a madeira do bacurizeiro por ser resistente e de grande porte (20-30 metros), é intensamente utilizada na construção de embarcações e de casas em muitas áreas de ocorrência natural desta espécie (Souza et al., 2007).

No contexto das espécies frutíferas em ascensão no Brasil, não se pode ignorar a ação antrópica atuando no processo de seleção, deriva genética e na distribuição de diversidade genética. Dentre as frutíferas de interesse econômico, o bacurizeiro como espécie semi-domesticada (Clement, 1999), corre grande risco de erosão genética. Dessa forma, estima-se que boa parte da variabilidade genética existente desta espécie já tenha sido eliminada devido ao desmatamento indiscriminado, rápido avanço das fronteiras agrícolas para implantação de grandes culturas e crescimento das áreas urbanas em suas respectivas áreas de ocorrência (Lima, 2007). Para o uso sustentável dos recursos

dessa espécie, é extremamente importante o conhecimento a respeito de sua variabilidade genética, no sentido de permitir que estratégias de conservação eficientes sejam traçadas e que sejam garantidos subsídios para futuros programas de melhoramento genético (Souza et al., 2016).

Existem diversos tipos de marcadores moleculares que permitem o desenvolvimento de estudos neste sentido (Ouborg et al., 2010). O recente desenvolvimento de tecnologias voltadas a genômica, como o sequenciamento de nova geração (*Next Generation Sequencing-NGS*), permite que os estudos evolutivos assumam uma abordagem em escala genômica (Hohenlohe et al., 2010). Dessa forma, é possível identificar locos que estão sofrendo a ação desses fatores evolutivos, tais como seleção, mutação, recombinação, deriva genética, fluxo gênico e endogamia (Luikart et al., 2003).

O advento do sequenciamento de nova geração (NGS) trouxe conseqüentemente diversas abordagens capazes de identificar milhares de marcadores em quase todo o genoma de interesse, mesmo para as espécies em que pouca ou nenhuma informação genética esteja disponível (Davey et al., 2011). Dentre as técnicas atualmente disponíveis destaca-se a genotipagem por sequenciamento (*Genotyping-by-Sequencing - GBS*), pois além de apresentar um protocolo simplificado em relação às outras técnicas, é adequada para estudos populacionais, tendo em vista que pode ser utilizado para qualquer espécie, com baixo custo por amostra e sem necessidade de genoma de referência (Poland et al., 2012). Dessa forma, a técnica de GBS possibilita a identificação de milhares de polimorfismos de nucleotídeo único (*Single Nucleotide Polymorphism - SNPs*), distribuídos aleatoriamente em todo o genoma, em que é realizada a genotipagem de populações em larga escala de forma rápida e precisa (Davey & Blaxter, 2010; Poland et al., 2012).

Em relação à caracterização molecular, incluindo estudos de genética de populações, pouco tem sido feito com relação ao bacurizeiro, podendo-se citar os estudos de Souza et al. (2013), Pena (2016), Pontes et al. (2017), Santos et al. (2017) e Pena et al. (2020), com marcadores ISSR (*Inter Simple Sequence Repeats*), bem como os de Paraense et al. (2020) que desenvolveram os marcadores microssatélites utilizados por Nascimento et al. (no prelo) baseado em oito marcadores microssatélites nucleares (ncSSR) e três cloroplastidiais (cpSSR). Tais estudos ainda se apresentam defasados em aspectos que podem ser melhor analisados pela genômica populacional, de forma complementar a esses trabalhos. Vale ressaltar que não há estudos abrangendo toda a área de ocorrência da espécie. Dessa forma, o presente estudo ampliou a amostragem realizada por estudos anteriores, e pretendeu realizar um estudo mais abrangente utilizando pela primeira vez a técnica de genotipagem por sequenciamento (GBS) com a tecnologia de sequenciamento nova geração (NGS) e obtenção de marcadores SNPs. Os marcadores SNPs possibilitaram a determinação da estrutura e diversidade genética de seis populações presentes nos dois principais biomas de ocorrência do bacurizeiro, Floresta Amazônica e Cerrado.

2. OBJETIVOS

2.1. Objetivo Geral

Avaliar a diversidade e estrutura genética e genômica do bacurizeiro como subsídio para a conservação da espécie *Platonia insignis* Mart.

2.2. Objetivos específicos

Estimar o nível de diversidade genética nas populações de bacurizeiro utilizando marcadores SNPs;

Avaliar a estrutura genética e genômica entre populações de bacurizeiro utilizando marcadores SNPs;

Descrever as condições de conservação e/ou vulnerabilidade em que se encontram as populações de bacurizeiro, nas distintas áreas de estudo,

Propor estratégias adequadas para a coleta e conservação dessa espécie.

3. REVISÃO DE LITERATURA

3.1. Aspectos taxonômicos e distribuição geográfica do bacurizeiro

O bacurizeiro (*Platonia insignis* Mart.), pertence à família Clusiaceae, subfamília Clusioideae e representa a única espécie inserida no gênero monotípico *Platonia* Mart. (Cronquist, 1981). Possui habito arbóreo e quando encontrado em florestas primárias densas atinge tamanho de 15 a 30 metros (Figura 1B), enquanto as de menor porte (Figura 1A), são encontradas em áreas de cultivo aberto, geralmente de vegetação secundária e de transição (Clement & Venturieri, 1990; Lorenzi, 1992). O tronco é espesso e quando lesionado apresenta exsudato de coloração amarelada (Loureiro et al., 1979). As folhas de *P. insignis* são simples, pecioladas e coriáceas, e variam de acordo com o limbo, ápice, e forma da base; as flores são hermafroditas terminais e de coloração branco-rósea a amarelada (Figura 1C). A antese é diurna e a polinização ornitófila, realizada por psitacídeos (papagaios, curicas e maritacas) (Maués et al., 1996). O tipo de reprodução assexual quando observada para essa espécie demonstra ampla eficiência, pois é realizado em sua maioria de forma bem-sucedida a partir de rebrotamentos feitos da raiz. No entanto, é uma planta predominantemente alógama, com autoincompatibilidade esporófitica (Maués et al., 1996; Cavalcante, 1996).

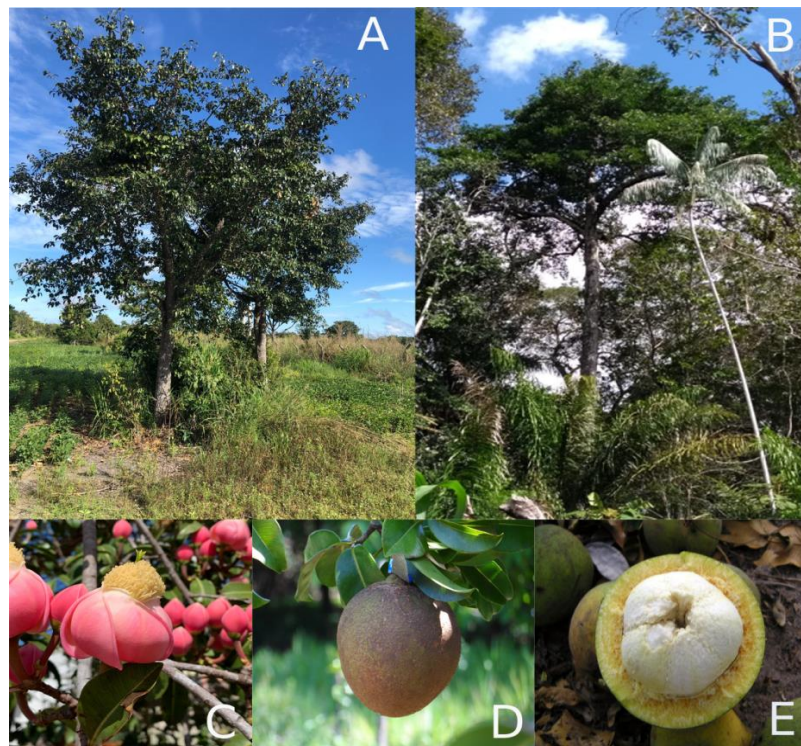


Figura 1. (A) Bacurizeiro da população de Chapadinha no Estado do Maranhão, (B) Bacurizeiro da população de Itacoatiara no Estado do Amazonas, (C) Flores do bacurizeiro. Fonte: acervo próprio de fotos. (D) Fruto no bacurizeiro. (E) fruto maduro de bacuri (*Platonia insignis*). Fonte: Embrapa, 2019.

Os frutos são do tipo baga, com variação no formato entre oblata a oblonga e sua casca (epicarpo e mesocarpo) rígida destaca-se pela coloração externa em verde, verde-amarronzado e amarelo-citrino (Figura 1D). Internamente, a polpa (endocarpo) de sabor e aroma agradável é um produto apreciado na região norte e nordeste do Brasil, a qual é reconhecida pela sua coloração branca a branco-amarelada de aspecto espesso. É carnosa e totalmente aderida à semente separada do pericarpo (Figura 1E) (Mourão et al., 1995; Cavalcante, 1996).

O Estado do Pará é considerado o centro de diversidade do bacurizeiro, que tem como sua origem o bioma Amazônico (Clement, 1999). Sua dispersão natural seguindo em direção ao norte, Roraima, apresenta-se em vegetação secundária, diferentemente dos Estados Acre, Amapá e Amazonas onde ocorre em vegetação primária (Carvalho & Nascimento, 2017). Na direção nordeste, o Estado do Maranhão e norte do Piauí apresentam populações do bacurizeiro em Mata de transição entre Floresta Amazônica e a Caatinga do semiárido nordestino (Carvalho & Nascimento, 2017). (Figura 2).

As árvores de *P. insignis* são densamente distribuídas pelos cerrados e chapadões, que são áreas com intensa atividade agrícola, desmatamento e incêndios florestais (Santos et al., 2017). Devida a sua alta resistência a estresse hídrico e notável capacidade de regeneração, o bacurizeiro obteve sucesso na ocupação dessas áreas (Nascimento et al., 2007; Araújo et al., 2018). Além das fronteiras Brasileiras, o bacurizeiro em menor densidade encontra-se na Guiana, Guiana Francesa, Suriname, Colômbia, Venezuela e raramente na Amazônia Peruana (Patiño, 2002; Nascimento et al., 2007).

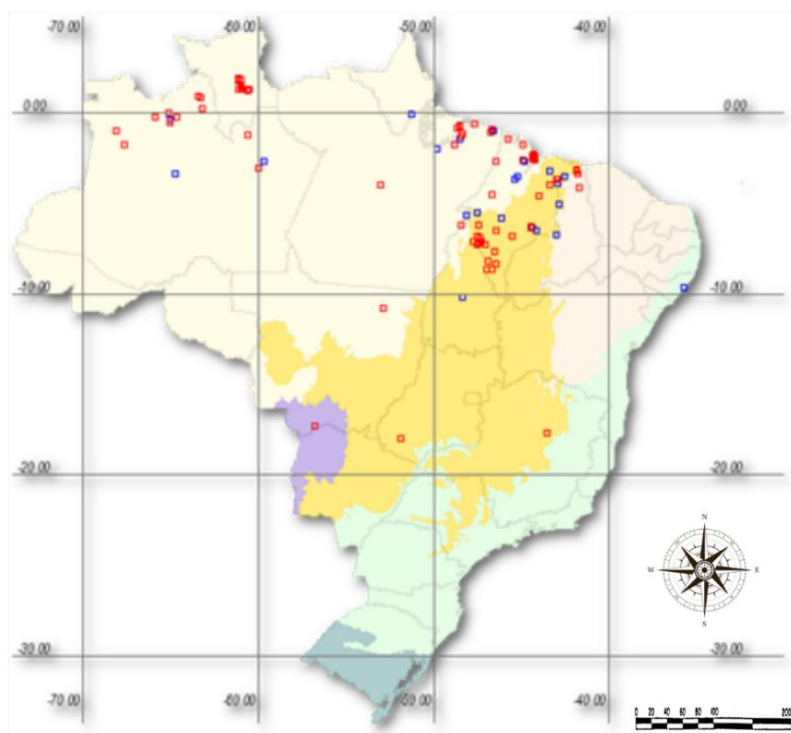


Figura 2. Mapa de distribuição de *Platonía insignis* no Brasil, Fonte: SpeciesLink (<http://smlink.cria.org.br/>)

3.2. Importância econômica e conservação

O fruto do bacurizeiro é rico em aminoácidos, vitaminas (B₁, B₂ e C) e sais minerais (fósforo, cálcio, ferro e potássio) (Morton, 1987), sendo bastante consumido *in natura* ou na forma de geleias, sucos e sorvetes (Souza et al., 2000; Shanley et al., 2017). Embora a polpa seja o principal produto do bacurizeiro, é possível aproveitar também a casca para a preparação de doces e cremes (Souza et al., 2000). Das suas sementes pode ser extraído um óleo bastante utilizado na medicina tradicional como anti-inflamatório (Cavalcante, 1996). Da casca do fruto pode-se extrair o azeite, caracterizado por conter ácido palmítico (44,2 % a 65,4 %) e ácido oleico (26,5 % a 37,8 %). Como a casca contém grande quantidade de pectina (5,0 %), após a separação da resina, essa pode ser usada como fonte de pectina para fabricação de geleias (Vollachica et al., 1996). A madeira do bacurizeiro é resistente e de coloração bege amarelada, sendo muito utilizada na construção de embarcações e de casas em muitas áreas de ocorrência natural desta espécie (Souza et al., 2007).

Dada as qualidades e a frequente procura pelos seus frutos, de acordo com Alves et al. (2020), o bacuri que era uma espécie considerada “alimento do mato”, uma referência à obra Macunaíma, de Mario de Andrade, atualmente assume uma nova posição no mercado de polpas e se torna um alimento das cidades. No cenário atual, a polpa de maior valor no mercado é a do fruto do bacurizeiro, em que o preço estimado do quilo da polpa chegou a ser cinco vezes maior em 2018 comparado a 2005 (Alves et al., 2020). A princípio o manejo era feito a partir da coleta dos frutos de ocorrência natural e atualmente observa-se novas estratégias como a indução do manejo dos rebrotamentos naturais e pomares desenvolvidos por agricultores do Estado do Pará (Alves et al., 2020). Dessa forma, transformando o cenário dessa frutífera para um bem próximo futuramente, tal como as frutíferas consumidas em larga escala pela população, como o cupuaçu (*Theobroma grandiflorum* (Willd. ex. Spreng.) Schum.), açaí (*Euterpe oleracea* Mart.), abacaxi (*Ananas comosus* L.), entre outros (Alves et al., 2020; Homma et al., 2018).

O potencial do bacurizeiro na recuperação de áreas degradadas pode ser vantajoso para a realização do manejo dessa espécie, no sentido de que representa fonte de renda e alimento para os agricultores e atua positivamente na recuperação de ecossistemas destruídos (Araujo et al., 2018; Alves et al., 2020). A alta resistência a solos deficientes, estresse hídrico e sua alta capacidade de regeneração natural, podem ser observadas através dos maciços formados de bacurizeiros no Estado do Maranhão. No Nordeste as populações densas de bacuri são vistas em vegetação secundária, que são caracterizadas justamente por apresentarem sinais de degradação pela ação antrópica (Homma et al., 2018; Alves et al., 2020).

As populações naturais do bacurizeiro em vegetação primária são menores, cerca de no máximo oito indivíduos, pois ocorre um indivíduo por hectare tomando distância uma planta da outra

por volta de 40 metros. Na vegetação secundária as populações do bacurizeiro possuem número maior, chegando a 200 indivíduos por hectare formando os chamados maciços, porém não significa que essas populações apresentam diversidade genética (Homma et al., 2010). Nesse sentido, de acordo com Homma et al. (2010), apesar de apresentar uma forte regeneração natural nas áreas de vegetação secundária, dada as brotações por raízes que facilitam o manejo, a distribuição do bacurizeiro pode ao mesmo tempo influenciar na baixa variabilidade genética da população. Essa situação pode ocorrer devido aos indivíduos que apresentam a mesma origem materna, ou seja, são todos provenientes da mesma planta-mãe, o que é agravado pela autoincompatibilidade da espécie. Essa população então seria formada por clones autoincompatíveis dependentes de pólen vindo de outras populações para que ocorra a produção do fruto.

No contexto das espécies frutíferas em ascensão no Brasil, não se pode ignorar a ação antrópica atuando no processo de seleção, deriva genética e a influência na distribuição de diversidade genética. As populações semi-domesticadas passaram por diversos eventos sequenciais de deriva (efeito fundador) o que reduz sua diversidade genotípica, mas sua diversidade fenotípica é aumentada pelo acúmulo de alelos devido a caracteres selecionados pelos humanos. A diversidade fenotípica do bacurizeiro é bem relatada na literatura atual, o que confirma essa afirmação. Nesse sentido, a domesticação ainda não completa de uma espécie tende a apresentar distribuições mais amplas, o que permite a introgressão com outras populações selvagens ou semi-domesticadas da mesma espécie. Tal introgressão pode oferecer alelos adicionais para os caracteres sujeitos à seleção, aumentando assim sua diversidade genética (Clement et al., 2010).

Contudo, a intensa exploração dos recursos naturais brasileiros deixou marcas profundas na biodiversidade do país. Sob um olhar conservacionista a história revela como se deu esse processo até os dias atuais, pois muitas das espécies nativas foram negligenciadas e sofreram perdas sucessivas durante os anos posteriores a colonização (Daniel, 2004). Nesse contexto, o bacurizeiro atualmente conhecido pela importância comercial dos seus frutos, no passado representava uma espécie de grande importância madeireira. Dessa forma, consistindo numa espécie de dupla aptidão, ou seja, aquela que pode ser explorada ao todo, do fruto desde a casca até a polpa e o tronco, o qual era utilizado na construção de embarcações (Homma et al., 2007). Essa espécie vem sofrendo ao decorrer das décadas grandes perdas já relatadas desde o seu passado madeireiro, fato esse que demonstra o quanto as populações e a diversidade presente nelas pode estar ameaçada (Lustosa, 1976; Souza et al., 2000; Santos et al., 2017). Dessa forma, novas estratégias para a conservação devem ser traçadas, para que seja realizado o manejo sustentável do bacurizeiro. Nesse sentido, os estudos de genômica populacional auxiliam na definição de estratégias eficientes para a conservação dessa espécie (Médail & Baumel, 2018).

3.3. Genômica populacional e conservação de bacurizeiro

A genômica populacional é uma transição da genética de populações, com o intuito de estudar amplamente o genoma de um organismo (Hazzouri et al., 2014). De um modo mais abrangente visa o entendimento de processos que influenciam nas frequências gênicas decorrentes de processos evolutivos (seleção natural, mutação, deriva genética e fluxo gênico) (Hartl & Clark, 2010; Luikart et al., 2003). A definição mais restrita de Black et al. (2001) para genômica populacional, consiste no entendimento de que a amostragem de todo o genoma é um meio para separar os efeitos específicos dos locos (mutação, seleção e recombinação) dos efeitos do genoma (deriva genética, fluxo gênico e endogamia). Nesse sentido, a microevolução pode ser melhor compreendida a partir dessa visão, além da possibilidade de inferir de forma confiável a história filogenética e demográfica das populações. De acordo com Krutovskii (2006), a genômica populacional dentro de um contexto climático em constante mudança, se torna fundamental para entender a base da adaptação genética e a divergência entre as populações. Nesse contexto, a Genômica populacional abre novos horizontes para o entendimento da evolução.

A genética de populações até a década de 1990 apresentava estudos com poucos locos gênicos devida à falta de recursos tecnológicos em genômica, que tornassem possível lidar com um número maior de locos (Bragg et al., 2015). O desenvolvimento crescente de novas tecnologias após esse período tornou possível a análise e descrição de polimorfismos. Assim, a genômica populacional tornou-se um termo cada vez mais frequente (Ellegren, 2014). Nessa perspectiva, o sequenciamento de nova geração (NGS) e a bioinformática revolucionaram a genômica. As plataformas Illumina (HiSeq ou MiSeq), os programas de filtragem (pipelines) e análise de dados tornaram possível estudos genômicos em larga escala. E foi observado o aprimoramento das técnicas para NGS, como genotipagem a partir dos SNPs, que foram desenvolvidas, dentre outras técnicas, para a redução de complexidade, como o GBS (Genotyping-By-Sequencing), RADseq (Restriction site-associated DNA sequencing) e ddRADseq Double digest restriction-site associated DNA sequencing) (Fuentes-Pardo & Ruzzante, 2017). Nesse sentido, as análises utilizando recursos genômicos (sequenciamento de nova geração, marcadores SNPs e técnicas de GBS), tem demonstrado análises mais precisas e de maior acurácia (Allendorf et al., 2010). Os avanços nas tecnologias para o sequenciamento de nova geração (NGS) são imprescindíveis, pois permitem o descobrimento de inúmeros marcadores ao longo do genoma de interesse, ainda que essas populações possuam pouca ou nenhuma informação genética (Davey et al., 2011). Com a expansão dos estudos genômicos para populações naturais, questões voltadas para a conservação podem ser respondidas com mais acurácia (Kohn et al., 2006). As medições mais precisas e extensas que utilizam o estudo em larga escala do genoma possibilitam melhor analisar a variação genética baseada em locos neutros, para que seja possível prever a variação

genética e a base adaptativa. Além disso, é possível o entendimento mais preciso da história demográfica, com base em dados genômicos, de espécies ameaçadas, o que proporciona melhor distinção de processos evolutivos a longo prazo em relação a mudanças recentes causadas pela ação antrópica (Kohn et al., 2006; Hazzouri et al., 2014).

Num contexto de espécies arbóreas frutíferas não domesticadas e semi-domesticadas, suas populações são o reflexo da ação das forças evolutivas agindo em diferentes níveis para a adaptação, especiação, demografia e ação antrópica (Aravanopoulos, 2014). As populações de espécies arbóreas se encontram em ambientes heterógenos, que influenciam na grande diversidade genética dessas espécies, para que ocorra a resposta adaptativa a essas diferenças dentro do habitat em que cada espécie pertence, ou seja, o reservatório gênico permite que as adaptações aconteçam levando à evolução (Neale & Kremer, 2011). A taxa alta de fluxo gênico frequente para espécies de hábito arbóreo, conta com a forte seleção adaptativa local dado a elevada diferenciação populacional de características adaptativas essenciais (Savolainen & Pyhäjärvi, 2007; Hazzouri et al., 2013).

As árvores possuem características que as posicionam como modelos para estudos dentro da genômica populacional, pois são sésseis de vida longa, ocupam ambientes diversificados e distribuição extensa. Além dessas características, as espécies arbóreas historicamente foram submetidas as contrações e expansões durante os ciclos glaciais, apresentando variação para adaptação relacionada a respostas climáticas. Por fim, apresenta de forma expressiva baixo desequilíbrio de ligação (LD) (Krutovskii, 2006). O estudo com espécies arbóreas frutíferas com a abordagem de genômica populacional da conservação, permite ampla precisão de estimativas e parâmetros genéticos demográficos (Graudal et al., 2014). O tamanho efetivo (N_e) é fundamental entre esses parâmetros, pois com base na teoria neutra e da evolução molecular, a diversidade reflete o equilíbrio entre a taxa de mutação e a perda de variação devido ao efeito de deriva genética (Kohn et al., 2006). A quantidade de diversidade numa população é determinada principalmente pelo tamanho efetivo, tornando esse conceito a base para a comparação dos efeitos da variação na história evolutiva dos indivíduos da população (Charlesworth, 2009; Kohn et al., 2006). A caracterização da variação genética adaptativa a partir de dados genômicos apresenta resolução maior, o que permite avaliar a ocorrência da seleção natural, e de forma mais precisa, inferir padrões demográficos, dado o aumento da avaliação de locos neutros (Hohenlohe, 2013).

3.4. Marcadores moleculares utilizados para estudos de estrutura e diversidade genética de *P. insignis*

No contexto do estudo de genética populacional com espécies arbóreas e frutíferas, o bacurizeiro foi estudado utilizando marcadores ISSR (Inter-Simple Sequence Repeat) e SSR (Simple-sequence repeats), que são limitados no sentido de quantidade de marcadores e locos amostrados (Souza et al., 2013; Pena, 2016; Pontes et al., 2017; Paraense et al., 2020; Pena et al., 2020; Nascimento et al., no prelo). Nesse sentido a genômica populacional amplia a escala de amostragem do genoma, demonstrando representatividade em termos de quantidade de marcadores e de regiões genômicas (Ellegren, 2014).

Os diferentes tipos de marcadores moleculares utilizados em genética de populações, para o estudo da estrutura populacional de *P. insignis*, revelaram padrões e características importantes para o entendimento da evolução dessa planta arbórea (Souza et al., 2013; Pena, 2016; Pontes et al., 2017; Paraense et al., 2020; e Pena et al., 2020). Marcadores ISSR foram utilizados para a caracterização molecular do bacurizeiro por Souza et al. (2013), que observaram alto nível de diversidade e diferenciação genética entre 10 populações do Maranhão e Piauí. Outro estudo realizado por Pontes et al. (2017), com base em marcadores ISSR, caracterizou 78 acessos da Ilha de Marajó no Estado do Pará, onde foi identificado baixa diferenciação genética entre e dentro das populações (progênies). A estrutura e conservação da paisagem que influenciam a diversidade genética do bacurizeiro foi estudada por Santos et al. (2017), com marcadores ISSR. Nesse estudo foram utilizados acessos da Reserva Extrativista da Chapada Limpa, Município de Chapadinha no Estado do Maranhão, em que a variação genética se apresentou de forma elevada (85,2%) dentro das populações e baixa (14,7%) entre as populações. Além disso, o trabalho ressaltou que o bacuri apresenta indícios de que está sofrendo diretamente com as consequências dos eventos de incêndios intensivos e redução de vegetação primária.

Recentemente, Pena et al. (2020), utilizando marcadores ISSR, caracterizaram 31 acessos de diferentes regiões do Pará, em que a diversidade genética foi predominantemente observada dentro das populações (98%). O estudo com base em marcadores microssatélites realizado por Paraense et al. (2020), voltado para o desenvolvimento de 50 *primers* para o bacurizeiro, tornou possível a caracterização do banco de germoplasma da Embrapa Amazônia Oriental para essa espécie por Pena (2016). Os dados foram obtidos através da utilização de 13 *primers* de microssatélites, observando-se valores elevados de heterozigosidade esperada (H_E) quando comparados a heterozigosidade observada (H_O). Os valores altos do índice de fixação (f), de 0,275, sugerem a ocorrência de endogamia.

Baseado em oito marcadores microssatélites nucleares (ncSSR) e três cloroplastidiais (cpSSR), Nascimento et al. (no prelo) avaliaram sete populações naturais provenientes dos dois principais

biomas em que a espécie *P. insignis* se encontra, Amazônia e Cerrado. Os valores elevados de heteroziguidade foram encontrados para as populações amazônicas e a do município de Chapadinha no Estado do Maranhão. Contudo, esse estudo se destacou em relação aos outros trabalhos citados com marcadores moleculares, pois foram identificadas correlações significativas entre as distâncias geográficas e genéticas.

Nesse contexto, os estudos ainda se apresentam defasados em aspectos que podem ser melhor analisados pela genômica populacional, de forma a complementar esses trabalhos (Luikart et al., 2003; Ellegren, 2014). Os estudos de genética de populações com *P. insignis* ainda são escassos, e até o momento, nenhum apresentou o uso de nova geração de sequenciamento de DNA para obtenção de SNPs, com o intuito de estimar a estrutura genômica das populações de bacurizeiro.

3.5. Marcadores SNPs e a técnica de “genotipagem por sequenciamento” (GBS)

As técnicas para obtenção de marcadores moleculares são utilizadas amplamente em estudos de diversidade genética. Entretanto, o uso de marcadores convencionais como ISSR e SSR necessitam de uma amostragem maior de indivíduos por população, além de um número razoável de marcadores utilizados. Por outro lado, as técnicas atuais para a obtenção de marcadores SNPs são eficientes e utilizam maior número de marcadores exigindo um menor número de acessos por população (Willing et al., 2012). A princípio, a primeira técnica com o objetivo da obtenção de maior quantidade de marcadores SNPs foi a RAD (*Restriction-site Associated DNA*), por meio de microarranjos (microarrays) (Miller et al., 2007). Entretanto, é necessário para microarranjos o conhecimento prévio do genoma de referência, o que representa um impedimento para estudos que não possuem essa informação (Peterson et al., 2012). Posteriormente, essa técnica foi adaptada às tecnologias de sequenciamento de nova geração, assumindo a denominação RADseq (Baird et al., 2008). Os marcadores de DNA associados a um sítio de restrição (RAD) correspondem a fragmentos de DNA associados a sítios de reconhecimento por enzimas de restrição. Estes marcadores possuem variações em relação a técnica, uma delas é ddRADseq (*Double-digest restriction-site-associated DNA sequencing*) (DaCosta & Sorenson, 2014). Dessa forma, não há necessidade de um genoma de referência, pois o número de leitura de sequência é maximizado (Hohenlohe et al., 2011; Peterson et al., 2012).

Dentre as técnicas de redução de complexidade do DNA para a montagem de bibliotecas genômicas, a técnica de GBS destaca-se, pois representa uma técnica com protocolo simplificado em relação as técnicas RAD (Poland et al., 2012). A abordagem da genotipagem por sequenciamento é adequada para estudos populacionais, tendo em vista que pode ser utilizada para qualquer espécie,

com baixo custo por amostra e sem necessidade de genoma de referência, obtendo alto rendimento por meio de sequenciamento de nova geração (NGS) (Elshire et al., 2011).

Para a obtenção dos marcadores SNPs pela técnica GBS, a amostra de DNA genômico é digerida com duas enzimas de restrição, de corte frequente e raro. Dessa forma, a partir do corte realizado pelas enzimas são gerados fragmentos com extremidades coesivas, conhecidos como “Sticky Ends”, que permitem a ligação da sequência de interesse. Os adaptadores (barcodes) são *primers* ligados à uma sequência conhecida, funcionando como um código de barras para a identificação das sequências durante o processamento dos dados. Os barcodes são inseridos em cada amostra contendo cada indivíduo, e posteriormente são misturadas (pool) para obtenção das bibliotecas genômicas multiplexadas. Posteriormente, os fragmentos são então amplificados por meio de PCR, purificados e sequenciados em plataforma Illumina (Poland et al., 2012). Não há necessidade da fragmentação aleatória (random shearing), o que difere da técnica RAD-seq convencional e reduz o custo de produção das bibliotecas e complexidade do genoma (Puritz et al., 2014; DaCosta & Sorenson, 2014).

A plataforma Illumina permite que seja feito o sequenciamento massivo das amostras e o rápido descobrimento de milhares a milhões de SNPs. Geralmente grande parte dos marcadores SNPs são bialélicos, possuem taxa de mutação 10^{-8} a 10^{-9} (Freeland, 2005) e ocorrem amplamente ao longo de todo o genoma. Isto posto, é considerado que os SNPs demonstram grande potencial para estudos com genômica de populações (Willing et al., 2012).

Nesse contexto, a utilização de marcadores SNPs em estudos com genômica e diversidade genética de bacurizeiro permite que sejam avaliados possíveis desvios no modo como os marcadores neutros se comportam (Davey & Blaxter, 2010). O reconhecimento desses locos permite que as sequências do genoma sejam identificadas, o que possibilita detectar se estas sequências estão associadas a padrões de diversificação dentre as populações, além de fornecer informações em relação a organização da diversidade genômica. Dessa forma, os estudos que assumem uma abordagem genômica com marcadores SNPs, por meio da técnica GBS, podem fornecer informações práticas para orientar futuros esforços para a conservação dos recursos genéticos do bacurizeiro e direcionar estratégias para o cultivo e melhoramento dessa espécie.

4. MATERIAIS E MÉTODOS

4.1. Coleta do material vegetal

Para a definição das áreas de coleta, foram consultados os mapas de distribuição geográfica com auxílio do SpeciesLink (<http://splink.cria.org.br/>) e GBIF: *The Global Biodiversity Information Facility* (<https://www.gbif.org>). Foram coletados 39 indivíduos de bacurizeiro distribuídos entre os Estados do Amazonas (Município de Itacoatiara), Pará (Municípios de Bragança e Ilha do Marajó), Roraima (Municípios de Rorainópolis e Nova Colina), Maranhão (Município de Chapadinha) e Piauí (Município de Teresina) (Tabela 1; Figura 3). As folhas foram coletadas e mantidas em sacos do tipo zip lock contendo sílica até a chegada ao Laboratório de Genética Ecológica de Plantas (LAGEP), no Departamento de Genética da Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP). Após a chegada, as amostras foram armazenadas em freezer -30 °C. Ressalta-se que foi realizado o registro das atividades de acesso do bacurizeiro (cadastro nº A744818), junto ao Sistema Nacional de Gestão do Patrimônio Genético (SISGen) visando regularização das atividades de pesquisa propostas.

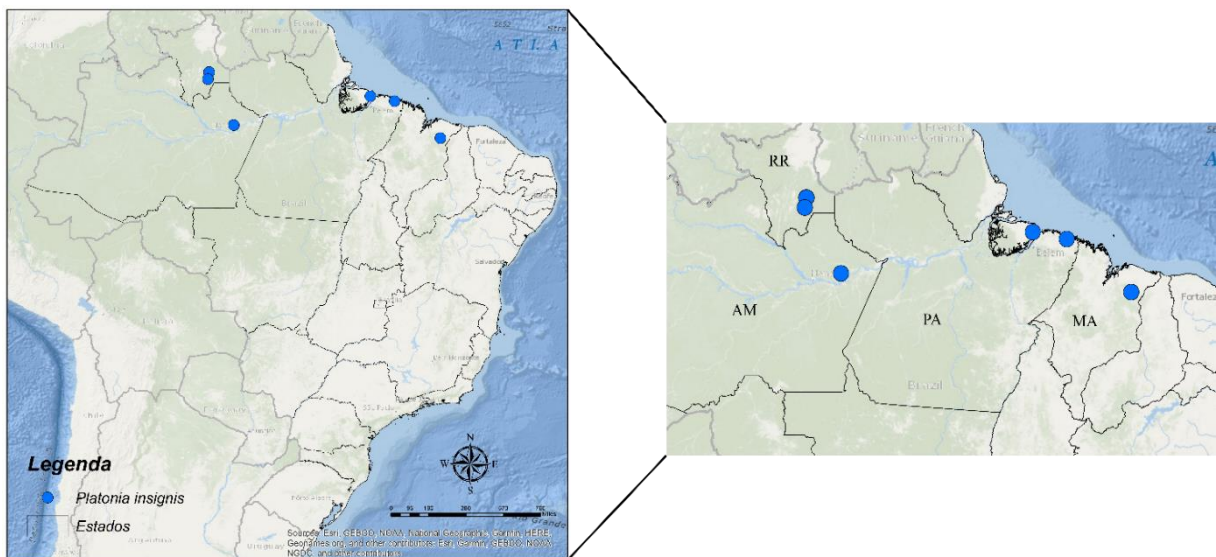


Figura 3. Mapa com os locais de coleta dos acessos de bacurizeiro (*Platonia insignis*), nos Estados brasileiros do Amazonas (AM), Roraima (RR), Pará (PA) e Maranhão (MA).

Tabela 1. Lista das populações de bacurizeiro (*Platonia insignis*) coletados, com detalhes de Estados, municípios, número de indivíduos, bioma e coordenadas geográficas.

Estados	Município	Número de Indivíduos	Bioma	Latitude	Longitude
Amazonas	Itacoatiara	5	Amazônia	02°47'47,0"S	58°34'15,1"W
Pará	Bragança	3	Amazônia	01°03'27,5"S	46°44'9,91"W
	Ilha do Marajó	5	Amazônia	00°41'37,0"S	48°31'07,0"W
Maranhão	Chapadinha	5	Cerrado	03°44'30,0"S	43°21'37,0"W
Roraima	Rorainópolis	12	Amazônia	01°03'14,7"N	60°23'10,4"W
	Nova Colina	9	Amazônia	00°32'29,9"N	60°27'50,5"W
		39			

4.2. Extração de DNA genômica

A extração de DNA genômico dos indivíduos foi realizada no Laboratório de Genética Ecológica de Plantas do Departamento de Genética, na ESALQ/USP, Piracicaba-SP. O DNA foi extraído das folhas por meio do protocolo descrito por Inglis et al. (2018) com algumas modificações. As amostras de folhas foram maceradas em nitrogênio líquido, de forma manual. Após maceradas, as amostras passaram por quatro pré-lavagens com o tampão contendo sorbitol (100 mM Tris-HCl pH8.0, 0.35 M sorbitol, 5 mM EDTA pH 8.0, 1% (massa/volume) Polyvinylpyrrolidona (PVP). Logo após a adição do sorbitol foi realizada a adição de 1% (volume do tampão) de 2-mercaptoetanol. As amostras foram centrifugadas em temperatura ambiente a 7.000 x g durante cinco minutos e tiveram o sobrenadante descartado ao fim de cada lavagem. Após as lavagens com sorbitol, foram adicionados 700 µl de tampão de lise CTAB (100 mM Tris-HCl pH 8.0, 3 M NaCl, 2% CTAB, 20 mM EDTA e 1% (do volume do tampão) de PVP). As amostras tiveram a adição de 1% (do volume do tampão) de 2-mercaptoetanol e seguiram ao banho-maria pré-aquecido a 65 °C, onde permaneceram por 30 minutos sendo levemente homogeneizadas a cada cinco minutos. Depois de serem retiradas do banho-maria, as amostras foram resfriadas por cinco minutos à temperatura ambiente e foram então adicionados 800 µl de cloroformio: álcool isoamílico (24:1 v/v; CIA). As amostras foram homogeneizadas suavemente e centrifugadas a 13.000 x g durante 10 minutos em temperatura ambiente. Ao fim da centrifugação, a fase superior aquosa das amostras foi transferida para um novo tubo contendo 10% do volume, do sobrenadante pipetado, de acetato de sódio 3 M pH 5.2 e 66% do volume de Isopropanol a -20 °C. Após a adição do isopropanol as amostras seguiram para o freezer a -20 °C durante o período de uma hora. O pellet foi obtido por centrifugação a 13.000 x g durante 10 minutos à temperatura ambiente e o sobrenadante foi descartado com cuidado para não perturbar o pellet formado. O pellet obtido foi submetido a uma lavagem com o volume de 1 ml de álcool 70% sob centrifugação a 13.000 x g durante 10 minutos. Ao fim da centrifugação as amostras tiveram o excesso de álcool removido e permaneceram na bancada durante uma hora para secarem. O DNA foi ressuspendido em 25 µl de

tampão TE e tratado com 0.1 mg ml⁻¹ de RNase A e mantido em freezer a -20 °C até o momento de uso.

4.3. Normalização das amostras

A normalização do DNA genômico das amostras do bacurizeiro foi feita para que todas as amostras apresentassem a mesma concentração de DNA em 20 ng. Para tanto, a análise do DNA em relação à qualidade e normalização, primeiramente foi feita através de eletroforese em gel de agarose 1% (w/v) corado com Gel Red. O DNA foi avaliado em relação a quantidade com base no phage λ molecular size standards (Invitrogen) em diferentes concentrações (20, 50 and 100 ng μL^{-1}) e validada com fluorômetro de alta sensibilidade (Invitrogen), o Qubit 4, que apresenta elevado nível de precisão. A partir dos valores em nanogramas indicados pelo Qubit 4 para cada amostra, a concentração foi adequada conforme o parâmetro de nivelamento de 20 ng para todas as amostras, sendo necessário diluir e concentrar a quantidade de DNA conforme a necessidade para atingir o valor de nivelamento.

Dentre as amostras extraídas muitas não apresentaram qualidade e quantidade aceitável para a normalização, dessa forma foram normalizadas com sucesso 39 amostras para o bacurizeiro. O processo de normalização garante que o número de *reads* ou leituras de sequência, não apresente grande variação na quantidade para cada acesso, eliminando o viés que pode ocorrer durante a filtragem de dados, caso um acesso apresente mais leituras que outro. Este é um trabalho minucioso, sendo uma etapa de grande importância, onde deve ser garantido que a concentração de DNA seja equivalente em todos os acessos, evitando-se desta forma o descarte desnecessário de sequências redundantes durante a filtragem dos SNPs e, conseqüentemente, os problemas que reduzem o número de SNPs obtidos.

4.4. Montagem da biblioteca genômica pela técnica de GBS.

As amostras normalizadas foram inicialmente digeridas com duas enzimas de restrição, *Pst*I e *Mes*I, que foram testadas previamente para o bacurizeiro seguindo o protocolo otimizado de Poland et al. (2012). Os fragmentos digeridos foram ligados a sequências adaptadoras específicas (*barcodes*) realizada com a enzima NEB T4 DNA ligase #M0202, com o tampão NEB Buffer4 e adição de ATP. A concentração elevada de enzima T4 foi utilizada para garantir que a ligação de todos os fragmentos ocorresse de maneira adequada. Os produtos provenientes da reação de ligação, com a devida identificação por meio dos *barcodes*, foram multiplexados e submetidos à amplificação por PCR. O enriquecimento dos fragmentos com adaptadores, por meio da reação em cadeia da polimerase (PCR),

utilizando *primers* com sequências complementares aos adaptadores utilizados na preparação da biblioteca, permitiu a obtenção de uma biblioteca contendo 39 amostras. Posteriormente, as amostras foram sequenciadas em uma linha (ou *flowcell*) por meio de sequenciador com tecnologia Illumina.

As bibliotecas de GBS foram realizadas no laboratório de Diversidade Genética e Melhoramento da ESALQ/USP. O sequenciamento dos produtos de PCR foi realizado na plataforma HiSeq2500 Illumina, junto à empresa EcoMol, no Centro de Genômica da ESALQ/USP.

4.5. Filtragem dos dados sequenciados e identificação de SNPs

A filtragem foi realizada a partir de pipeline computacional projetado especificamente para trabalhar com dados de GBS. Dessa forma, o software Stacks v. 1.42 (Catchen et al., 2011) usado para a filtragem e identificação de SNPs é computacionalmente robusto, eficiente e flexível na montagem *de novo* com leituras curtas. O software Stacks executa procedimentos com dados de milhares de indivíduos e incorpora um banco de dados MySQL e front-end da web para visualização, gerenciamento e modificação de dados.

A filtragem inicialmente envolveu a etapa de controle de qualidade das sequências a partir do componente `process_radtags`, onde foram eliminadas as sequências que poderiam comprometer a veracidade da análise. A separação das sequências de cada indivíduo foi feita de acordo com os diferentes barcodes (demultiplex), em seguida as sequências remanescentes foram analisadas pelo componente `ustacks`. A função desse componente é identificar possíveis locos presentes no material genético de cada indivíduo, com os parâmetros `-m 3 -M 2` e `-N 2`. O próximo componente utilizado pelo programa foi o `cstacks`, com o parâmetro `-n 2`, responsável por criar um catálogo com todos os possíveis locos identificados e verificar quais destes locos possui cada indivíduo analisado. Após esta etapa, dois componentes foram utilizados para o cruzamento de informações entre os locos obtidos para cada indivíduo e os locos do catálogo (`sstacks`) e remoção dos locos com menor probabilidade (`rxstacks`, `--lnl_lim -10`). O componente `populations` foi utilizado para a filtragem final dos marcadores SNPs. Foram retidos apenas um SNP por tag, com profundidade $\geq 3X$; $MAF \geq 0,01$; SNP presente em pelo menos 60% das amostras dentro de populações; SNP presente em pelo menos 3 das 4.

4.6. Análises estatísticas

A Análise Discriminante de Componentes Principais (DAPC) realizada a partir do pacote `adegenet` (Jombart, 2008) no programa R (R Development Core Team, 2015). O DAPC utiliza a Análise de Componentes Principais (PCA) como etapa anterior a Análise Discriminante (DA), para que seja

realizada a transformação dos dados, o que garante que não ocorra a perda desnecessária de informação. Dessa forma, a DAPC apresenta a estruturação genética populacional, em que a variação é minimizada dentro dos grupos e maximizada entre os grupos (Jombart et al., 2010). O número de agrupamentos a partir da DAPC foi calculado pelo método K-means, que executa diferentes probabilidades de números de agrupamentos (Clusters). Dentre os modelos de probabilidade gerados pelo método K-means, o método Bayesian Information Criterion (BIC) considerou o melhor suporte estatístico em relação ao número e natureza dos grupos. Os modelos considerados nessas análises consideraram inicialmente a não adequação aos pressupostos de Hardy-Weinberg e a ausência de ligação entre os locos (Pritchard et al., 2000).

A relação genética entre os 39 acessos do bacurizeiro foi feita pela análise de agrupamento apresentada no dendrograma obtido pelo método Neighbor-Joining, que utiliza as distâncias genéticas de Nei (Nei, 1972) e utilizado o pacote ape (Paradis & Schliep, 2018) no programa R (R Development Core Team, 2015). A partir do programa FigTree v.1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>) foi feita a edição do dendrograma gerando duas versões, uma representando as distâncias refletidas nos ramos e outra somente com os grupos formados com a identificação de cada acesso para melhor visualização. As matrizes par-a-par de F_{ST} baseadas nas populações avaliadas e a partir dos grupos delimitados pela análise de DAPC, foram calculadas a partir do pacote poppr (Kamvar et al., 2014) e visualizadas por meio do pacote corrplot v. 0.84 (Wei & Simko, 2017) no programa R (R Development Core Team, 2015). Foi também estimado o fluxo gênico (Nm) entre populações, utilizando os valores obtidos de F_{ST} , baseado na fórmula $Nm = (1 - F_{ST})/4F_{ST}$ (Slatkin & Barton, 1989).

Os parâmetros de diversidade genética, como o número total de alelos (A), as heterozigosidades observada (H_O) e esperada (H_E) e o coeficiente de endogamia (f), estimados a partir do pacote hierfstat (Goudet & Jombart, 2020) e poppr (Kamvar et al., 2014) no programa R (R Development Core Team, 2015), foram calculados de acordo com os grupos delimitados pelo DAPC e para as populações avaliadas. A avaliação da distribuição da variabilidade genética entre e dentro de populações, bem como entre e dentro dos grupos definidos pela DAPC, foi estimada através da Análise de Variância Molecular (AMOVA) por meio dos pacotes hierfstat (Goudet & Jombart, 2020) e poppr (Kamvar et al., 2014; Kamvar et al., 2015) no programa R (R Development Core Team, 2015). Com o objetivo de verificar se estava ocorrendo o isolamento por distância, foi realizado o teste de Mantel com o auxílio do pacote ade4 (Chessel et al., 2004; Dray & Dufour, 2007; Dray et al., 2007; Bougeard e Dray, 2018; Thioulouse et al., 2018), visando avaliar a correlação entre a divergência genética a partir dos valores de F_{ST} da matriz par a par entre populações e a distância geográfica, gerada a partir das coordenadas geográficas, e construída com o auxílio do pacote geodist (Padgham & Sumner, 2020).

5. RESULTADOS

5.1. Diversidade e estrutura genética e genômica do bacurizeiro

A análise dos dados moleculares resultou em 2.031 marcadores SNPs, os quais foram submetidos a análises estatísticas robustas, destacando-se a Análise Discriminante de Componentes Principais (DAPC). Nesse sentido, um número baixo de componentes já retém informação suficiente para explicar a variação dentre um conjunto de dados de milhares de marcadores (Jombart et al., 2010). Para a análise DAPC foram retidos 38 componentes principais, dentro dos quais 8 explicaram 73,2% da variação total dos dados. O método K-means permitiu identificar quatro grupos genéticos (Figuras 4a e 4b), através da indicação do K ótimo. Os grupos formados apresentaram clara diferenciação e elevada estruturação. Os grupos 1 (cor laranja), 2 (cor vermelha), 3 (cor verde) e 4 (cor azul) agruparam 5, 21, 4 e 9 acessos do bacurizeiro, respectivamente.

O grupo 1 apresentou-se geneticamente distante e isolado dos demais grupos determinados pelo DAPC, os quais se encontram estreitamente relacionados. A outra forma que a análise de DAPC pode ser visualizada é através do gráfico em duas dimensões, com a densidade dos acessos e permite que seja observado a sobreposição entre os grupos formados pela análise de DAPC (Figura 4 b). Dentre os quatro grupos, os que apresentaram sobreposição foram os grupos 2, 3 e 4, em que o grupo 4 apresenta-se como um componente de intersecção entre o grupo 2 e 3. O grupo 4 em relação ao pico de densidade, apresenta o maior pico de interação dentre os três grupos (Figura 4b) e é o que está localizado mais próximo ao centro no gráfico de dispersão dos agrupamentos (Figura 4a). O grupo com o segundo maior pico é o 2 seguido do grupo 1, e o menor pico de densidade é representado pelo grupo 3, pois os acessos estão mais distribuídos no espaço. O maior pico observado do grupo 4 representa um indivíduo concentrado num mesmo espaço em sobreposição ao grupo 3 e outros dois picos de densidade menores que não se sobrepõem aos grupos que fazem intersecção, entretanto uma parte com baixa densidade faz a intersecção com os grupos 2 e 3. O grupo 1 se encontra densamente distribuído num espaço restrito com maior distância dos grupos 2, 3 e 4, sendo que dentre os três grupos está mais próximo do grupo 2 (Figura 4b).

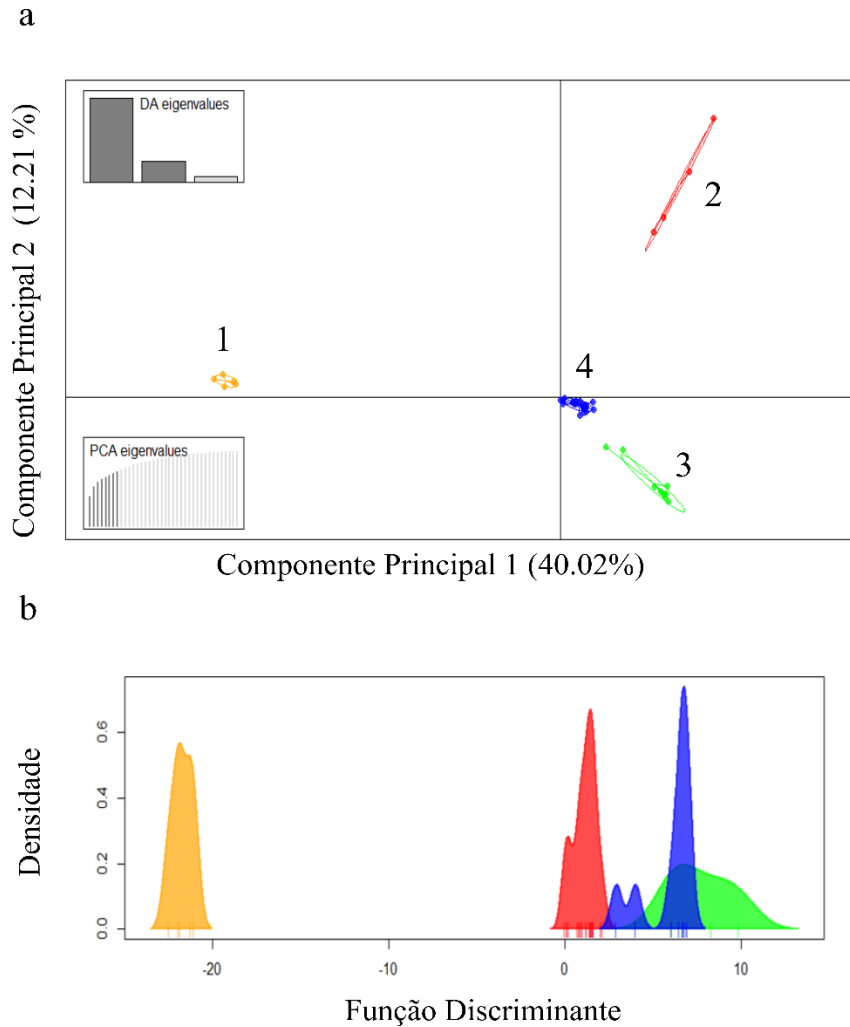


Figura 4: Análise Discriminante de Componentes Principais (DAPC) realizada com base em 2.031 marcadores SNPs para 39 acessos de Bacurizeiro (*Platonia insignis*), representados por: grupo 1 (laranja), grupo 2 (vermelho), grupo 3 (verde), grupo 4 (azul). a) Gráfico de dispersão dos agrupamentos formados pelo método de K-means. b) Gráfico de densidade dos acessos de acordo com os agrupamentos estabelecidos ao longo da primeira função discriminante obtida na DAPC.

O nível de mistura entre os grupos definidos pela análise de DAPC foi baixo, o que indica forte estruturação e diferenciação entre os grupos de acordo com o Estado a que cada grupo pertence (Figura 5). O grupo 1, formado exclusivamente por indivíduos da população de Chapadinha do Estado do Maranhão, região Nordeste, não apresentou mistura com os demais grupos pertencentes aos Estados da região Norte. O grupo 2 foi formado pelas populações de Bragança e Ilha do Marajó, do Estado do Pará, e Nova Colina e Rorainópolis (3 acessos), do Estado de Roraima. Já uma parcela maior dos acessos de Rorainópolis (8 indivíduos), formaram o grupo 4, incluindo um acesso da população de Itaquiara. O grupo 3 foi formado majoritariamente por indivíduos de Itaquiara, do Estado do Amazonas.

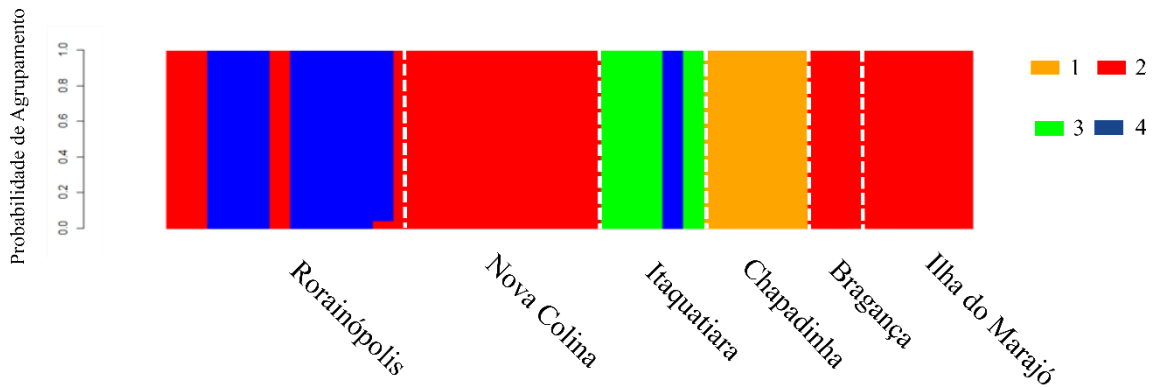


Figura 5: Estrutura genética das populações de bacurizeiro (*Platonia insignis*) com base em 2.031 SNPs demonstrando estrutura genética com a probabilidade da relação entre membros por grupo definido na análise de DAPC, onde as linhas pontilhadas indicam as populações estudadas, sendo Rorainópolis e Nova Colina no Estado de Roraima, Itaquatiara no Estado do Amazonas, Chapadinha no Estado do Maranhão, Bragança e Ilha do Marajó no Estado do Pará.

O dendrograma sem enraizamento representa a análise de agrupamento dos 39 acessos de bacurizeiro, sem a pretensão de definir as relações evolutivas entre os indivíduos. Os grupos representados demonstram especificamente como cada acesso se posiciona (Figura 6) e suas distâncias genéticas (Figura 7). Os grupos genéticos reconhecidos na análise de agrupamento totalizaram cinco: A, B, C, D e E. Os grupos definidos pelo DAPC mostraram coerência com o dendrograma, exceto o grupo 2 que foi dividido em dois grupos: B - com todos os acessos da Ilha do Marajó e Bragança, pertencentes ao Estado do Pará, e C - com os nove indivíduos de Nova Colina e dois de Rorainópolis, pertencentes ao Estado de Roraima. O grupo A com os acessos da população de Chapadinha no Maranhão, tanto na análise de DAPC quanto no dendrograma, formam um grupo distante geneticamente dos acessos da região Norte do Brasil, mostrando-se mais próximos somente das populações do Estado do Pará (Grupo B) (Figuras 6 e 7). O agrupamento D, referente à população de Itaquatiara, Estado do Amazonas, está numa posição intermediária na representação do dendrograma, pois se encontra entre as populações do Estado de Roraima e Pará (Figuras 6 e 7). Rorainópolis (grupo E), com oito dos nove acessos presentes no grupo 4 do DAPC (Figura 5), conforme o dendrograma (Figura 6), está mais próximo geneticamente da população de Nova Colina no mesmo Estado de Roraima. De acordo com as análises de agrupamento pelo método *Neighbor-Joining* visualizadas nos dendrogramas, fica nítida a correlação entre distâncias genéticas e geográficas entre as populações amostradas.

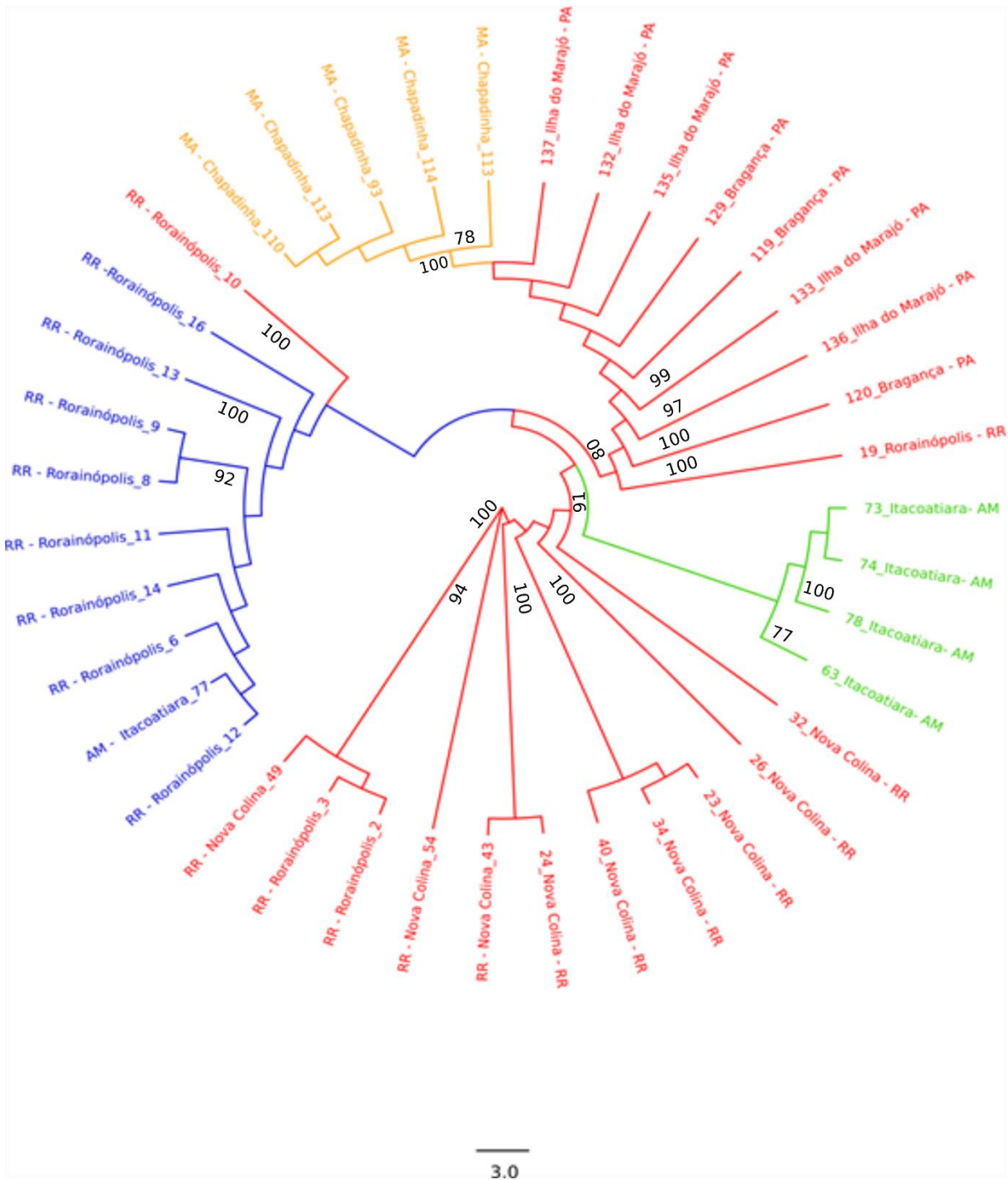


Figura 6: Análise de agrupamento realizada pelo método *Neighbor-Joining*, com base em 2.031 marcadores SNPs, para os 39 acessos de bacurizeiro (*Platonis insignis*) de acordo com os grupos identificados pela Análise Discriminante de Componentes Principais (DAPC).

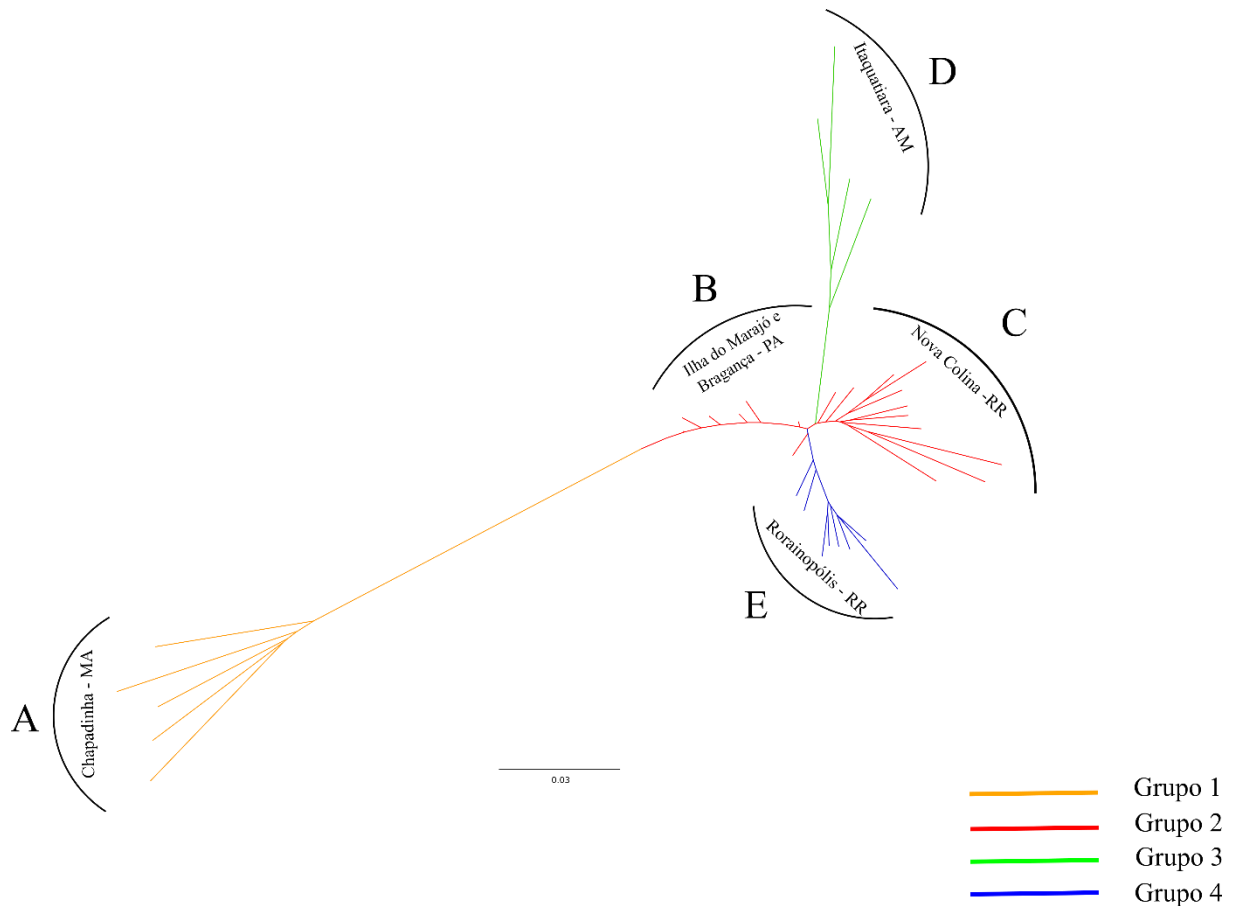


Figura 7: Análise de agrupamento para 39 acessos do bacurizeiro (*Platonina insignis*) a partir da árvore filogenética construída pelo método de *Neighbor-Joining*, com base em 2.031 marcadores SNPs, caracterizada pela estruturação dos grupos identificados pela Análise Discriminante de Componentes Principais (DAPC).

A análise de variância molecular (AMOVA) indica que a variação genética explicada pela variação entre populações é superior à variação dentro das populações, tanto entre os grupos definidos pelo DAPC, quanto entre as populações avaliadas (Tabela 2). Os valores de F_{ST} superiores a 0,25 são indicativos de elevada diferenciação genética. Assim, os valores de F_{ST} para as populações (0,68) e para os grupos definidos pela análise de DAPC (0,59) apresentaram-se próximos, significativos e são mais que o dobro do valor de 0,25. Esse fato demonstra que os grupos do DAPC possuem um grau de compatibilidade relevante com as populações coletadas, de acordo com a região geográfica em que cada uma se encontra e possuem um valor extremamente elevado de diferenciação genética. O teste de Mantel identificou uma correlação positiva e significativa ($r = 0,85$; $p < 0,005$) entre a distância geográfica e a divergência genética, mostrando, dessa forma, que quanto maior a distância entre as populações maior a diferenciação genética entre elas, ou seja, comprovando a existência de um isolamento por distância (Tabela 3).

Tabela 2.: Análise de variância molecular (AMOVA) para o bacurizeiro (*Platonia insignis*) baseada em 2.031 marcadores SNPs realizada para as populações avaliadas e para os grupos da Análise Discriminante de Componentes Principais (DAPC). Valores em negrito indicam as estatísticas correspondentes a diferenciação entre grupos (F_{ST}) em cada AMOVA.

Fonte de Variação	Grau de liberdade	Soma dos quadrados	Coefficiente de Variação	Porcentagem da Variação (%)	Estatística F
Entre Populações	3	15987.13	478.761	68.34	$F_{ST} = 0.68^*$
Dentro de Populações	33	7317.50	221.742	31.65	
Entre Grupos DAPC	3	3373.08	126.89	59.79	$F_{ST} = 0.59^*$
Dentro dos Grupos DAPC	35	2986.95	85.341	40.21	

*Significativo a $p < 0,05$ ($p\text{-value} = 5,00E^{-05}$)

Tabela 3: Matriz par-a-par baseada nos valores de diferenciação genética (F_{ST}) entre os 39 acessos de *P. insignis* analisados agrupados por população.

Populações	Rorainópolis	Nova Colina	Itaquatiara	Chapadinha	Bragança	Ilha do Marajó
Rorainópolis						
Nova Colina	0,28					
Itaquatiara	0,33	0,29				
Chapadinha	0,62	0,60	0,55			
Bragança	0,71	0,69	0,48	0,04		
Ilha do Marajó	0,69	0,68	0,55	0,11	0,16	

Tabela 4: Matriz par-a-par de distância geográfica entre as populações coletadas, na diagonal superior, e na diagonal inferior os valores de F_{ST} calculados entre as populações. Os valores apresentados de distância estão na unidade de quilômetros.

Populações	Rorainópolis	Nova Colina	Itaquatiara	Chapadinha	Bragança	Ilha do Marajó
Rorainópolis		57,327	471,269	1967,065	1537,262	1335,089
Nova Colina	0,28		424,992	1960,842	1538,301	1336,714
Itaquatiara	0,33	0,29		1693,688	1330,608	1142,344
Chapadinha	0,62	0,60	0,55		478,586	665,402
Bragança	0,71	0,69	0,48	0,04		202,441
Ilha do Marajó	0,69	0,68	0,55	0,11	0,16	

A análise par-a-par baseada nos valores de diferenciação genética (F_{ST}) entre os grupos definidos pela análise de DAPC (Figura 8), apresentaram alta diferenciação entre os grupos 1 e 4 (0,69) e entre os grupos 3 e 4 (0,66). De acordo com os valores de F_{ST} , a menor diferenciação observada se encontra entre os grupos 2 e 4 (0,39) e entre os grupos 2 e 3 (0,43). A diferenciação genética atribuída aos grupos do DAPC (Figura 8), conforme os valores encontrados de F_{ST} , são coerentes com as

distâncias genéticas de Nei avaliadas no dendrograma (Figura 7). A matriz par-a-par a partir do F_{ST} entre as populações (Tabela 3; Figura 9) indica que os valores mais baixos de diferenciação genética foram encontrados entre as populações de Chapadinha e Bragança (0,04), Chapadinha e Ilha do Marajó (0,11) e Ilha do Marajó e Bragança (0,16). Os valores mais altos de F_{ST} foram observados entre as populações de Bragança e Rorainópolis (0,71), da Ilha do Marajó e Rorainópolis (0,69), de Bragança e Nova Colina (0,69), e da Ilha do Marajó e Nova Colina (0,68) (Figura 9).

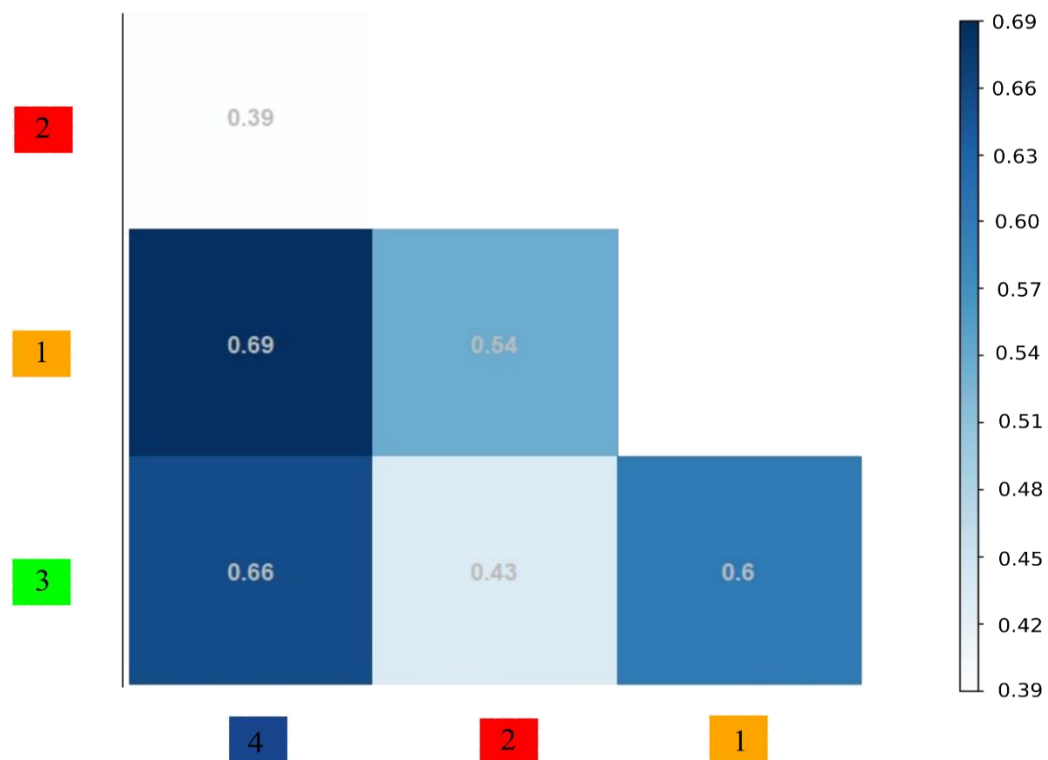


Figura 8: Matriz par-a-par baseado nos valores de diferenciação genética (F_{ST}) para as populações de bacurizeiro (*Platonia insignis*) definidas na análise discriminante de componentes principais (DAPC).

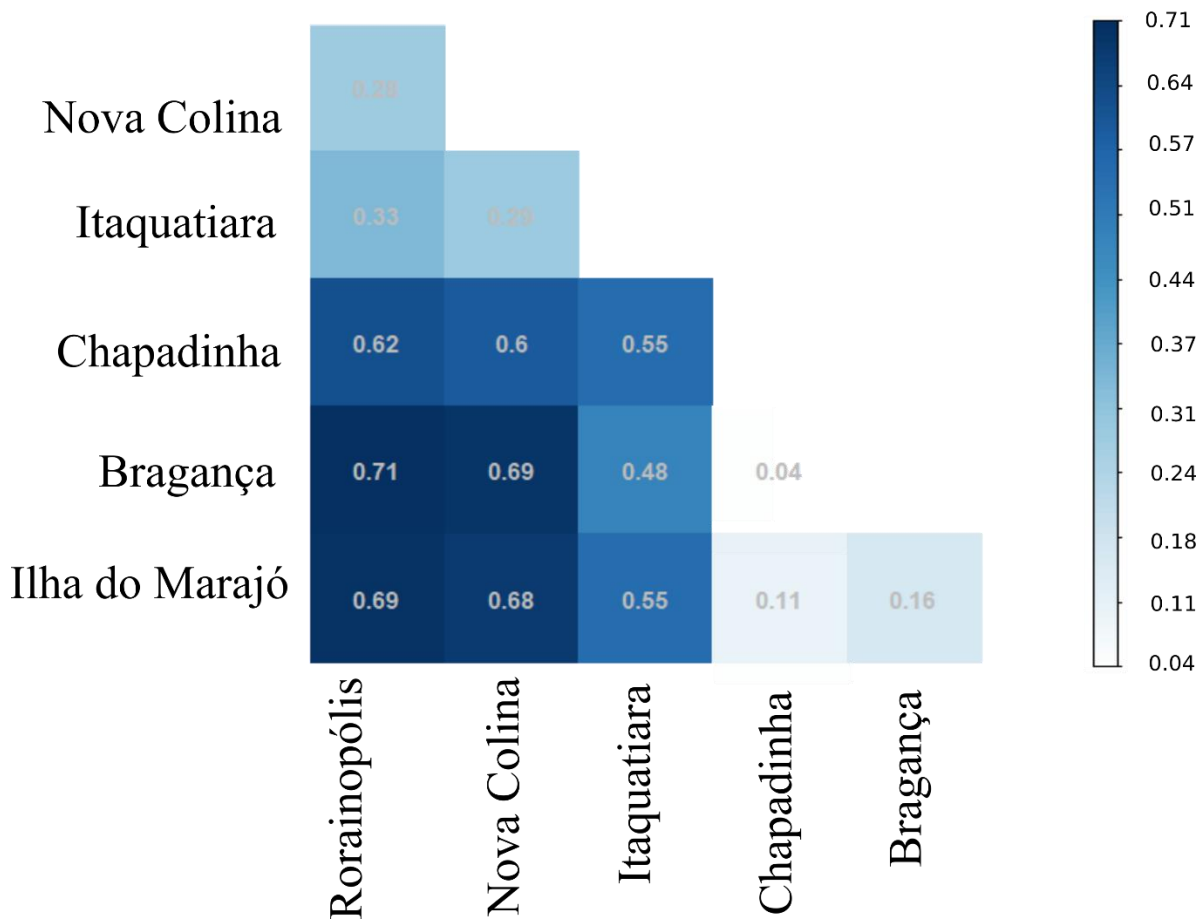


Figura 9: Matriz par-a-par baseada nos valores de diferenciação genética (F_{ST}) entre os 39 acessos de *P. insignis* analisados agrupados por população.

Os valores estimados de fluxo gênico, a partir dos valores de diferenciação genética (F_{ST}) entre as populações mostram que há um baixo fluxo gênico entre as três populações dos Estados de Roraima e Amazonas, variando de 0,51 a 0,64. Essas populações mostraram menor fluxo gênico com as duas populações do Pará ($N_m = 0,10$ a $0,20$), na mesma magnitude que o fluxo gênico observado entre as mesmas com a população de Chapadinha, no Estado do Maranhão, na região Nordeste, variando de 0,15 a 0,20. Já Chapadinha mostrou elevado fluxo gênico com as populações do Pará, sendo $N_m = 6,00$ com a população de Bragança e $N_m = 2,02$ com a população da Ilha do Marajó. E um fluxo gênico mais baixo mas também considerado elevado foi observado entre as duas populações do Pará ($N_m = 1,31$), entre Bragança e Ilha do Marajó.

Tabela 4: Matriz com os valores de fluxo gênico (Nm), calculados a partir dos valores de F_{ST} calculados entre as populações de *P. insgnis*.

Populações	Rorainópolis	Nova Colina	Itaquatiara	Chapadinha	Bragança
Rorainópolis					
Nova Colina	0,64				
Itaquatiara	0,51	0,61			
Chapadinha	0,15	0,17	0,20		
Bragança	0,10	0,11	0,27	6,00	
Ilha do Marajó	0,11	0,12	0,20	2,02	1,31

A heterozigosidade observada para as populações avaliadas e para os grupos definidos pelo DAPC apresentaram, em sua maioria, valores inferiores quando comparados aos da heterozigosidade esperada (Tabela 5). O coeficiente de endogamia negativo nas populações Bragança e Chapadinha e nos grupos 1 e 4, definidos pela análise de DAPC, indicam excesso de heterozigosidade, porém os acessos de Bragança apresentaram valores baixos de heterozigosidade esperada. De acordo com a heterozigosidade esperada, dentre os grupos da DAPC o grupo 1 apresentou maior valor de diversidade genética. O grupo 1 é formado pela população regional de Chapadinha que, consequentemente, apresentou o valor de heterozigosidade esperada mais alto entre as populações avaliadas, seguido pela população de Itaquatiara do Estado do Amazonas.

O número de alelos entre os agrupamentos definidos pela DAPC apresentou valores próximos, em que o menor valor foi encontrado para o grupo 4 e os maiores valores para os grupos 1 e 2. As populações, em sua maioria, apresentaram valores próximos de número de alelos, exceto pelas populações do Estado do Pará com valores baixos em relação as outras regiões. O número de acessos amostrado de cada população não demonstrou uma possível interferência nos parâmetros de diversidade, pois a população de Chapadinha com maior número de alelos e de heterozigosidade observada e esperada possui apenas cinco acessos. Esse fato contrasta com as populações do Estado de Roraima, que mesmo com número superior de acessos apresentou valores menores de números de alelos e heterozigosidade.

O índice de fixação para os grupos 2 e 3, de 0,4295 e 0,3242, respectivamente, são mais próximos de 1,0 o que indica a presença de endogamia nesses dois grupos (Tabela 5). No caso das populações avaliadas, Itaquatiara apresentou índice de fixação de 0,5885, indicando maior grau de endogamia quando comparada às demais populações, seguida das populações de Rorainópolis (0,2589) e Nova Colina (0,2512), e da Ilha do Marajó, com o menor grau de endogamia.

Tabela 5. Parâmetros de diversidade genética para 2.031 marcadores SNPs avaliados em 39 acessos de *P. insignis* das regiões Norte e Nordeste do Brasil, considerando os grupos identificados pela Análise Discriminante de Componentes Principais e as populações avaliadas [Rorainópolis e Nova Colina, do Estado de Roraima (RR), Itaquatiara do Estado do Amazonas (AM), Bragança e Ilha do Marajó do Estado do Pará (PA) e Chapadinha do Estado do Maranhão (MA)].

Grupo	N*	A	H_o	H_E	f
1	5	2981	0,2078	0,1746	-0,1904
2	21	2764	0,0561	0,0983	0,4295
3	4	2339	0,0562	0,0832	0,3242
4	9	2184	0,0461	0,0435	-0,0585
Média		2567	0,0915	0,0999	0,1262
Rorainópolis - RR	12	2310	0,0493	0,0665	0,2589
Nova Colina - RR	9	2326	0,0563	0,0752	0,2512
Itaquatiara - AM	5	2566	0,0519	0,1262	0,5885
Bragança - PA	3	609	0,0572	0,0424	-0,3500
Ilha do Marajó - PA	5	676	0,0644	0,0649	0,0071
Chapadinha - MA	5	2981	0,2078	0,1746	-0,1904
Média		1911	0,0811	0,0916	0,0942

*Número de acessos (N), número total de alelos (A), heterozigidade observada (H_o), heterozigidade esperada (H_E) e coeficiente de endogamia (f).

6. DISCUSSÃO

6.1. Estrutura e diversidade genética do bacurizeiro em seus principais biomas e regiões de ocorrência no Brasil

A estruturação genética das seis populações amostradas, conforme a análise de DAPC e as distâncias genéticas observadas no dendrograma, refletem a relação das populações com as regiões geográficas. Os grupos formados em ambas as análises são compatíveis com a localidade de cada população e a pouca mistura observada entre elas indica o isolamento dessas populações e, conseqüentemente, a diferenciação entre elas. Parte elevada da variabilidade genética (68%) se concentra entre as populações de bacurizeiro, principalmente, entre as regiões (Norte e Nordeste) analisadas. Logo, a variabilidade genética das populações de bacurizeiro está altamente estruturada ($F_{ST} = 0,59$; $p < 0,001$). A estruturação de acordo com a região, sugerindo o isolamento por distância é evidenciado pelo teste de Mantel com 0,85 de correlação positiva e significativa entre distâncias genéticas e geográficas. O fato da maior parte da variabilidade genética se concentrar entre as populações, e não dentro de populações, é um dado que contrapõe todos os trabalhos de caracterização molecular publicados com a espécie *P. insignis* até o momento. É possível que este resultado se deva ao fato de que estamos avaliando populações muito distantes, de diferentes regiões e biomas, embora um estudo semelhante (Nascimento et al., no prelo) com marcadores SSR, também envolvendo populações das regiões Norte e Nordeste, tenha detectado que a maior parte da variabilidade se encontra dentro de populações (72%) e dentro de grupos DAPC (78%).

As distâncias genéticas observadas no dendrograma, no agrupamento formado pela análise de DAPC e a diferenciação apresentada na matriz par-a-par de F_{ST} indicam o quanto população de Chapadinha na região Nordeste está afastada geneticamente de forma significativa dos acessos pertencentes à região Norte do Brasil. O agrupamento 1 delimitado pelo DAPC é exatamente o grupo formado pela população de Chapadinha, o que indica a influência das características geográficas e fitofisionômicas para a estrutura da população. Os acessos desta população apresentam-se isolados dos acessos das populações do Norte, exceto pelas duas populações do Estado do Pará geneticamente mais próximas, com valores baixos de diferenciação obtidos na matriz par-a-par de F_{ST} e maior fluxo ($Nm = 6,00$) gênico entre elas. Vale ressaltar que mesmo com a distância genética menor comparada as outras populações amostradas, os acessos da Ilha do Marajó e Bragança, do Pará, ainda apresentam distância genética considerável da população de Chapadinha, conforme verificado no dendrograma (Figura 7). A correlação entre distâncias genéticas e geográficas confirmada pelo teste de Mantel e a forte estruturação entre a região Norte e Nordeste observada nesse estudo e no estudo de Nascimento et al. (no prelo), deixa claro o isolamento pela distância entre as populações

Ao Norte do Brasil há predominância do bioma da Floresta Amazônica, o que torna os Estados como o Pará, Amazonas e Roraima ricos em espécies frutíferas de interesse econômico. O Pará, considerado como centro de diversidade do bacurizeiro (Silva & Donato, 1993; Cavalcante, 1996; Clement, 1999), tem acessos de diferentes regiões com sua diversidade conservada no Banco Ativo de Germoplasma (BAG) da Amazônia Oriental. Pontes et al. (2017) e Pena et al. (2020) avaliaram populações da ilha do Marajó pertencentes ao BAG da Amazônia Oriental com o objetivo de realizar a caracterização molecular com marcadores ISSR. Os dados obtidos mostraram nos dois estudos que as populações possuem maior variação genética dentro do que entre as populações e baixa diferenciação genética entre as localidades tanto da Ilha do Marajó quanto de outras localidades do Estado do Pará. A matriz par-a-par de F_{ST} no presente estudo revelou elevada diferenciação genética entre as duas populações de Bragança e Ilha do Marajó ($F_{ST}= 0,16$), uma vez que F_{ST} acima de 0,15 já é considerado um valor alto, segundo Hartl & Clark (2010) e contrastando com os trabalhos anteriores, o que pode ser explicado pela barreira física estabelecida pelo oceano entre a Ilha e o continente. Entretanto, o agrupamento dos acessos dessas populações observado na análise de DAPC demonstrou haver semelhanças genéticas entre essas duas populações e o dendrograma reforçou esse resultado apresentando menor distância genética entre as mesmas. Entendem-se que apesar de se diferenciarem geneticamente, de acordo com a matriz par-a-par de F_{ST} , as duas populações apresentam fluxo gênico elevado ($Nm = 1,31$), que pode ser atribuído ao deslocamento de frutos entre os dois grupos feito pelos habitantes dessas regiões.

No Estado de Roraima, o bacurizeiro apresenta-se em vegetação secundária, diferentemente do Estado do Amazonas onde ocorre em vegetação primária (Carvalho & Nascimento, 2017). As populações amostradas de Nova Colina/RR e Rorainópolis/RR estão presentes em vegetação secundária e Itaquiara/AM em vegetação primária, assim como já observado em estudos anteriores para os dois Estados. As três populações demonstraram sinais de que sofreram reduções, pois conforme relatos de moradores locais as populações dessas áreas continham maior número de indivíduos, porém ocorreram eventos de queimada para expansão de área de pastagem. A diferenciação genética (F_{ST}) presente entre essas três populações geograficamente próximas foi muito alta, entre 0,28 e 0,33, sendo que a maior diferenciação foi entre Rorainópolis e Itaquiara ($F_{ST} = 0,33$). Na análise de DAPC cada uma dessas três populações agrupou separadamente, sendo que Rorainópolis apresentou a maior parte dos seus acessos presentes no grupo 4, Itaquiara no grupo 3 e Nova Colina no grupo 2, junto com acessos do Estado do Pará. As distâncias genéticas entre essas três populações são próximas, conforme visto no dendrograma e é coerente com a distância geográfica, porém apresentam alta diferenciação o que as dividiu em três grupos distintos. A elevada diferenciação mesmo em populações próximas geograficamente está relacionado com o baixo fluxo gênico observado entre essas populações, de 0,51 a 0,64. De forma semelhante foi observado para o urucum

(*Bixa orellana* Var.), uma espécie arbórea que exibiu baixo fluxo gênico entre suas populações ($Nm = 0,545$), mesmo entre populações geograficamente próximas (Dequigiovanni et al., 2018). Nesse caso, Dequigiovanni et al. (2018) concluiu que a estrutura genética é fortemente determinada pelo efeito de deriva genética quando observado o baixo fluxo gênico, assim como Wright (1931) estabelece que a taxa de $Nm = 1,0$ é o mínimo necessário para impedir que ocorra o efeito de deriva genética.

O agrupamento dos acessos de Nova Colina com os acessos das duas populações do Pará, no grupo 2, pode representar uma consequência das retrações conhecidas do “Arco Pleistocênico”, onde possivelmente essas populações se tornaram disjuntas e mantiveram um pool gênico semelhante (Prado & Gibbs, 1993; Pennington et al., 2000). Na Floresta Amazônica existe a presença de áreas isoladas e disjuntas, semelhantes às do Cerrado, presentes em áreas do Estado de Roraima próximas aos Estados do Amazonas e Pará (Pennington et al., 2000), como a população de Nova Colina. Na idade do gelo, como visto na teoria dos refúgios ecológicos (Haffer, 1982; Prance, 1973), houve a disseminação de condições de seca formando as manchas disjuntas em áreas mais úmidas (Pennington et al., 2000). Nesse sentido, as duas populações uma vez juntas no passado, e posteriormente separadas, se diferenciaram, o que pode explicar os valores da matriz par-a-par de F_{ST} com valores de 0,68 e 0,69 entre a população de Nova Colina/RR e as populações do Estado do Pará. No entanto deve-se levar em consideração que o número total de alelos (A), das duas populações do Estado do Pará são reduzidos ($A = 609$ e 676) em relação aos das demais populações (entre $A = 2310$ até 2981), pois durante a filtragem foram eliminadas sequências com baixa qualidade dessas populações. Nesse sentido, os valores de F_{ST} entre 0,68 e 0,69 podem ter se apresentado maiores do que são realmente, dada a menor quantidade de alelos utilizados na análise. Esse desbalanceamento pode ter ocorrido, e como consequência, elevando a diferenciação entre a população de Nova Colina/RR com as da Ilha do Marajó/PR e Bragança/PR.

Com relação aos parâmetros de diversidade genética, as populações avaliadas apresentaram em sua maioria valores extremamente baixos a moderados de diversidade. A população de Chapadinha no Estado do Maranhão apresentou maior índice de diversidade em relação as outras populações amostradas, seguida da população de Itaquiara, do Estado do Amazonas. As duas populações de Roraima (Nova Colina e Rorainópolis) e do Pará (Ilha do Marajó e Bragança) foram as que apresentaram índices extremamente baixos de variabilidade genética. Além dos baixos níveis de diversidade, principalmente nas populações da região Norte do bioma amazônico, os níveis de endogamia são alarmantes. A diversidade genética defasada pode indicar que essas populações são formadas em parte por clones, pois conforme Homma et al. (2010) essa situação ocorre quando os indivíduos apresentam a mesma origem materna e a autoincompatibilidade de *P. insignis* agrava essa situação.

A região Nordeste, representada nesse estudo pelos acessos de Chapadinha do bioma Cerrado, apresentou o maior valor de heterozigiosidade esperada ($H_E = 0,1746$) e índice de fixação negativo ($f = -0,1904$), valores bem superiores quando comparados às populações amostradas da região Norte, bioma Amazônia. No estudo de Nascimento et al. (no prelo), a partir de marcadores microssatélites com acessos provenientes das regiões Norte e Nordeste, foram obtidos valores superiores a estes para todas as populações, sendo que para a população de Chapadinha, a H_E foi de 0,559 e o índice de fixação de -0,131. Esse fato é esperado, considerando a diferença do número de locos utilizados e o tipo de marcador. Apesar disso, observa-se que a população de Chapadinha apresenta estruturação e ainda detém valores de diversidade consideráveis e que em ambos os estudos foi observada a ausência de endogamia. Os valores de diversidade genética encontrados dentro das populações (85,2%) inseridas na região da Reserva Extrativista da Chapada Limpa, também em Chapadinha, MA, e os índices altos de diversidade observados ($H_E = 0,133$ até 0,533) no estudo de Santos et al. (2017), indicam o bom Estado de preservação dessa área de proteção ambiental. No entanto, o autor ressalta que a diferenciação genética do bacurizeiro pode ter sofrido diretamente com as consequências dos eventos de incêndios intensos e a redução da vegetação primária. Dessa forma, temos que os valores de diversidade nas populações de Chapadinha, comparando os três estudos (presente estudo; Nascimento et al. no prelo; Santos et al. 2017) se mantem entre valores moderados a altos, dada a presença da área de proteção ambiental delimitada, como um fator de redução dos efeitos da erosão genética. Os valores moderados de diversidade observados nesse estudo evidenciam que as populações dessa área já apresentam os efeitos da ação antrópica.

As populações de Bragança e Ilha do Marajó apresentaram baixa diversidade genética com valores de H_E de 0,0424 e 0,0649, respectivamente, porém, os valores de fixação demonstraram que ambas as populações não apresentam endogamia. Sanches et al. (2016), avaliando 78 indivíduos de uma população natural de *P. insignis* de Bragança com seis marcadores ISSR, encontraram 68 % dos indivíduos identificados como clones. Dessa forma, a população no presente estudo de Bragança por apresentar variabilidade baixa pode significar um indício da presença de indivíduos que são também clones como no estudo anterior. A expressiva endogamia em populações de bacurizeiro é inédita até o momento dentre os estudos já publicados, sendo que o maior valor encontrado de fixação para a população de Itaquiara ($f = 0,5885$) evidencia um desvio no equilíbrio populacional dessa espécie arbórea. Segundo Luikart et al. (1998), a redução do tamanho efetivo de uma população no processo de gargalo genético recente, como o observado para as populações de bacurizeiro, pode estar relacionada com as queimadas frequentes observadas nas áreas de ocorrência dessa espécie. A diferenciação acelerada observada a partir dos valores de F_{ST} , como consequência do efeito de gargalo genético, indicam o isolamento geográfico da população de Chapadinha que sofre constantemente com as queimadas e o avanço da ocupação urbana. As populações da região Norte ainda apresentam

fluxo gênico entre si, como observado entre Nova Colina e Rorainópolis ($N_m = 0,64$); entre Itaquiara e Rorainópolis ($N_m = 0,51$); entre Itaquiara e Bragança ($N_m = 0,27$); e entre Itaquiara e Ilha de Marajó ($N_m = 0,20$), porém os valores baixos de diversidade e altos de endogamia possivelmente indicam o início do isolamento dessas populações.

O conhecimento a respeito do sistema reprodutivo possibilita, quando em conjunto com os dados de estrutura genética, o entendimento da dinâmica populacional. Para uma espécie predominantemente alógama com autoincompatibilidade esporofítica como o bacurizeiro (Maués et al., 1996; Cavalcante, 1996), não é esperado que ocorra baixa variabilidade e endogamia. Entretanto, de acordo com os dados apresentados existe um forte indicativo de que essas populações se apresentam isoladas e reduzidas, fazendo com que ocorra o efeito de gargalo genético progressivo, o qual se torna intenso à medida em que essas populações frágeis são submetidas a ação antrópica. De acordo com Saraiva et al. (2013), a depressão por endogamia e autopolinização podem estar relacionadas com as anormalidades nas estruturas reprodutivas de *P. insignis*. Os estudos a respeito da polinização dessa espécie ainda são escassos, com a maioria deles realizados em condições não naturais. A realização de autopolinização por *P. insignis* é uma possibilidade que deve ser considerada, indicando que essa espécie está em transição para o sistema misto (Barrett et al., 2009). Nesse sentido, a endogamia observada nessas populações fornece indícios de que o bacurizeiro teve seu sistema reprodutivo afetado devido a pressões ambientais. A ação dos polinizadores não eficiente dada a condições de degradação do habitat e à redução desses visitantes às flores do bacurizeiro, tal como a redução drástica das populações do bacurizeiro ao longo de muitos anos, levam a adaptações no sistema reprodutivo de predominantemente alógamo e com autoincompatibilidade esporofítica para um que começa a apresentar sinais de que está ocorrendo autofecundação.

6.2. A fragilidade ambiental das populações do bacurizeiro e as práticas conservacionistas sustentáveis

As populações de bacurizeiro do Nordeste Brasileiro são pertencentes ao Cerrado, podendo ser encontradas nas zonas de transição de Floresta-savana e ao sul da Floresta Amazônica (Clement & Venturieri, 1990). Nessas regiões com áreas descampadas o bacurizeiro é encontrado em maior densidade em vegetações caracterizadas como secundária de terra firme, o que foi observado para a população de Chapadinha no Estado do Maranhão. No passado essa população pertenceu a uma área de proteção ambiental (PA). Atualmente, parte da população se encontra na reserva extrativista da Chapada Limpa, local contemplado pelo presente estudo, onde os acessos coletados sofrem com o avanço urbano. O uso e a ocupação de áreas de proteção, apesar das restrições, não impede que

ocorram atividades irregulares (Morelli et al., 2009), como a exploração madeireira e abertura de roçados para a atividade agrícola. Na literatura é visto o quanto as populações de bacurizeiros maranhenses são afetadas pelo uso do glifosato, principal herbicida utilizado nas plantações de soja (*Glycine max* (L.) Merr.) (Homma et al., 2018). Nesse caso, dada a ação do glifosato as plântulas de bacurizeiro não se desenvolvem como deveriam e acarreta na não renovação dos indivíduos da população, o que pode ser atenuado pelos solos (de arenosos a argilosos) pobres e ácidos, dessa região (Souza, 2011). Desse modo, conseqüentemente, as perdas devidas à soma de todos os fatores citados e os eventos de queimadas, contribuem continuamente para perdas sem precedentes da diversidade genética, além de impedir a ação da regeneração do solo exercida pelos bacurizeiros.

A intensa exploração do bacurizeiro em seu passado madeireiro, relatado desde a década de 70, já indicava a fragilidade dessa espécie. Ao longo das décadas, a situação se agravou com a expansão das áreas urbanas, derrubadas para a formação de pastagem e o cultivo de espécies como a soja (Lustosa, 1976; Souza et al., 2000; Santos et al., 2017). As populações de *P. insignis* encontradas no Brasil concentram-se principalmente nas regiões Norte e Nordeste do país. No Estado do Pará suas populações são reconhecidas amplamente pelas populações locais e facilmente encontradas nos principais mercados de frutas. Nos Estados de Roraima e Amazonas o mesmo não se repete, pois, as populações foram extremamente reduzidas e dificilmente os frutos do bacurizeiro são encontrados nos mercados de frutas e são pouco reconhecidos pela população local. No processo de coleta dos acessos nesse estudo, os moradores antigos das áreas de coleta relataram que com o passar dos anos observaram a escassez do bacuri nessas áreas e que as grandes populações que conheciam foram queimadas para dar lugar a pastagem. Durante as coletas, a equipe comprovou tais relatos, pela dificuldade em se encontrar áreas de ocorrência da espécie. A situação mais drástica encontrada durante a busca pelas populações do bacurizeiro no Estado do Amazonas, foi no município de Novo Airão. Um morador antigo dessa região levou a equipe até uma população conhecida pelas numerosas árvores de *P. insignis*, porém ao chegar no local restava apenas um indivíduo quase inteiramente queimado e o local onde se encontravam os outros bacurizeiros deu lugar a uma área de pastagem (Figura 10).



Figura 10. Imagens feitas por satélite obtidas no Google Earth à esquerda (disponível em <http://earth.google.com/>, 2021) e à direita fotos realizadas pela equipe de coleta para o presente estudo, da única árvore remanescente da população do município de Novo Airão/AM.

As populações do bacurizeiro, quando encontradas próximas a áreas de proteção ambiental ou de Reserva Indígena, podem apresentar maior representatividade. Conforme relato de uma descendente indígena que participou de parte da coleta no Estado do Amazonas, em São Guilherme onde fica a antiga reserva à qual a família dela pertencia, as populações de bacurizeiro são bem conhecidas. O local é próximo da Estação Ecológica Juami-Japurá, o que permite que essas populações estejam menos sujeitas a ação humana do que outras populações que estão em regiões mais suscetíveis a degradação ambiental. Essa informação está de acordo com registros de coletas botânicas nessa região, que pode ser visto em GBIF: The Global Biodiversity Information Facility (<https://www.gbif.org>). A maior diversidade encontrada para a população de Chapadinha no Estado do Maranhão demonstra que mesmo quando afetada pela urbanização acelerada da região, por estar próxima a uma área de preservação, consegue manter a variabilidade dado ao possível fluxo gênico entre a população que se encontra dentro da área de preservação.

Os resultados de baixíssima diversidade encontrados dentro das populações e elevada endogamia na região Norte do Brasil é um reflexo das queimadas e desmatamento acelerado causando danos sem precedentes para a variabilidade genética dessa espécie. A população de Nova colina/RR, assim como a de Rorainópolis/RR se encontram no ultimo estagio da ação antrópico, com a formação de grandes bancos de areia devida ao empobrecimento do solo (Figura 11), que no caso já é naturalmente pobre em floresta amazônica quando afastados de corpos d'água que trazem nutrientes (Neves et al., 2001). Por tanto, medidas efetivas para a preservação dessas populações devem ser tomadas com urgência. No estudo de Nascimento et al. (no prelo) com populações naturais de regiões semelhantes às desse estudo foram sugeridas estratégias importantes que podem ser tomadas, entre elas a conscientização a respeito da derrubada dos bacurizeiros e a colheita precoce de seus frutos.

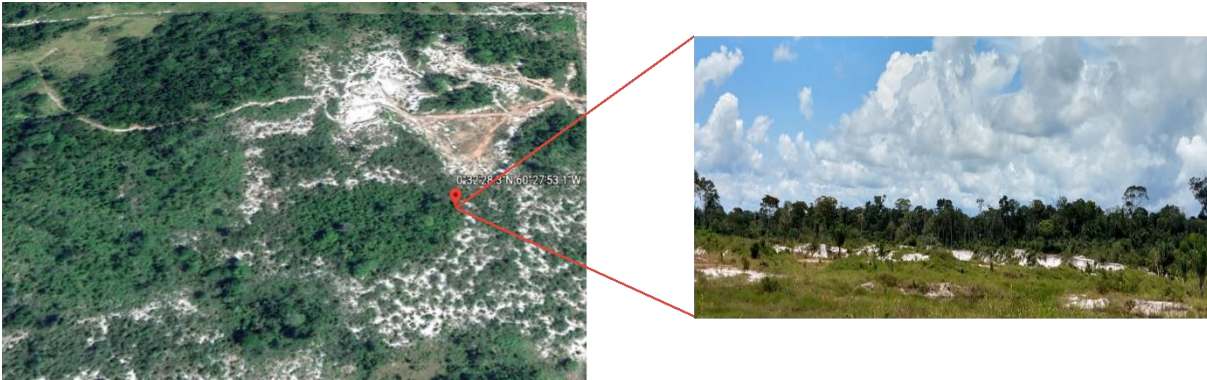


Figura 11. Imagens feitas por satélite obtidas no Google Earth à esquerda (disponível em <http://earth.google.com/>, 2021) e à direita fotos realizadas pela equipe de coleta para o presente estudo, da população do município de Nova Colina/RR.

Programas de educação ambiental enfatizando o reconhecimento das espécies frutíferas para a população local, podem ajudar no processo de conscientização, e somando a outras medidas mais voltadas a políticas públicas, podem ser mais efetivos. A derrubada dos bacurizeiros para a implementação de áreas de pastagem ou de cultivo de outras espécies de interesse agrônômico é extremamente problemático não só para a espécie *P. insignis*, mas para seu polinizador com a perda de seu habitat (Homma et al., 2007). Nesse sentido, a implementação de sistemas agroflorestais pode ser uma estratégia eficiente para a recuperação das áreas desmatadas e gerar subsídios para as populações locais. Esforços para que os bancos de germoplasma de *P. insignis* compreendam a variabilidade suficiente devem ser realizados, pois há claramente a necessidade da representatividade de acessos das populações do Estado de Roraima e Amazonas antes que essas populações sejam extintas.

7. CONCLUSÕES

- Por meio dos marcadores SNPs foram detectados índices de diversidade genética moderados para a população de Chapadinha no Estado do Maranhão e Itaquiara no Estado do Amazonas. No entanto, índices de fixação elevados e baixa diversidade alarmantes foram observadas para as populações Amazônicas, bioma que é considerado como de origem do bacurizeiro.
- As populações avaliadas do bacurizeiro pertencentes aos estados do Amazonas, Pará, Maranhão e Roraima apresentam elevada estruturação genética. Foi detectada a ocorrência de isolamento por distância. A maior parte da variabilidade genética que se concentra entre as populações de bacurizeiro, o que é inédito até o momento em estudos publicados.
- A ação antrópica intensa que tem levado à redução drástica dessas populações vem ocorrendo há anos e se manteve acelerada nas últimas décadas, reduzindo os níveis diversidade do bacurizeiro. O gargalo genético observado para essas populações e características anômalas das estruturas reprodutivas, juntamente com elevados níveis de endogamia observados nesse estudo levam ao entendimento de que é possível um modo de reprodução mista para essa espécie. O que representa, para uma espécie predominantemente alógama, a adaptação ao ambiente fragmentado e destruído.
- Novas medidas efetivas devem ser tomadas para proteção dessas populações frágeis que tem sua diversidade genética em risco, as quais não devem apenas se limitar a melhora da representatividade das populações dos estados de Roraima e Amazonas nos Bancos de germoplasma. A ampliação de estudos com maior número de populações e regiões, conscientização da população e implementação de sistemas agroflorestais para recuperação das áreas desmatadas significam um novo olhar estratégico para a manutenção da diversidade do bacurizeiro.

REFERÊNCIAS

- Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, *11*(10), 697-709.
- Alves, R., & Modesto Junior, M. D. S. (2020). Roça sem fogo: da tradição das queimadas à agricultura sustentável na Amazônia. *Embrapa Amazônia Oriental-Livro técnico (INFOTECA-E)*. Embrapa, Brasília - DF, 184.
- Araújo, E., Vasconcelos, L., Alves, R., & de Souza, V. A. B. (2018). *Platonia insignis*: bacuri. *Embrapa Meio-Norte-Capítulo em livro científico (ALICE)*. Embrapa, Brasília - DF.
- Aravanopoulos, F. A. (2014). Genomics of trees. *Tree biotechnology*. CRC Press, Boca Raton, 514-557.
- Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, *3*(10), e3376.
- Barrett, S. C., Ness, R. W., & Vallejo-Marín, M. (2009). Evolutionary pathways to self-fertilization in a tristylous plant species. *New Phytologist*, *183*(3), 546-556.
- Black IV, W. C., Baer, C. F., Antolin, M. F., & DuTeau, N. M. (2001). Population genomics: genome-wide sampling of insect populations. *Annual review of entomology*, *46*(1), 441-469.
- Bragg, J. G., Supple, M. A., Andrew, R. L., & Borevitz, J. O. (2015). Genomic variation across landscapes: insights and applications. *New Phytologist*, *207*(4), 953-967.
- Carvalho, J. E. U., & Nascimento, W. M. O. (2017). Bacuri: *Platonia insignis*. *Embrapa Amazônia Oriental-Fôlder/Folheto/Cartilha (INFOTECA-E)*.
- Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*(3), 171-182.
- Cavalcante, P. B. (1996). Frutas Comestíveis da Amazônia. 6ªed. Belém: CNPq/Museu Paraense Emílio Goeldi, 279p.
- Accessed 14 Aug. 2019.
- Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 195-205.
- Chessel, D., Dufour, A., & Thioulouse, J (2004). The ade4 Package – I: One-Table Methods. R News, v. 4(1), 5–10. Disponível em: <https://cran.r-project.org/doc/Rnews/>.
- Clement, C. R. (1999). 1492 and the loss of Amazonian crop genetic resources. II. Crop biogeography at contact. *Economic Botany*, *53*(2), 203-216.
- Clement, C. R., & Venturieri, G. A. (1990). Bacuri and cupuaçu. In: Nagy, S.; Shaw, P.E.; Wardowsky, W. (Ed.). Fruits of tropical and subtropical origin: composition, properties and uses. Lake Alfred: Florida Science Source Inc., p.178-192.

- Clement, C. R., Crito-Araujo, M., D'Eeckenbrugge, G. C., Alves-Pereira, A., Picanço-Rodrigues, D. (2010). Origin and domestication of native Amazonian Crops. *Diversity*, 2, 72-106.
- Cronquist, A. (1981) An Integrated System of Classification of Flowering Plants. Columbia University Press, New York, 248-250.
- DaCosta, J. M., & Sorenson, M. D. (2014). Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PLoS One*, 9(9), e106713.
- Daniel, J. (2004). *Tesouro descoberto no máximo rio Amazonas*. Rio de Janeiro: Contraponto, v.2, 2004. 597p.
- Davey, J. W., & Blaxter, M. L. (2010). RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, 9(5-6), 416-423.
- Davey, J. W., Hohenlohe, P. A.; Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12(7), 499-510.
- Dequigiovanni, G., Ramos, S. L. F., Alves-Pereira, A., Fabri, E. G., Picanço-Rodrigues, D., Clement, C. R., ... & Veasey, E. A. (2018). Highly structured genetic diversity of *Bixa orellana* var. *urucurana*, the wild ancestor of annatto, in Brazilian Amazonia. *PLoS one*, 13(6), e0198593.
- Dray, S., & Dufour, A. (2007). The ade4 Package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, v. 22, n. 4, p. 1–20. doi: 10.18637/jss.v022.i04.
- Dray, S., Dufour, A., & Chessel, D. (2007). The ade4 Package – II: Two-Table and K-Table Methods. *R News*, 7(2), 47–52. Disponível em: <https://cran.r-project.org/doc/Rnews/>.
- Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, 29(1), 51-63.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6(5), e19379.
- Embrapa (2019). Empresa Brasileira de Pesquisa Agropecuária, Disponível em: <<http://www.embrapa.br>>. Acesso em 24/01/2019.
- Freeland, J. R. (2005). *Molecular Ecology*. Hoboken: Wiley-Blackwell. p. 400.
- Fuentes-Pardo, A. P., & Ruzzante, D. E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Molecular Ecology*, 26(20), 5369-5406.
- GOOGLE. Google Earth website. <http://earth.google.com/>, 2021.
- Goudet, J.; & Jombart, T. (2020). hierfstat: Estimation and Tests of Hierarchical F-Statistics. R package version 0.5-7, 2020. Disponível em: <https://CRAN.R-project.org/package=hierfstat>

- Graudal, L., Aravanopoulos, F., Bennadji, Z., Changtragoon, S., Fady, B., Kjær, E. D., ... & Vendramin, G. (2014). Global to local genetic diversity indicators of evolutionary potential in tree species within and outside forests. *Forest Ecology and Management*, 333, 35-51.
- Haffer, J. (1982) General aspects of the Refuge Theory. Biological diversification in the tropics (ed. by G. T. Prance). Colombia University Press, New York, 6–24
- Hartl, D. L., & Clark, A. G. (2010). Princípios de Genética de Populações. 4ª ed. Porto Alegre: Artmed, 2010. 660p.
- Hazzouri, K. M., Escobar, J. S., Ness, R. W., Killian Newman, L., Randle, A. M., Kalisz, S., et al. (2013). Comparative population genomics in *Collinsia* sister species reveals evidence for reduced effective population size, relaxed selection, and evolution of biased gene conversion with an ongoing mating system shift. *Evolution*, 67(5), 1263–1278.
- Hazzouri, K. M., Purugganan, M. D., & Flowers, J. M. (2014). Population genomics of plant species. In *Advances in Botanical Research*, 69, 311-334.
- Hohenlohe, P. A., Amish, S. J., Catchen, J. M., Allendorf, F. W., & Luikart, G. (2011). Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular Ecology Resources*, 11, 117-122.
- Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, 6(2), e1000862.
- Hohenlohe, P. A., Day, M. D., Amish, S. J., Miller, M. R., Kamps-Hughes, N., Boyer, M. C., & Luikart, G. (2013). Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Molecular Ecology*, 22(11), 3002-3013.
- Homma, A. K. O., de Carvalho, J. E. U., & de Menezes, A. J. E. A. (2010). Bacuri: fruta amazônica em ascensão. *Embrapa Amazônia Oriental-Artigo em periódico indexado (ALICE)*. *Ciência Hoje*, 46, 271.
- Homma, A. K. O., de Carvalho, J. E. U., de Matos, G. B., & de Menezes, A. J. E. A. (2007). Manejando a planta e o homem: os bacurizeiros do nordeste paraense e da Ilha de Marajó. *Embrapa Amazônia Oriental-Artigo em periódico indexado (ALICE)*, *Amazônia: Ciência & Desenvolvimento*, 2(4), 119-135.
- Homma, A. K. O., de Menezes, A. J. E. A., de Carvalho, J. E. U., & de Matos, G. B. (2018). Manejo e plantio de bacurizeiros (*Platonia insignis* Mart.): a experiência no manejo e domesticação de um recurso da biodiversidade amazônica. *Inclusão Social*, 12(1), 48-57.

- Inglis, P. W., Pappas, M. D. C. R., Resende, L. V., & Grattapaglia, D. (2018). Fast and inexpensive protocols for consistent extraction of high quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. *PLoS One*, *13*(10), e0206085.
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403-1405. doi: 10.1093/bioinformatics/btn129.
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 94.
- Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in genetics*, *6*(208).
- Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281. doi:10.7717/peerj.281.
- Kohn, M. H., Murphy, W. J., Ostrander, E. A., & Wayne, R. K. (2006). Genomics and conservation genetics. *Trends in Ecology & Evolution*, *21*(11), 629-637.
- Krutovskii, K. V. (2006). From population genetics to population genomics of forest trees: integrated population genomics approach. *Genetika*, *42*(10), 1304.
- Lima M.C. (2007). Bacuri: (*Platonia insignis* Mart., Clusiaceae): Agrobiodiversidade. São Luis, Instituto Interamericano de Cooperação para a Agricultura.
- Lorenzi, H. (1992). Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. Nova Odessa: Plantarum.
- Loureiro, A. A., da Silva, M. F., & Alencar, J. D. C. (1979). *Essências madeireiras da Amazônia* (No. 634.9811 L892). Instituto Nacional de Pesquisas da Amazônia, Manaus (Brasil).
- Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, *4*(12), 981-994.
- Luikart, G., Sherwin, W.B., Steele, B.M., & Allendorf, F.W. (1998). Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology* *7*, 963-974.
- Lustosa, A. (1976). *No estuário amazônico: à margem da visita pastoral*. Belém: Conselho Estadual de Cultura.
- Maués, M. M., Venturieri, G. C., Souza, L. A., & Nakamura, J. (1996). Identificação e técnicas de criação de polinizadores de espécies vegetais de importância econômica no estado do Pará. In: Embrapa. *Centro de Pesquisa Agroflorestal da Amazônia Oriental (Belém, PA). Geração de tecnologia agroindustrial para o desenvolvimento do Trópico Úmido*. Belém: Embrapa-CPATU: IICA, 17-55. (Embrapa-CPATU. Documento, 85).

- Médail, F., & Baumel, A. (2018). Using phylogeography to define conservation priorities: The case of narrow endemic plants in the Mediterranean Basin hotspot. *Biological Conservation*, 224, 258-266.
- Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, 17(2), 240-248.
- Morelli, F., Setzer, A., & Jesus, S. C. (2009). Focos de queimadas nas unidades de conservação e terras indígenas do pantanal 2000- 2008 (Burning outbreaks in conservation units and indigenous lands of the Pantanal 2000-2008, in Portuguese). *Geografia*, 34, 681-695.
- Morton, J. F. (1987). *Fruits of warm climates*. JF Morton. Florida Flair Books, Miami.
- Mourão, K. S. M. M. & Beltrati, C. M. (1995) Morfologia dos frutos, sementes e plântulas de *Platonia insignis* Mart. (Clusiaceae). II. Morfoanatomia dos frutos e sementes maduros. *Acta Amazônica*, 25(1), 33-45.
- Nascimento, W. F., Dequigiovanni, G., Bertocco, C. G., & Veasey, E. A. Nuclear and chloroplast reveal high genetic diversity and structure in *Platonia insignis* Mart., an endangered species native to the Brazilian Amazon. *Acta Botânica Brasileira*, no prelo.
- Nascimento, W. M. O., Carvalho, J. E. U., & Müller, C. H. (2007). Ocorrência e distribuição geográfica do bacurizeiro. *Revista Brasileira de Fruticultura*, 29, 657–660.
- Neale, D. B., & Kremer, A. (2011). Forest tree genomics: growing resources and applications. *Nature Reviews Genetics*, 12(2), 111-122.
- Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, 106(949), 283-292.
- Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583-590.
- Neves, E. J., Martins, E. G., & Reissmann, C. B. (2001). Deposição de serrapilheira e de nutrientes de duas espécies da Amazônia. *Embrapa Florestas-Artigo em periódico indexado (ALICE, Boletim de Pesquisa Florestal)*. Colombo, 43, 47-60.
- Ouborg, N. J., Pertoldi, C., Loeschcke, V., Bijlsma, R. K., & Hedrick, P. W. (2010). Conservation genetics in transition to conservation genomics. *Trends in Genetics*, 26(4), 177-187.
- Padgham, M., & Sumner, M. D. (2020). geodist: Fast, Dependency-Free Geodesic Distance Calculations. R package version 0.0.6. <https://CRAN.R-project.org/package=geodist>
- Paradis E., & Schliep K. (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526-528.
- Paraense, L. C. R., Pena, D. N., Darnet, S. H., de Miranda Rodrigues, S., de Menezes, I. C., & Moura, E. F. (2020). First genomic microsatellite markers developed for *Platonia insignis* (Clusiaceae), a Brazilian fruit tree. *Molecular Biology Reports*, 47, 2985-2989.

- Patiño, V. M. (2002). *Historia y dispersión de los frutales nativos del Neotrópico*. Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. 665 p. (Publicación CIAT no. 326)
- Pena DN. 2016. *Divergência genética entre acessos de bacurizeiro (Platonia insignis Mart.) conservados na Embrapa Amazônia Oriental por meio de marcadores microssatélites*. Master Thesis, Universidade Federal Rural da Amazônia, Belém, Brazil.
- Pena, D. N., Moura, E. F., Rodrigues, S. M., Oliveira, M. S. P., Sanches, J. P., & Moura, M. F. (2020). Molecular characterization of a germplasm bank of *Platonia insignis* Mart.: a fruit tree. *Genetic Resources and Crop Evolution*, 67(2), 411-420.
- Pennington, T. R., Prado, D. E., & Pendry, C. A. (2000). Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography*, 27(2), 261-273.
- Pennington, T. R., Prado, D. E., & Pendry, C. A. (2000). Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography*, 27(2), 261-273.
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, 7(5), e37135.
- Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PloS One*, 7(2), e32253.
- Pontes, L. C. G., Moura, E. F., Moura, M. F., Rodrigues, S. D. M., Oliveira, M. D. S. P. D., Carvalho, J. E. U. D., & Therrier, J. (2017). Molecular characterization of progenies of bacurizeiro (*Platonia insignis*) from Marajó Island, northeastern Amazon. *Acta Amazonica*, 47(4), 293-300.
- Prado, D. E., & Gibbs, P. E. (1993). Patterns of species distributions in the dry seasonal forests of South America. *Annals of the Missouri Botanical Garden*, 80(4), 902-927.
- Prance, G. T. (1973) Phytogeographic support for the theory of Pleistocene forest refuges in the Amazon Basin, based upon evidence from distribution pattern in Caryocaraceae, Chrysobalanaceae, Dichapetalaceae and Lecythidaceae. *Acta Amazonica*, 3, 5–28.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Puritz, J. B., Matz, M. V., Toonen, R. J., Weber, J. N., Bolnick, D. I., & Bird, C. E. (2014). Demystifying the RAD fad. *Molecular Ecology*, 23(24), 5937-5942.
- R Development Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available in: <http://www.R-project.org/>.

- Sanches, J., Cunha, E., Rodrigues, S. D. M., Oliveira, M., Nahon, A., & Pena, D. (2016). Caracterização genética de uma população natural de bacurizeiro (*Platonia insignis* Mart.) por meio de marcadores ISSR. In *Embrapa Amazônia Oriental-Artigo em anais de congresso (ALICE)*. In: SEMINÁRIO DE INICIAÇÃO CIENTÍFICA, 20.; SEMINÁRIO DE PÓS-GRADUAÇÃO DA EMBRAPA AMAZÔNIA ORIENTAL, 4., 2016, Belém, PA. Anais. Belém, PA: Embrapa Amazônia Oriental, 2016.
- Santos, E. M., de Jesus Nascimento, C. A., Muniz, F. H., & Barros, J. E. D. R. S. (2017). Genetic and population diversity of bacuri (*Platonia insignis* Mart.) in Chapada Limpa Extractive Reserve, Maranhão State, Brazil. *African Journal of Biotechnology*, 16(50), 2317-2325.
- Saraiva, R. C., de Albuquerque, P. C., & Girnos, E. C. (2014). Floral and vegetative morphometrics of three *Platonia insignis* Mart.(Clusiaceae) populations, a native tree from the Brazilian Amazon. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, 148(4), 666-674.
- Savolainen, O., & Pyhäjärvi, T. (2007). Genomic diversity in forest trees. *Current Opinion in Plant Biology*, 10(2), 162-167.
- Shanley, P., Clement, C. R., Carvalho, J. E. U., de Homma, A. K. O., & Menezes, A. J. E. A. (2016). "Amazonian fruits: How farmers nurture nutritional diversity on farm and in the forest," in *Tropical Fruit Tree Diversity: Good Practices for in situ and on-Farm Conservation*, (Abingdon: Routledge), 147–160.
- Silva, S., & Donato, H. (1993). *Frutas do Brasil*. São Paulo: Imprensa de Arte, Projetos e Edições Artísticas, 50p.
- Simko, V. R & Wei, T. package "corrplot" (2017): Visualization of a Correlation Matrix (Version 0.84). Available from <https://github.com/taiyun/corrplot>.
- Slatkin, M.; & Barton, N.H. (1989). A comparison of three indirect methods for estimating average levels of gene flow. *Evolution*, 43: 1349-1368.
- Souza, I. D. B., Souza, V. A. B., & Lima, P. S. C. (2013). Molecular characterization of *Platonia insignis* Mart. ("Bacurizeiro") using inter simple sequence repeat (ISSR) markers. *Molecular Biology Reports*, 40(5): 3835-3845.
- Souza, I. G. B. (2011). Caracterização morfológica e molecular do bacurizeiro (*Platonia insignis* Mart.). MSc Thesis, Universidade Federal do Piauí, Brazil.
- Souza, I. G., de Souza, V. A. B., Damasceno, K. J., & da Costa Lima, P. S. (2016). Multivariate analysis of 'bacuri' reproductive and vegetative morphology. *Comunicata Scientiae*, 7(2), 232-240.
- Souza, V. A. B., Vasconcelos, L. F. L., & Araújo, E. C. E. (2007). Recursos genéticos do bacurizeiro na região Meio Norte do Brasil. *Bacurizeiro: agrobiodiversidade*. São Luís: IICA, 65-101.

- Souza, V. A. B., Vasconcelos, L. F. L., Araújo, E. C. E., & Alves, R. E. (2000). O bacurizeiro (*Platonia insignis* Mart.). Jaboticabal, Funep.
- Thioulouse, J., Dray, S., Dufour, A., Siberchicot, A., Jombart, O. T., & Pavoine, S. (2018). Multivariate Analysis of Ecological Data with ade4. Springer. doi: 10.1007/978-1-4939-8850-1.
- Villachica H., Carvalho J. E. U., Müller CH., Diaz SC, & Almanza M. (1996). Frutales y hortalizas promissórias de La Amazonia. Lima, Tratado de Cooperación Amazônica.
- Willing, E. M., Dreyer, C., & Van Oosterhout, C. (2012). Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PloS One*, 7(8), e42649.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16(2), 97.