



**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRICULTURA E BIODIVERSIDADE**

**ESTUDOS GENÉTICOS-GENÔMICOS EM *Hancornia speciosa*
GOMES: ÁRVORE MEDICINAL E FRUTÍFERA NATIVA DA
AMÉRICA DO SUL**

VALDINETE VIEIRA NUNES



**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRICULTURA E BIODIVERSIDADE**

VALDINETE VIEIRA NUNES

**ESTUDOS GENÉTICOS-GENÔMICOS EM *Hancornia speciosa* GOMES: ÁRVORE
MEDICINAL E FRUTÍFERA NATIVA DA AMÉRICA DO SUL**

Tese apresentada à Universidade Federal de Sergipe, como parte das exigências do Curso de Doutorado em Agricultura e Biodiversidade, área de concentração em Agricultura e Biodiversidade, para obtenção do título de “Doutora em Ciências”.

Orientadora
Prof^a. Dr^a. Renata Silva-Mann

SÃO CRISTÓVÃO
SERGIPE – BRASIL
2021

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE**

N972e Nunes, Valdinete Vieira.
Estudos genéticos-genômicos em *Hancornia speciosa* Gomes:
árvore medicinal e frutífera nativa da América do Sul / Valdinete
Vieira Nunes; orientadora Renata Silva-Mann. – São Cristóvão,
SE, 2021.
115 f.: il.

Tese (doutorado em Agricultura e Biodiversidade) –
Universidade Federal de Sergipe, 2021.

1. Mangaba. 2. Bibliometria. 3. Genômica. I. Silva-Mann, Renata,
orient. II. Título.

CDU 634.61

VALDINETE VIEIRA NUNES

**ESTUDOS GENÉTICOS-GENÔMICOS EM *Hancornia speciosa* GOMES: ÁRVORE
MEDICINAL E FRUTÍFERA NATIVA DA AMÉRICA DO SUL**

Tese apresentada à Universidade Federal de Sergipe, como parte das exigências do Curso de Doutorado em Agricultura e Biodiversidade, área de concentração em Agricultura e Biodiversidade, para obtenção do título de “Doutora em Ciências”.

APROVADA em 25 outubro de 2021.

Prof. Dr. Arie Fitzgerald Blank
UFS

Dr^a. Elisa Serra Negra Vieira
Embrapa

Prof^a. Dr^a. Heloisa Oliveira dos Santos
UFLA

Prof^a. Dr^a. Maria Imaculada Zucchi
IAC-APTA

Prof^a. Dr^a. Renata Silva-Mann
UFS
(Orientadora)

SÃO CRISTÓVÃO
SERGIPE – BRASIL

*A minha mãe Elenilde (in memoriam), minha
primeira professora, meu exemplo de
generosidade, humildade e amor.*
Dedico

AGRADECIMENTOS

Agradeço a DEUS, o meu Senhor e Salvador! Criador inefável, a Vós entrego todos os meus dias, planos e sonhos, para que tudo seja feito conforme a Tua vontade.

À minha família, pelo amor e compreensão. Cada um de vocês contribuiu enormemente em algum momento para que eu pudesse estudar, sem o apoio de vocês não seria possível trilhar esse caminho.

À minha orientadora, Dra. Renata Mann, pela orientação, confiança, motivação e por todos os ensinamentos.

À Universidade Federal de Sergipe, ao Programa de Pós-Graduação em Agricultura e Biodiversidade e as Agências de Fomento, CAPES, CNPq e FAPITEC/SE, pela estrutura e suporte financeiro para o desenvolvimento dessa tese.

Ao Programa Nacional de Cooperação Acadêmica (PROCAD) - (AUXPE 2971/2014) pela oportunidade de realizar o doutorado sanduíche, em especial agradeço aos professores Dr. José Baldin Pinheiro (ESALq/USP) e Dra. Maria Imaculada Zucchi (IAC/APTA) pelo acolhimento e orientação durante o desenvolvimento das atividades.

À formidável Banca Examinadora composta por minha orientadora Profa. Dra. Renata Mann, Prof. Dr. Arie Blank (UFS), Profa. Dra. Maria Zucchi (IAC/APTA), Dra. Elisa Serra Negra (Embrapa Florestas) e Profa. Dra. Heloisa Oliveira (UFLA), pela disponibilidade em compartilhar conhecimento e contribuir para melhorar a tese.

Ao Prof. Dr. Fernando Pedroni (UFMT), Prof. Dr. Sérgio Martins (UFS), a Doutoranda Luíse Andrade (PPGAGRI/UFS) e ao Eng. Agrônomo João Pedro Ferreira pela generosidade e apoio para a realização das coletas.

Aos amigos do Grupo de Pesquisa em Conservação, Melhoramento e Gestão de Recursos Genéticos (GENAPLANT), em especial a Natali Santana, Sara de Pádua, Izabel de Jesus e Lucas Alexandre Rocha.

Às lindas e queridas amigas Juliana Lopes, Crislaine Calazans, Maria Fernanda Torres e Sara Dayan, com vocês tive a oportunidade e alegria de compartilhar muitos momentos agradáveis e conhecimento durante a pós-graduação. Obrigada meninas pelo carinho, paciência e apoio, sobretudo nos dias mais difíceis.

A todos os professores e colegas do PPGAGRI.

A cada pessoa que em algum momento contribuiu para a realização deste trabalho, meu muito obrigada.

SUMÁRIO

LISTA DE FIGURAS	i
LISTA DE TABELAS	iii
LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS	iv
RESUMO	vi
ABSTRACT	vii
1. INTRODUÇÃO GERAL	8
2. REVISÃO DE LITERATURA	10
2.1 Bioeconomia e recursos genéticos	10
2.1.1 <i>Hancornia speciosa</i> Gomes	12
2.2 Diversidade em recursos genéticos vegetais.....	15
2.3 Inter Simple Sequence Repeat (ISSR)	17
2.4 <i>Single Nucleotide Polymorphism</i> (SNP).....	19
2.5 Modelos de distribuição de espécies (SDM)	22
3. REFERÊNCIAS BIBLIOGRÁFICAS	28
4. ARTIGO 1: POTENCIAL FARMACOLÓGICO, ALIMENTÍCIO E DADOS MOLECULARES DE <i>Hancornia speciosa</i> : UMA REVISÃO SISTEMÁTICA.....	40
RESUMO.....	40
ABSTRACT.....	41
4.1. Introdução	41
4.2. Material e Métodos	42
4.2.1. Mapeamento científico.....	44
4.2.2. Mapeamento tecnológico	44
4.2.3 Mapeamento de dados moleculares.....	44
4.3. Resultados.....	44
4.3.1. Mapeamento científico	44
4.3.2 Mapeamento tecnológico	50
4.3.3 Mapeamento de dados moleculares.....	52
4.4. Discussão	53
4.5. Conclusões	57
4.6. Referências Bibliográficas	57
5. ARTIGO 2: DIVERSIDADE GENO-FENOTÍPICA EM UMA POPULAÇÃO NATURAL DE <i>Hancornia speciosa</i> GOMES: IMPLICAÇÕES PARA A CONSERVAÇÃO E MELHORAMENTO	65
RESUMO.....	65
ABSTRACT.....	66
5.1. Introdução	66
5.2. Material e Métodos	67
5.2.1 Amostragem da população	67
5.2.2 Variabilidade de frutos, sementes e plântulas	68
5.2.3 Análise molecular.....	69
5.2.4 Análises populacionais.....	69
5.3. Resultados.....	70
5.3.1 Variabilidade de frutos, sementes e plântulas	70
5.3.2 Análises populacionais.....	74
5.4. Discussão	75
5.5. Conclusões	79
5.6. Referências Bibliográficas	79

6. ARTIGO 3: GENÔMICA E ESTRUTURA GENÉTICA DE POPULAÇÕES DE <i>Hancornia speciosa</i> : FONTE DE ALIMENTOS FUNCIONAIS E FÁRMACOS	84
RESUMO.....	84
ABSTRACT.....	85
6.1. Introdução	85
6.2. Material e Métodos	86
6.2.1 Modelos de distribuição de espécies	86
6.2.2 Populações amostradas para análises moleculares.....	88
6.2.3 Extração do DNA, preparação e sequenciamento da biblioteca genômica.....	89
6.2.4 Processamento das sequências, filtragem e identificação dos SNPs.....	90
6.2.5 Diversidade e estrutura genômica	90
6.2.6 Detecção e caracterização de SNPs <i>outliers</i>	91
6.3 Resultados	91
6.3.1 Modelos de distribuição de espécies	91
6.3.2 Diversidade genômica e estrutura genética	92
6.3.3 Caracterização de SNPs supostamente sob seleção	97
6.4 Discussão	98
6.5. Conclusão.....	102
6.6 Referências Bibliográficas	102
7. CONSIDERAÇÕES FINAIS	108
ANEXOS	109

LISTA DE FIGURAS

ARTIGO 1

Figura		Página
1	Fluxograma do processo de mapeamento científico, tecnológico e de dados moleculares para <i>Hancornia speciosa</i> Gomes.	43
2	Rede de cocitação de documentos relacionados com <i>Hancornia speciosa</i> Gomes.	45
3	Mapa mundial de colaboração para publicações sobre <i>Hancornia speciosa</i> Gomes.	46
4	Mapa temático de publicações com <i>Hancornia speciosa</i> Gomes.	47
5	Códigos de barras de gene de cloroplasto - <i>rbcL</i> (A) e gene <i>matK</i> (B) depositados no Barcode of Life Data System para a espécie <i>Hancornia speciosa</i> Gomes. Fonte: BOLD System, 2020.	52
6	Estrutura tridimensional de proteínas modeladas por homologia para <i>Hancornia speciosa</i> Gomes. A. Proteína de montagem do fotossistema I; B. Subunidade c da ATP sintase. C. Subunidade 9 do centro de reação do fotossistema I; D. Preteína Z do centro de reação; E. Centro ferro-enxofre do fotossistema I; F. Proteína T do centro de reação do fotossistema II; G. Subunidade 6 do citocromo b6-f; H. Proteína D do fotossistema II. Fonte: SIB, 2020.	53

ARTIGO 2

Figura		Página
1	Distribuição espacial de indivíduos de <i>Hancornia speciosa</i> Gomes em área de coleta, município de Pirambu, Sergipe, Brasil.	68
2	Frutos (a), coloração da polpa (b) e variáveis biométricas (c) para frutos de <i>Hancornia speciosa</i> Gomes obtidos em população natural.	70
3	Variância fenotípica no tamanho de frutos de <i>Hancornia speciosa</i> Gomes obtido em população natural.	70
4	Massa da polpa (MP), massa de sementes (MS), número de sementes (NS) e rendimento da polpa (RP) para frutos de <i>Hancornia speciosa</i> Gomes obtidas em população natural.	71
5	Correlação de Spearman para as variáveis diâmetro longitudinal (DL), diâmetro transversal (DT), massa de polpa (MP), massa de sementes (MS), número de sementes (NS) e rendimento da polpa (RP), para frutos de <i>Hancornia speciosa</i> Gomes.	72
6	Sementes (a), embrião (b) e variáveis biométricas (c) para sementes de <i>Hancornia speciosa</i> Gomes obtidas em população natural.	72
7	Plântula normal (a), plântulas anormais (b), poliembrionia (c) e variáveis biométricas (d) para plântulas de <i>Hancornia speciosa</i> Gomes obtidas em população natural.	73
8	Correlação de Spearman para as variáveis comprimento (C), largura (L) e espessura (E) das sementes e comprimento da parte aérea (CPA), da raiz (CR) e massa seca de plântulas (MSP) de <i>Hancornia speciosa</i> Gomes.	73
9	Dendrograma gerado pelo método UPGMA, baseado na similaridade de Jaccard, para população natural de <i>Hancornia speciosa</i> Gomes.	74
10	Análise de Coordenadas Principais (PCoA) para população natural de <i>Hancornia speciosa</i> Gomes.	75

ARTIGO 3

Figura		Página
1	Registro da ocorrência de <i>Hancornia speciosa</i> na América do Sul após limpeza das coordenadas.	87
2	Variáveis bioclimáticas que foram incluídas no estudo ($r \leq 0,70$).	88
3	Mapa ilustrando as populações naturais de <i>Hancornia speciosa</i> que foram amostradas.....	89
4	Distribuição geográfica potencial de <i>Hancornia speciosa</i> , considerando o cenário climático atual. A distribuição potencial é indicada por um gradiente de cor variando de vermelho (alta adequabilidade) a azul (baixa adequabilidade).	92
5	Estimativas de diversidade genômica e endogamia baseada em 4.691 SNPs neutros de seis populações naturais de <i>Hancornia speciosa</i>	94
6	Heatmap para as estimativas de F_{ST} par-a-par baseadas em 4.691 marcadores SNP neutros para seis populações naturais de <i>Hancornia speciosa</i> . Regiões de coloração mais intensa na figura indicam maior F_{ST} ...	95
7	Análise Bayesiana com marcadores SNPs para 74 indivíduos de <i>Hancornia speciosa</i> . Probabilidades de atribuição dos indivíduos aos agrupamentos com base em 4.691 marcadores SNP neutros (A) e com base em 355 marcadores SNP outliers (B).....	96
8	Análises Discriminantes de Componentes Principais (DAPC) com marcadores SNPs para 74 indivíduos de <i>Hancornia speciosa</i> . Gráfico de dispersão (A) e gráfico de densidade (C) com base em 4.691 marcadores SNPs com comportamento neutro; Gráfico de dispersão (B) e gráfico de densidade (D) com base em 355 marcadores SNPs com comportamento outlier.	97
9	Resumo da análise de 355 SNPs supostamente sob seleção no Blast2GO para populações naturais de <i>Hancornia speciosa</i>	98

LISTA DE TABELAS

ARTIGO 1

Tabela		Página
1	Classificação Internacional de Patentes para produtos tecnológicos desenvolvidas com <i>Hancornia speciosa</i> Gomes.	51

ARTIGO 2

Tabela		Página
1	Primers ISSR usados para o estudo de diversidade de <i>Hancornia speciosa</i> Gomes e suas respectivas sequências e temperatura de anelamento (T° A).	69

ARTIGO 3

Tabela		Página
1	Parâmetros de diversidade genômica usando 4.691 SNPs neutros de seis populações naturais de <i>Hancornia speciosa</i> . Número de indivíduos (N), número médio de alelos (A), número de alelos efetivos (Ne) e alelos privados (Ap).	93
2	Análise de variância molecular (AMOVA) baseada em 4.691 marcadores SNP neutros para seis populações naturais de <i>Hancornia speciosa</i> e apenas para as populações da região Nordeste.	94
3	Identificação e anotação funcional dos marcadores SNPs supostamente sob seleção para populações naturais de <i>Hancornia speciosa</i>	98

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

A	Número médio de alelos
AGEs	Produtos finais da glicação avançada
AMOVA	Análise de variância molecular
Ap	Alelos privados
BIC	<i>Bayesian Information Criterion</i>
BLAST	<i>Basic Local Alignment Search Tool</i>
BOLD System	<i>Barcode of Life Data System</i>
C	Comprimento
CPA	Comprimento da parte aérea
cpDNA	Genoma cloroplastidial
CR	Comprimento da raiz
DAPC	<i>Discriminant Analysis of Principal Componentes</i>
DL	Diâmetro longitudinal
DNA	Ácido desoxirribonucleico
DT	Diâmetro transversal
E	Espessura
ENM	<i>Ecological Niche Model</i>
EPO	<i>European Patent Office</i>
EST	<i>Expressed Sequence Tags</i>
EUA	Estados Unidos da América
f	Coefficiente de endogamia de Wright
GBIF	<i>Global Biodiversity Information Facility</i>
GBS	<i>Genotyping-by-sequencing</i>
GenBank	<i>Genetic Sequence Database</i>
GO	<i>Gene Ontology</i>
H_E	Heterozigosidade esperada
H_O	Heterozigosidade observada
I	Índice de Shannon
IBGE	Instituto Brasileiro de Geografia e Estatística
IDT	<i>Integrated DNA Technologies</i>
INPI	Instituto Nacional da Propriedade Industrial
ISSR	<i>Inter Simple Sequence Repeat</i>
L	Largura
LATIPAT- Espacenet	Pesquisa de Patentes da Espanha e América Latina
$MaxEnt$	Entropia máxima
MCMC	<i>Markov Chain Monte Carlo</i>
MCTIC	Ministério da Ciência, Tecnologia, Inovações e Comunicações
MMA	Ministério do Meio Ambiente
MP	Massa da polpa
MS	Massa de sementes
MSP	Massa seca de plântulas
N	Número de indivíduos
N_a	Número médio de alelos observados
N_e	Tamanho efetivo da população
NGS	<i>Next Generation Sequencing</i>
NS	Número de sementes
ODMAP	<i>Overview, Data, Model, Assessment, Prediction</i>

ODS	Objetivos de Desenvolvimento Sustentável
PCoA	<i>Principal Coordinate Analysis</i>
PCR	<i>Polymerase Chain Reaction</i>
PEVS	Produção da Extração Vegetal e da Silvicultura
PFNM	Produto Florestal Não Madeireiro
RGV	Recursos Genéticos Vegetais
ROC	<i>Receiver Operating Characteristic</i>
RP	Rendimento de polpa
SDM	<i>Species Distribution Model</i>
SIB	<i>Swiss Institute of Bioinformatics</i>
SNP	<i>Single Nucleotide Polymorphism</i>
SVM	<i>Support Vector Machine</i>
TSS	<i>True Sill Statistic</i>
UBC	University of British Columbia
UPGMA	<i>Unweighted Pair-Group Method Arithmetic Average</i>
WIPO	<i>World Intellectual Property Organization</i>

RESUMO

NUNES, Valdinete Vieira. **Estudos genéticos-genômicos em *Hancornia speciosa* Gomes: árvore medicinal e frutífera nativa da América do Sul.** São Cristóvão: UFS, 2021. 115p. (Tese – Doutorado em Agricultura e Biodiversidade).*

Hancornia speciosa é uma árvore frutífera e medicinal nativa da América do Sul, no Brasil compõe a lista de plantas prioritárias para o desenvolvimento de pesquisas. Nesse contexto, o presente estudo teve como objetivos: i) mapear o estado da arte por meio de artigos, patentes e dados moleculares da espécie e elencar perspectivas para novas pesquisas; ii) avaliar a variabilidade fenotípica de frutos, sementes e plântulas e estimar a diversidade genética em população natural de Sergipe, visando a conservação e melhoramento genético; e iii) adotar abordagem integrada de modelos de distribuição de espécie e genômica populacional baseada em polimorfismo de um único nucleotídeo (SNP), para identificar áreas com adequabilidade ambiental, a diversidade genômica e estrutura genética das populações naturais de *H. speciosa*. A produção científica, tecnológica e disponibilidade de dados moleculares foram prospectados em bancos públicos. Em uma população natural as variáveis fenotípicas de frutos, sementes e plântulas foram mensuradas e a diversidade genética estimada por marcadores moleculares de repetição de sequência inter-simplicada. Com base em registros de ocorrências e variáveis bioclimáticas usando modelos de distribuição foram identificadas áreas de maior adequabilidade e amostradas seis populações naturais, que foram genotipadas por sequenciamento para análise da diversidade e estrutura genômica. A partir dos dados compilados e analisados, o Brasil é o país de origem das 310 publicações e 29 patentes que foram produzidas, majoritariamente por pesquisadores afiliados a universidades públicas. Os dados moleculares disponíveis em bancos públicos para a espécie são sequências de nucleotídeos e proteínas do genoma cloroplastidial. Na caracterização fenotípica em população nativa os frutos, sementes e plântulas possuem variação a ser explorada. Quanto aos parâmetros genéticos populacionais, obteve-se: $H_e = 0,40$, $I = 0,50$, $N_a = 2,0$, $N_e = 1,71$ e 55,29% de similaridade. As áreas com maior adequabilidade (0,75 a 1,00) estão na Mata Atlântica (litoral nordestino) e Cerrado (Brasil central). Foram identificados 5.046 SNPs (4.691 neutros e 355 *outliers*), os quais permitiram inferir que existe alta diversidade e excesso de heterozigotos em todas as populações e 315 alelos privados. Maior variação genética ocorre dentro das populações e existe diferenciação genética entre as populações do Nordeste e Centro-Oeste. Para a estrutura genética (SNPs neutros) foram formados três grupos e a população de Coxim é a mais divergente. O conjunto de SNPs *outliers* revela uma estruturação genética formada por três grupos que estão correlacionados com os biomas em que estão inseridas as populações. Para dois SNPs *outliers* foi possível identificar anotação funcional. A compilação e análise das informações científicas, tecnológicas e moleculares realizadas permitiram identificar as linhas de pesquisas desenvolvidas, as lacunas do conhecimento sobre a espécie e contribuem para direcionar novas pesquisas. A população de Sergipe apresentou alta diversidade genética e variabilidade fenotípica com potencial para fornecer propágulos para o enriquecimento de populações naturais e restauração florestal. Os resultados obtidos pela combinação do modelo de distribuição de espécies e genômica populacional contribuíram para avançar no conhecimento sobre a espécie e para o desenvolvimento de estratégias visando o manejo, conservação e melhoramento genético.

Palavras-chave: bibliometria, bioeconomia, genômica populacional, genotipagem por sequenciamento, mangaba, SNP.

* Comitê Orientador: Renata Silva-Mann – UFS (Orientadora).

ABSTRACT

NUNES, Valdinete Vieira. **Genetic-genomic studies in *Hancornia speciosa* Gomes: medicinal and fruit tree native to South America.** São Cristóvão: UFS, 2021. 115p. (Thesis - Doctor of Science in Agriculture and Biodiversity).*

Hancornia speciosa is a fruit and medicinal tree native to South America. In Brazil it makes up the list of priority plants for research development. In this context, the present study aimed to i) map state of the art through articles, patents, and molecular data of the species and list perspectives for further research; ii) evaluate the fruits, seeds, and seedlings phenotypic variability, and estimate the genetic diversity in the natural population of Sergipe, aiming at conservation and genetic improvement; and iii) adopt an integrated approach of species distribution models and population genomics based on single nucleotide polymorphism (SNP) to identify areas with environmental suitability, genomic diversity and natural *H. speciosa* populations genetic structure. The scientific and technological production and availability of molecular data prospected in public databases. In a natural population, the fruits, seeds, and seedlings phenotypic variables were measured, and the genetic diversity was estimated by inter simple sequence repeat markers. Based on occurrence records and bioclimatic variables using distribution models, areas of greatest suitability were identified, and six natural populations were sampled, which were genotyped by sequencing in genomic diversity and structure analysis. Based on the compiled and analyzed data, Brazil is the origin country of 310 publications and 29 patents that were produced mostly by researchers with public universities. The molecular data available in public banks are sequences for the specie are nucleotides and proteins from the chloroplastid genome. In the phenotypic characterization in the native population, the fruits, seeds, and seedlings have a variation to be explored. As for the population genetic parameters, the following were obtained: $H_e = 0.40$, $I = 0.50$, $N_a = 2.0$, $N_e = 1.71$ and 55.29% similarity. The areas with the greatest suitability (0.75 to 1.00) are the Atlantic Forest (northeastern coast) and Cerrado (central Brazil). We identified 5,046 SNPs (4,691 neutrals and 355 outliers), which allowed us to infer that there is high diversity and excess of heterozygotes in all populations and 315 private alleles. Greater genetic variation occurs within populations and there is genetic differentiation between populations in the Northeast and Midwest. For the genetic structure (neutral SNPs), three groups were formed and the population of Coxim is the most divergent. The set of outliers SNPs reveals a genetic structure formed by three groups that are correlated with the biomes in which the populations are inserted. For two outliers SNPs it was possible to identify functional annotation. The compilation and scientific, technological, and molecular information analysis carried out in this work allowed us to identify the lines of developed research, the gaps in knowledge about the specie and contribute to directing new research. The population of Sergipe showed high genetic diversity and phenotypic variability with the potential to provide propagules for the enrichment of natural populations and forest restoration. The results obtained by the combination of distribution models and population genomics contributed to advancing knowledge about the specie and developing strategies aimed at management, conservation, and genetic improvement.

Key-words: bibliometrics, bioeconomy, population genomics, sequencing genotyping, mangaba, SNP.

* Supervising Committee: Renata Silva-Mann – UFS (Advisor)

1. INTRODUÇÃO GERAL

Desafios como o aumento na demanda global por alimentos e o impacto potencial das mudanças climáticas, criaram a necessidade de repensar as estratégias de conservação adotadas pelos países (UNITED NATIONS, 2019). Nesse contexto, na última década o conceito de bioeconomia vem sendo discutido e apresentado como parte da solução para tais desafios, especificamente por propiciar uma mudança na sociedade e nos meios de produção, bem como pela aplicação da economia focada na utilização dos recursos de base biológica (LEWANDOWSKI, 2018).

A bioeconomia vai ao encontro da Agenda 2030 composta por 17 Objetivos de Desenvolvimento Sustentável (ODS) e 169 metas que foram estabelecidas pelas Nações Unidas (UNITED NATIONS, 2021). Países que fazem parte da União Europeia e os Estados Unidos foram os primeiros a formular suas estratégias nacionais de bioeconomia; e mais recente essa é uma agenda que vem sendo adotada por outros governos (EUROPEAN UNION, 2012; WHITE HOUSE, 2012).

Na América do Sul a bioeconomia tem enfoque no uso da biodiversidade, visto que nesse continente estão localizados seis *hotspots* de biodiversidade mundiais, a Floresta Amazônica, Mata Atlântica, Cerrado, Florestas Chilenas de Inverno-Valdivianas, Tumbes-Chocó-Magdalena e Andes Tropicais. No Brasil, a bioeconomia foi incluída como um dos eixos estratégicos a serem desenvolvidos e as ações estão pautadas na integração da conservação, consumo e produção para promover o desenvolvimento sustentável (MCTIC, 2016).

Tratando especificamente de árvores nativas, muitas espécies estão sendo perdidas, antes mesmo que sejam identificadas ou ainda muitas delas são subutilizadas. Para desenvolver ações propositivas visando a conservação e uso sustentável de espécies florestais, a caracterização é uma etapa fundamental, que engloba a correta identificação botânica, avaliação da variabilidade fenotípica, o estabelecimento de descritores morfológicos e seus usos. Tradicionalmente a caracterização ocorre por meio da avaliação de características morfológicas, entretanto, marcadores moleculares são ferramentas biotecnológicas que podem contribuir nesse processo para o conhecimento da taxonomia, estrutura e diversidade genética (HEIDARI *et al.*, 2019).

A diversidade genética tem papel importante porque possibilita a evolução e adaptação das espécies às mudanças bióticas e abióticas do ambiente, bem como a resistência a novos patógenos. É uma informação tão relevante que, em estudo sobre o estado atual dos recursos genéticos florestais mundiais, foram propostos Indicadores de Diversidade Genética de Árvores que abrangem, entre outros, o uso de marcadores moleculares (GRAUDAL *et al.*, 2020) que possibilitam a estimativa de parâmetros genéticos úteis para conhecer a situação atual da população bem como sua capacidade de manutenção nas gerações posteriores (BRUFORD *et al.*, 2017).

Os avanços na ciência permitiram a descoberta de diferentes marcadores moleculares que são aplicados em vários estudos, sobretudo em genética de populações. Dentre os marcadores amplamente utilizados, tem-se ISSR (*Inter Simple Sequence Repeat*) que são informativos e aplicados com frequência em estudos com espécies florestais, isso se deve ao fato do alto número de polimorfismos, que podem ser obtidos com esse marcador, sua alta reprodutibilidade e por não necessitar do conhecimento prévio do genoma da espécie (GOWAYED; ABD EL-MONEIM, 2021; ZIETKIEWICZ; RAFALSKI; LABUDA, 1994).

Marcadores SNP (*Single Nucleotide Polymorphism*) também são muito utilizados, esses marcadores são abundantes e podem ocorrer em diferentes regiões do genoma. Diferentes técnicas e métodos são adotadas para a genotipagem por SNP e, mais recente, as tecnologias de Sequenciamento de Nova Geração (*Next Generation Sequencing*, NGS), juntamente com ferramentas de bioinformática, são eficientes e robustas na descoberta de SNPs (TURCHETTO-ZOLET *et al.*, 2017).

Entre as abordagens, a genotipagem por sequenciamento (*Genotyping-by-sequencing - GBS*), um método de genotipagem robusto baseado no sequenciamento de representações parciais do genoma, tem sido aplicada em estudos populacionais, caracterização de germoplasma e genética de plantas, especificamente para espécies que não possuem um genoma de referência (HUANG, Y.-F. *et al.*, 2014).

Outra ferramenta que pode contribuir no planejamento de ações para conservação e o uso sustentável de espécies são os Modelos de Distribuição de Espécies (*Species Distribution Models - SDMs*), ferramentas poderosas que usam associações entre ocorrências e condições ambientais para estimar áreas com maior adequabilidade para espécies (FENG *et al.*, 2019).

Hancornia speciosa figura entre as muitas espécies que possuem valor potencial ou futuro e que podem contribuir para a estratégia nacional de bioeconomia. Suas populações naturais são amplamente distribuídas pelo país, utilizadas como fonte de alimentos e medicamentos. Conhecer a diversidade genética é um fator determinante para sua conservação e uso sustentável. Por ser uma espécie alógama e autoincompatível requer genótipos geneticamente diferentes para cruzamentos, produção de frutos e manutenção das populações (DARRAULT; SCHLINDWEIN, 2005).

Estudos para avaliar a diversidade de populações naturais foram realizados por meio de avaliações das características morfológicas (LÚCIO *et al.*, 2019) ou empregando marcadores ISSR (LUZ *et al.*, 2020). Entretanto, essas análises foram realizadas de forma isolada, sendo relevante estudos que combinem essas duas abordagens concomitantemente. Na literatura atual são poucos os estudos considerando populações localizadas em diferentes regiões do país; e marcadores SNPs não foram descritos e utilizados em estudos com a espécie até o momento.

Considerando a sua importância ecológica, alimentícia e farmacológica, bem como sua ampla ocorrência nos ecossistemas brasileiros e exploração sem plano de manejo, o presente estudo teve como objetivos: i) mapear o estado da arte por meio de artigos, patentes e dados moleculares da espécie visando elencar perspectivas para novas pesquisas; ii) avaliar a variabilidade fenotípica de frutos, sementes e plântulas e estimar a diversidade genética em população natural de Sergipe por meio de marcadores moleculares ISSR, visando fornecer informações que contribuam para a conservação e melhoramento genético; e iii) adotar uma abordagem integrada de SDM e genômica populacional baseada em SNPs, para identificar áreas com adequabilidade ambiental, a diversidade genômica e estrutura genética das populações naturais da árvore *H. speciosa*.

As hipóteses propostas foram que: i) existe uma baixa disponibilidade de dados em bancos públicos (científicos tecnológicos e moleculares) sobre a espécie; ii) alta diversidade genética em suas populações naturais e iii) diferenciação genética entre as populações das regiões Nordeste e Centro-Oeste.

2. REVISÃO DE LITERATURA

2.1 Bioeconomia e recursos genéticos

Estimativas das Nações Unidas apontam que o tamanho da população global será entre 8,5 e 8,6 bilhões em 2030 e para 2050 entre 9,4 e 10,1 bilhões (UNITED NATIONS, 2019). Esse aumento da população tem implicações importantes para o desenvolvimento econômico, social e para a sustentabilidade ambiental mundial, sendo agravados problemas como a fome, a desnutrição e a exploração de recursos naturais. Diante de tal cenário, governos mundiais, organizações civis e empresas têm como desafios alimentar uma população crescente, mitigar os impactos ambientais e sociais, enquanto buscam o crescimento econômico (MIRANDA; RUIZ-CABELLO; HURTADO, 2020).

Como parte da solução para estes desafios globais surge a bioeconomia, que consiste na evolução para um novo sistema econômico baseado no uso sustentável dos recursos biológicos e que engloba a produção a partir da base biológica e sua conversão em alimentos, rações e bioenergia (AGUILAR; WOHLGEMUTH; TWARDOWSKI, 2018; LEWANDOWSKI, 2018).

Existem diferentes políticas e estratégias relacionadas à bioeconomia que foram desenvolvidas por países altamente industrializados, economias em transição e países em desenvolvimento (Figura 1). Assim, não existe uma definição universalmente aceita de bioeconomia, visto que as definições em cada país variam conforme os ecossistemas e modelos socioeconômicos (AGUILAR; TWARDOWSKI; WOHLGEMUTH, 2019; AGUILAR; WOHLGEMUTH; TWARDOWSKI, 2018).

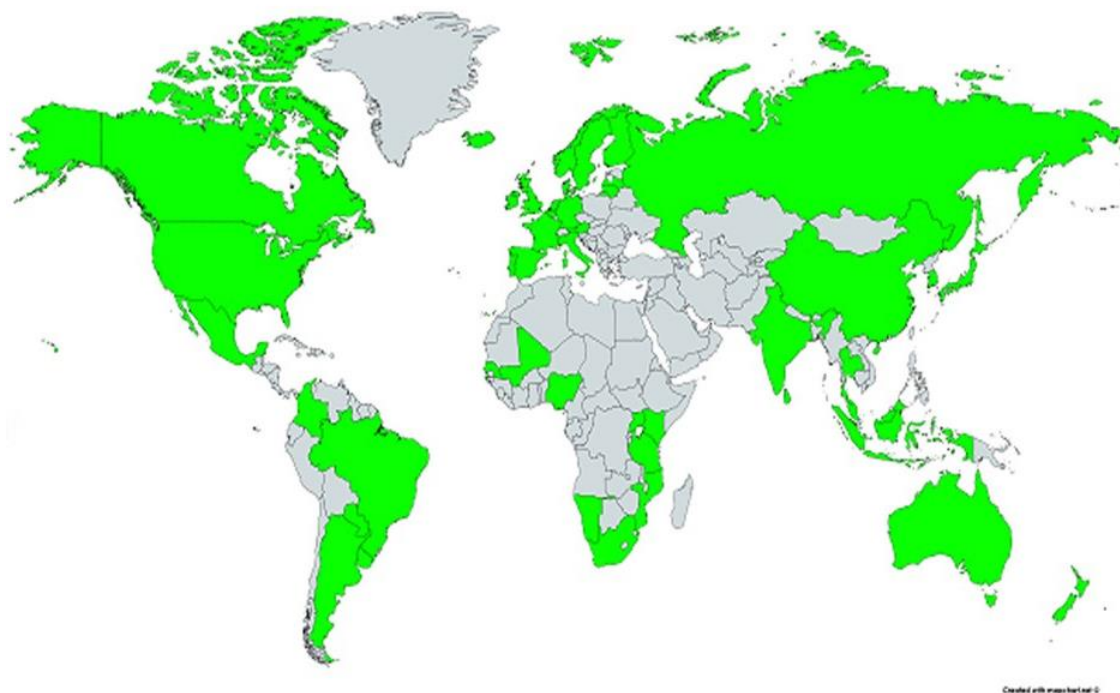


Figura 1. Países (verde) que possuem estratégias de política de bioeconomia. Fonte: Aguilar, Wohlgemuth e Twardowski (2019).

Em se tratando da agenda global para alcançar os Objetivos de Desenvolvimento Sustentável (ODS) das Nações Unidas, a adoção das estratégias de bioeconomia pelos países, de acordo com especialistas de instituições europeias e organizações internacionais, tem efeitos positivos sobre todos os objetivos, especialmente para os seguintes ODS's: 12 - Consumo e Produção Responsáveis, 9 - Indústria, Inovação e Infraestrutura, 13 - Ação Contra a Mudança do Clima e 7- Energia Limpa e Acessível (ISSA; DELBRÜCK; HAMM, 2019).

A União Europeia lançou a sua estratégia de bioeconomia em 2012 (*Innovating for Sustainable Growth: A Bioeconomy for Europe*), na qual entende por bioeconomia a “produção de recursos biológicos e a conversão desses recursos e fluxos de resíduos em produtos de valor agregado, como alimentos, rações, produtos de base biológica e bioenergia” (EUROPEAN UNION, 2012, p. 9, tradução autora). Nos anos seguintes alguns países europeus como a Alemanha, Finlândia e países nórdicos ocidentais, desenvolveram e publicaram suas estratégias de política nacional de bioeconomia (ALLEN *et al.*, 2015).

Também no ano de 2012, os Estados Unidos da América (EUA) publicaram o *National Bioeconomy Blueprint*, documento que estabeleceu os objetivos estratégicos do governo para a bioeconomia, que é definida como “aquela baseada no uso de pesquisas e inovação nas ciências biológicas para criar atividade econômica e benefício público” (WHITE HOUSE, 2012, p. 7, tradução autora). De acordo com o *Blueprint*, a bioeconomia oferece os meios para desenvolver produtos inovadores e que resultam em benefícios, como a melhoria da saúde humana, maior rendimento das safras de alimentos e a redução das emissões de carbono. Neste contexto, a biotecnologia é a característica-chave, vista como um impulsionador da bioeconomia nacional (FRISVOLD *et al.*, 2021).

Em termos de comercialização global, produtos da bioeconomia em 2014 representaram 13% do comércio internacional, com faturamento superior a US\$ 2.500 bilhões, dos quais US\$ 2.200 bilhões são das bioindústrias da União Europeia e US\$ 370 bilhões dos EUA (excluindo alimentos) (ISSA; DELBRÜCK; HAMM, 2019).

Na América do Sul, os princípios da bioeconomia também estão sendo adotados e existem iniciativas em vários países, nos quais o maior enfoque é dado para o uso da biodiversidade como fonte de medicamentos e alimentos funcionais pela aplicação da biotecnologia e a produção de biocombustíveis como o etanol (cana-de-açúcar) (SASSON; MALPICA, 2018).

No cenário mundial, o Brasil se destaca por ser megadiverso, seus recursos genéticos (vegetais, animais e microrganismos) e ecossistemas marinhos contribuem para o suprimento de alimentos, energia, biomateriais, moléculas e princípios ativos de interesse econômico (MMA, 2021). O país figura entre os líderes mundiais no agronegócio e na produção de biocombustíveis (OECD-FAO, 2020). Tais características conferem vantagens competitivas em comparação aos demais países e favorece a adoção de estratégias nacionais de bioeconomia.

A política nacional de bioeconomia foi publicada em 2016 pelo Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), entre as estratégias estão a elaboração de Planos de Ação de Ciência, Tecnologia e Inovação em biomas e em bioeconomia. Ainda de acordo com a estratégia nacional, é crucial a promoção de Ciência, Tecnologia e Inovação (CT&I) para conhecer, conservar e recuperar os ecossistemas e seus serviços ambientais. Portanto, as políticas públicas em CT&I para biomas e bioeconomia devem integrar a conservação, consumo e produção visando promover o desenvolvimento sustentável do país (MCTIC, 2016).

Entre os recursos genéticos, as espécies arbóreas são valiosos ativos para a bioeconomia brasileira, pela capacidade de fornecer produtos florestais madeireiros e não-madeireiros e serviços ambientais (MORI; LERTZMAN; GUSTAFSSON, 2017). Entretanto, alguns desafios precisam ser superados, por exemplo, o fato desses valiosos ativos serem pouco conhecidos, negligenciados e subutilizados. Nesse sentido, o governo brasileiro deu um passo inicial por meio da iniciativa “Plantas para o Futuro” do Ministério do Meio Ambiente (MMA), que busca identificar as espécies da flora brasileira de uso atual ou potencial, ampliar o conhecimento, despertar a preocupação sobre a conservação e a promoção do uso (CORADIN; SIMINSKI; REIS, 2011).

Como resultados dessa iniciativa, é possível oferecer informações claras às diferentes esferas do governo e da sociedade por meio de publicações contendo portfólios das espécies prioritárias para o desenvolvimento de pesquisas e uso para cada região do Brasil

(CORADIN; CAMILLO; PAREYN, 2018; CORADIN; SIMINSKI; REIS, 2011; VIEIRA; CAMILLO; CORADIN, 2016).

Entre as espécies florestais, a mangaba (*Hancornia speciosa* Gomes) foi inserida na categoria de Espécies Alimentícias em duas das listas publicadas, na região Nordeste (CORADIN; CAMILLO; PAREYN, 2018) e região Centro-Oeste (VIEIRA; CAMILLO; CORADIN, 2016).

2.1.1 *Hancornia speciosa* Gomes

H. speciosa (Apocynaceae), conhecida popularmente por mangabeira, é uma espécie arbórea, frutífera e medicinal nativa da América do Sul. Suas populações naturais estão amplamente distribuídas na região biogeográfica neotropical, especificamente do Nordeste ao Centro-Oeste do Brasil, Paraguai, Bolívia e Peru (TROPICOS, 2021).

Com base nas características morfológicas, esta espécie foi botanicamente caracterizada em seis variedades por Monachino (1945): *H. speciosa* var. *speciosa* Gomes - ocorre do Nordeste ao Norte e Centro-Oeste; *H. speciosa* var. *maximiliani* A. DC. e *H. speciosa* var. *lundii* A. DC. - Sudeste; *H. speciosa* var. *cuyabensis* Malme - Centro-Oeste (Mato Grosso); *H. speciosa* var. *gardneri* Mull. Arg. e *H. speciosa* var. *pubescens* (Nees & Mart.) Müll. Arg. - região central do Brasil. De acordo com a Flora do Brasil (2021), apenas duas variedades são oficialmente reconhecidas: *Hancornia speciosa* var. *speciosa* Gomes e *Hancornia speciosa* var. *pubescens* (Nees & Mart.) Müll.Arg.

Árvores de mangabeira podem chegar a até 15 metros, o tronco pode ser tortuoso ou reto com casca fendida, a copa é formada por ramos numerosos, separados e bem formados. Nas árvores da região Nordeste os ramos são mais finos e pendentes (Figura 2) (LEDERMAN *et al.*, 2000).



Figura 2. *Hancornia speciosa* em área natural no município de Coxim, Mato Grosso do Sul (Centro-Oeste) - A e no município de Pirambu, Sergipe (Nordeste) - B. Fonte: Autoria própria (2021).

As folhas são simples, alternas e opostas, com forma (elípticas, oblongo ou elíptico-lanceoladas nas duas extremidades) e tamanho variados com ausência (região Nordeste) ou presença de pilosidade (em algumas árvores da região Centro-Oeste) (Figura 3A e D). As flores são hermafroditas, brancas, em forma de campânula alongada (tubular), a inflorescência composta por duas a quatro flores por ápice e ocasionalmente flores isoladas (Figura 3B e E).

Os frutos são do tipo baga, de tamanho variado, formato elipsoide ou arredondado e exocarpo amarelo como pigmentação vermelha ou esverdeado (Figura 3C e F) (MONACHINO, 1945).



Figura 3. Folhas, flor e fruto de *Hancornia speciosa* em área natural no município de Coxim, Mato Grosso do Sul (Centro-Oeste) - A, B e C e no município de Pirambu, Sergipe (Nordeste) - D, E e F. Fonte: Autoria própria (2021).

Todas as partes da planta produzem látex (Figura 4), que foi objeto de pesquisas para a produção de borracha durante a Segunda Guerra Mundial, entretanto, após esse período, o reduzido valor econômico dessa borracha não estimulou a continuidade de pesquisas sobre suas características tecnológicas. Além disso, *H. speciosa* possui sistema laticífero com vasos inarticulado, necessitando portando de maior tempo para a restauração do látex, o que impede a exploração contínua da árvore (PINHEIRO et al., 2004). Atualmente, o látex é extraído por comunidades extrativistas e utilizado para fins medicinais. E as pesquisas realizadas possuem o objetivo de investigar suas propriedades farmacológicas.



Figura 4. Látex de *Hancornia speciosa*. Fonte: Plants of the World Online (2021).

H. speciosa é uma planta de clima tropical, com temperatura ideal de desenvolvimento entre 24-26 °C e a pluviosidade ideal pode estar entre 750 mm e 1.600 mm anuais. No entanto, pode ser encontrada em zonas com temperaturas e pluviosidades mínimas e máximas bem variadas. Ocorre naturalmente em áreas com solos arenosos ou areno-argilosos, ácidos, pobres em nutrientes e matéria orgânica, com baixa retenção de água, acidentados e pedregosos, ou seja, solos indesejados para fins agrícolas (PEREIRA *et al.*, 2016; SILVA JÚNIOR *et al.*, 2018).

A propagação de *H. speciosa* ocorre principalmente por sementes que apresentam comportamento fisiológico recalcitrante, perdendo a viabilidade em um curto espaço de tempo. A conservação da espécie tem ocorrido por meio de plantas vivas em áreas de ocorrência natural ou em bancos ativos de germoplasma mantidos por instituições de pesquisa, a exemplo, Universidade Federal de Goiás (57 acessos); Embrapa Amapá (86 acessos); Empresa Estadual de Pesquisa Agropecuária da Paraíba (540 acessos) e Embrapa Tabuleiros Costeiros (271 acessos) (SILVA JÚNIOR *et al.*, 2018).

H. speciosa é uma planta alógama e autoincompatível, para que ocorra a fecundação cruzada e a produção de frutos é necessário genótipos diferentes da espécie e polinizadores específicos (DARRAULT; SCHLINDWEIN, 2005). Suas populações naturais no bioma Cerrado e Mata Atlântica apresentam duas safras anuais. No bioma Cerrado, em populações localizadas no estado de Goiás, ocorre a perda de suas folhas no período seco, produção de flores e frutos durante as chuvas e brotação no período chuvoso, com produção de novembro e maio. Nesse intervalo de tempo, no mês de dezembro, os frutos maduros são encontrados no chão, os chamados frutos de caída (ALMEIDA *et al.*, 2018).

Na região Nordeste, em áreas de extrativismo no bioma Cerrado, nas fitofisionomias Cerrado *sensu stricto*, Cerradão e Mata Úmida, ocorre a presença de flores e frutos ao longo do ano, embora em quantidades muito baixas. Os picos das fenofases reprodutivas ocorrem durante os períodos de menor pluviosidade, e as fenofases vegetativas também ocorrem ao longo do ano de forma praticamente contínua (CAMPOS *et al.*, 2018).

No litoral nordestino, em remanescentes da Mata Atlântica na fitofisionomia de restinga, a safra de verão (janeiro a abril ou dezembro a abril - variação em função do estado) gera maior lucro, uma vez que a produção é gradativa e elevada, obedecendo à sazonalidade da fruta, que é de clima quente. A safra de inverno (maio a julho ou julho a outubro – variação em função do estado) tem menor produção e duração pela intensificação das chuvas nesse período, o que impede a maturação completa dos frutos (OLIVEIRA; ALOUFA, 2020; SANTOS; SOUZA, 2015).

O fruto é o principal produto florestal não madeireiro (PFNM) utilizado. É extraído por comunidades tradicionais que o consomem e o comercializam *in natura* e seus derivados, como doces, polpas e licores. Embora sejam muito apreciados pelas características sensoriais e pelo potencial como alimento nutricional e funcional, a espécie é subutilizada no país e a produção é oriunda do extrativismo em áreas naturais sem plano de manejo (VILAS BOAS *et al.*, 2021).

De acordo com dados oficiais, obtidos na Produção da Extração Vegetal e da Silvicultura (PEVS-2019) do Instituto Brasileiro de Geografia e Estatística (IBGE), a região Nordeste é responsável por 93,25% da produção nacional de mangaba, seguida pela região Sudeste responsável por 6,51% e Centro-Oeste 0,17% (IBGE, 2019). Para estas regiões existe uma portaria do Ministério da Agricultura, Pecuária e Abastecimento, nº 313/2019, que regulamenta os preços mínimos a serem praticados na comercialização dos frutos, que são R\$ 2,68/kg na região Nordeste e R\$ 1,68/kg para as regiões Sudeste e Centro-Oeste (BRASIL, 2019).

Entre os estados produtores de mangaba, a Paraíba tem a maior produção (735 t), seguido por Sergipe (381 t) e Bahia (232 t). Os estados de Pernambuco (6 t), Maranhão (3 t) e Goiás (3 t) são os com menor produção (Figura 5).

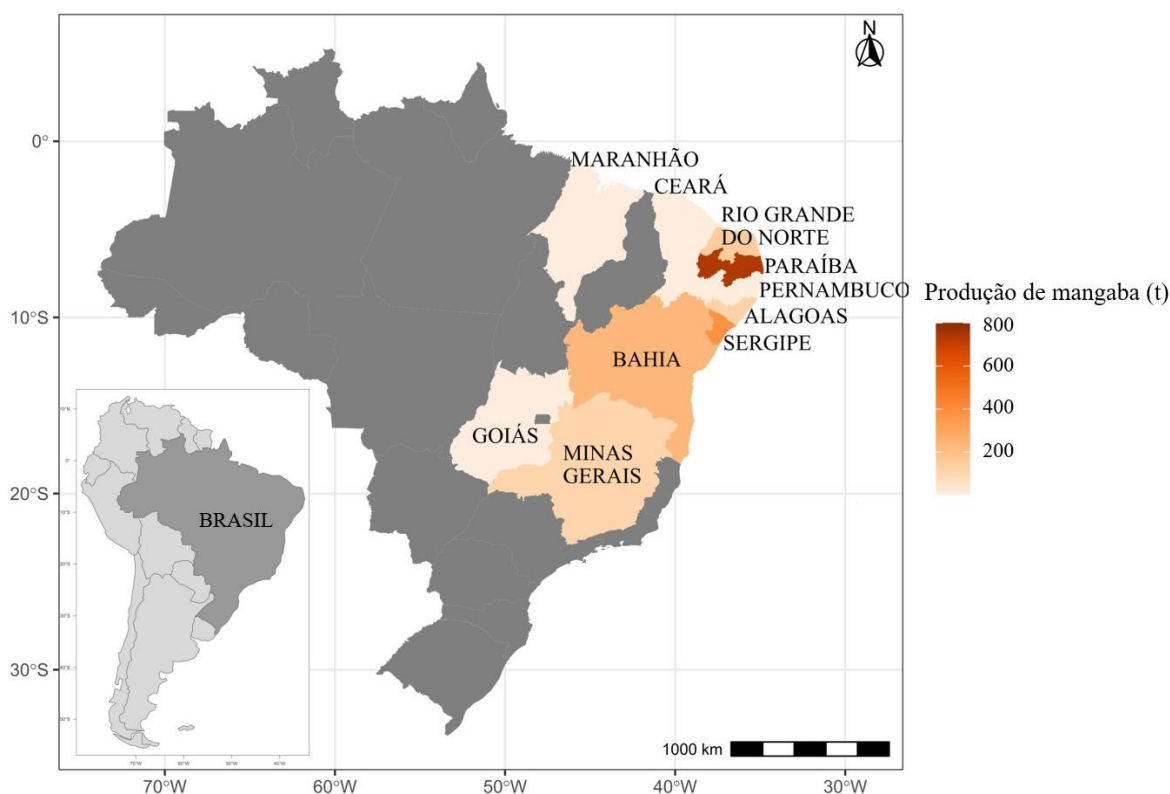


Figura 5. Espacialização da produção de mangaba no Brasil de acordo com a Produção da Extração Vegetal e da Silvicultura (PEVS-2019). Fonte: Elaborado a partir da PEVS-2019.

H. speciosa é uma planta que pode ser aproveitada por inteiro pois, além dos frutos como alimentos, comunidades extrativistas fazem uso do tronco para lenha, e raízes, cascas, folhas e látex como recurso medicinal.

A partir de levantamentos etnofarmacológicos, vários estudos foram desenvolvidos e comprovam a ação farmacológica das folhas (TOMAZI *et al.*, 2021) e látex, os quais são biocompatíveis com sistemas de vida e combinam propriedades angiogênicas, anti-inflamatórias e antibacterianas, portanto, pode ser um biomaterial para aplicações médicas (ALMEIDA *et al.*, 2014, 2019; BONETE *et al.*, 2020).

As sementes que desempenham um importante papel na propagação e sobrevivência da espécie, também são produtos promissores para o uso na produção de bio-óleo (SANTOS *et al.*, 2015). Conforme demonstrado por pesquisadores da Universidade Federal de Sergipe, sementes combinadas com resíduos de cana-de-açúcar possibilitaram o desenvolvimento de um bioplástico (PORTAL UFS, 2021).

H. speciosa é uma planta que possibilita diferentes usos, motivo que a torna um ativo para a bioeconomia brasileira. Garantir sua conservação e utilização de forma sustentável é imprescindível. Para tanto, é necessário o conhecimento da distribuição e extensão da diversidade genética disponível nas populações naturais que, entre outros fatores, é influenciada pelos níveis de evolução, sistema de reprodução, fatores ecológicos, geográficos e humanos (RAO; HODGKIN, 2002).

2.2 Diversidade em recursos genéticos vegetais

Os recursos genéticos vegetais (RGV) podem ser definidos como o material genético das plantas que possui uso potencial ou futuro, são estoques importantes de diversidade genética e constituem a matéria-prima básica para qualquer programa de melhoramento vegetal. Podem estar sob a forma de sementes, propágulos vegetativos (tubérculos, rizomas, entre outros), pólen, células, DNA e qualquer outro componente que contenha unidades

funcionais de hereditariedade (MIEDANER; KORZUN; BAUER, 2019; MURRAY, 2017; PATRA *et al.*, 2016).

Os RGV são a base da segurança alimentar, são também fonte de energia, ração animal, fibras e outros serviços ecossistêmicos (WAMBUGU; NDJIONDJOP; HENRY, 2018). Portanto, são recursos-chave para atender os Objetivos de Desenvolvimento Sustentável (ODS) das Nações Unidas, especificamente para: ODS 1 - Erradicação da pobreza; ODS 2 - Fome zero e Agricultura Sustentável; e ODS 13 - Ação Contra a Mudança Global do Clima.

Por toda a sua importância, a conservação e uso sustentável desses recursos são extremamente importantes e nunca foram tão urgentes, especificamente pelas pressões das atividades humanas que levam à degradação e erosão genética. Por exemplo, para florestas, de acordo como a Avaliação de Recursos Florestais Globais (FRA/FAO), entre 1990 e 2020 estima-se que 420 milhões de hectares foram perdidos pelo desmatamento (FAO, 2020).

Os RGV podem ser conservados basicamente por meio de duas estratégias, a conservação *in situ*, que consiste na manutenção da população dentro do seu ambiente de origem, de modo que os processos evolutivos continuem ocorrendo; e a conservação *ex situ*, na qual uma amostra dessa população é conservada fora do seu local de origem. Essa conservação pode ocorrer por meio de coleções de plantas vivas, bancos de sementes ou tecidos, desde que seja possível capturar e manter a variação genética da população (ARAVANOPOULOS, 2016).

A escolha de indivíduos ou populações para conservação, entre outros aspectos, deve considerar a diversidade genética, que pode ser entendida como a quantidade total de diferenças genéticas dentro das espécies, populações (entre e dentro das populações) e indivíduos (expressa por diferenças entre alelos nos cromossomos) (GRAUDAL *et al.*, 2014).

A diversidade genética é necessária para garantir a sobrevivência, adaptação e evolução das espécies sob condições de mudanças ambientais, bem como fornecer resiliência ao ataque de pragas e doenças. Além disso, é importante para a seleção artificial, reprodução e domesticação em programas de melhoramento genético (ARAVANOPOULOS, 2016; GOVINDARAJ; VETRIVENTHAN; SRINIVASAN, 2015).

Forças evolutivas como a seleção natural, mutação, migração e deriva genética resultam em modificações nas frequências alélicas e afetam a diversidade genética das populações. Do mesmo modo, a domesticação ou seleção artificial exerce efeitos sobre a diversidade genética, reduzindo-a quando comparada à presente na natureza por favorecer alguns alelos em detrimento de outros, aumentando a frequência dos alelos selecionados (BHANDARI *et al.*, 2017).

A diversidade genética presente nas populações de plantas pode ser acessada por meio de técnicas como os marcadores morfológicos, bioquímicos e de DNA (ácido desoxirribonucleico). Os marcadores morfológicos são baseados em características fenotípicas, a exemplo, cor da flor, formato da semente e pigmentação. Estes marcadores normalmente requerem grandes extensões de terra para experimentos de campo e disponibilidade de mão de obra (GOVINDARAJ; VETRIVENTHAN; SRINIVASAN, 2015).

Os marcadores moleculares bioquímicos são baseados na eletroforese de enzimas, a reação pode expressar-se na forma de isoenzimas - formas alélicas variantes de uma mesma enzima codificadas por locos diferentes; ou por aloenzimas - representam a expressão fenotípica de alelos no mesmo loco (ALFENAS *et al.*, 1991; KUMAR *et al.*, 2018).

Além dos marcadores bioquímicos, tem-se os marcadores moleculares baseados no DNA, que são ferramentas da biotecnologia amplamente utilizadas em diferentes plantas. Consistem em um loco genômico, detectado por meio de sonda ou iniciadores específicos (*primers*), que por sua presença distingue a característica cromossômica que representa (BARCACCIA *et al.*, 2011). Não sofrem efeito do ambiente, podem ser aplicados a diferentes partes do genoma (íntrons, éxons e regiões de regulação), não são afetados por efeitos

pleiotrópicos ou epistáticos e são capazes de distinguir polimorfismos que não produzem variação fenotípica (MONDINI; NOORANI; PAGNOTTA, 2009).

Quanto à capacidade de distinção de heterozigotos e homozigotos, os marcadores moleculares baseados em DNA podem ser codominantes quando permitem a distinção clara entre homozigotos e heterozigotos ou dominantes, quando não é possível tal distinção. Estes marcadores possibilitam estimar parâmetros genéticos importantes, como por exemplo, o tamanho efetivo da população (N_e), coeficiente de endogamia (f) e heterozigosidade esperada (H_E). Sendo que, ao usar marcadores dominantes, a estimativa do H_e só pode ocorrer partindo-se da premissa de que a população estudada se encontra em Equilíbrio de Hardy-Weinberg (EHW), ou seja, todos os indivíduos possuem probabilidades iguais de sobrevivência e/ou reprodução e na ausência de seleção, mutação ou migração, as frequências alélicas e genotípicas são constantes de uma geração para a outra (HARTL; CLARK, 2010).

O N_e é um parâmetro genético importante que se refere ao número de indivíduos que se reproduzem e deixam descendentes, transmitindo os genes para a próxima geração. É fundamental para fins evolutivos e conservação genética, porque permite inferir sobre a deriva genética, endogamia e permite a estimativa e comparação com o tamanho mínimo viável ou adequado que assegura a manutenção da viabilidade das populações e sustentabilidade genética na sua descendência (CHARLESWORTH, 2009; VENCOVSKY; CROSSA, 2003).

A H_E também é conhecida como a diversidade de Nei (1973) e é utilizada com frequência para medir a diversidade em populações. A estimativa desse parâmetro varia de zero a um e quanto mais alta, maior é a diversidade e ocorre a minimização da coancestria média (parentesco) (ALLENDORF; LUIKART, 2007).

O f mede a relação entre as heterozigosidades observadas e esperadas, valores próximos a zero indicam cruzamento aleatório, valores positivos indicam endogamia e valores negativos excesso de heterozigotos (HARTL; CLARK, 2010).

Estes e outros parâmetros genéticos estimados por marcadores moleculares são importantes para conhecer e monitorar a diversidade genética em plantas e podem contribuir na proposição de estratégias de conservação e melhoramento.

2.3 Inter Simple Sequence Repeat (ISSR)

Os marcadores ISSR (*Inter Simple Sequence Repeat*) são marcadores dominantes e estão distribuídos aleatoriamente no genoma. Trata-se de segmentos de DNA flanqueados em ambas as extremidades por microssatélites - regiões do genoma que consistem em motivos de DNA curtos repetidos várias vezes em uma sequência (ZIETKIEWICZ; RAFALSKI; LABUDA, 1994).

O uso de marcadores ISSR possibilita que vários segmentos aleatórios de DNA no genoma sejam amplificados em uma única reação, possibilitando a geração de um grande número de locos em todo o genoma de qualquer espécie, além disso não requer informações sobre as sequências de DNA das regiões alvo e possuem alta reprodutibilidade (BORNET; BRANCHARD, 2001; ZIETKIEWICZ; RAFALSKI; LABUDA, 1994).

A técnica é simples e requer baixa quantidade de DNA ($10 - 50 \text{ ng}\mu\text{L}^{-1}$), os *primers* ISSR são desenvolvidos com sequências repetitivas complementares às regiões de microssatélites no genoma e geralmente possuem de 16-25 pares de bases (bp), os mais usados são aqueles desenvolvidos na *University of British Columbia* (UBC) (NG; TAN, 2015).

Em linhas gerais, a condução de um experimento empregando marcadores ISSR para genotipagem consiste em: i) extração, quantificação e normalização do DNA genômico; ii) amplificação por reação em cadeia da polimerase (*Polymerase Chain Reaction* - PCR) com *primers* ISSR; iii) eletroforese dos produtos amplificados por PCR em gel de agarose ou poliacrilamida; iv) geração da matriz binária a partir dos fragmentos visualizadas no gel; e v) análises dos dados.

Na vasta literatura sobre ISSR, esses marcadores foram utilizados para diferentes objetivos e em diferentes espécies para avaliar a diversidade genética em trigo (GOWAYED; ABD EL-MONEIM, 2021), estudar relações genéticas entre espécies como as dos gêneros *Mentha*, *Malva* e *Lilium* (JEDRZEJCZYK; REWERS, 2018, 2020; ZHAO *et al.*, 2014), para confirmar a homogeneidade genética entre plântulas micropropagadas *in vitro* para cultivares de uva (NOOKARAJU; AGRAWAL, 2012) e em espécies ameaçadas como a orquídea *Dendrobium moschatum* (THAKUR *et al.*, 2016) e *Pittosporum eriocarpum* (TIKENDRA; AMOM; NONGDAM, 2019).

Por suas características, o uso de marcadores ISSR para conhecer a diversidade genética presente em recursos genéticos vegetais (RGV) é uma excelente opção, principalmente em estudos preliminares ou com espécies para as quais não existem informações disponíveis sobre o genoma. Em vários países (Figura 6A) são desenvolvidos estudos com ISSR em RGV desde a sua descoberta até os dias atuais e as principais áreas de aplicação são Biologia, Botânica, Diversidade Genética, Variação Genética e População (Figura 6B).

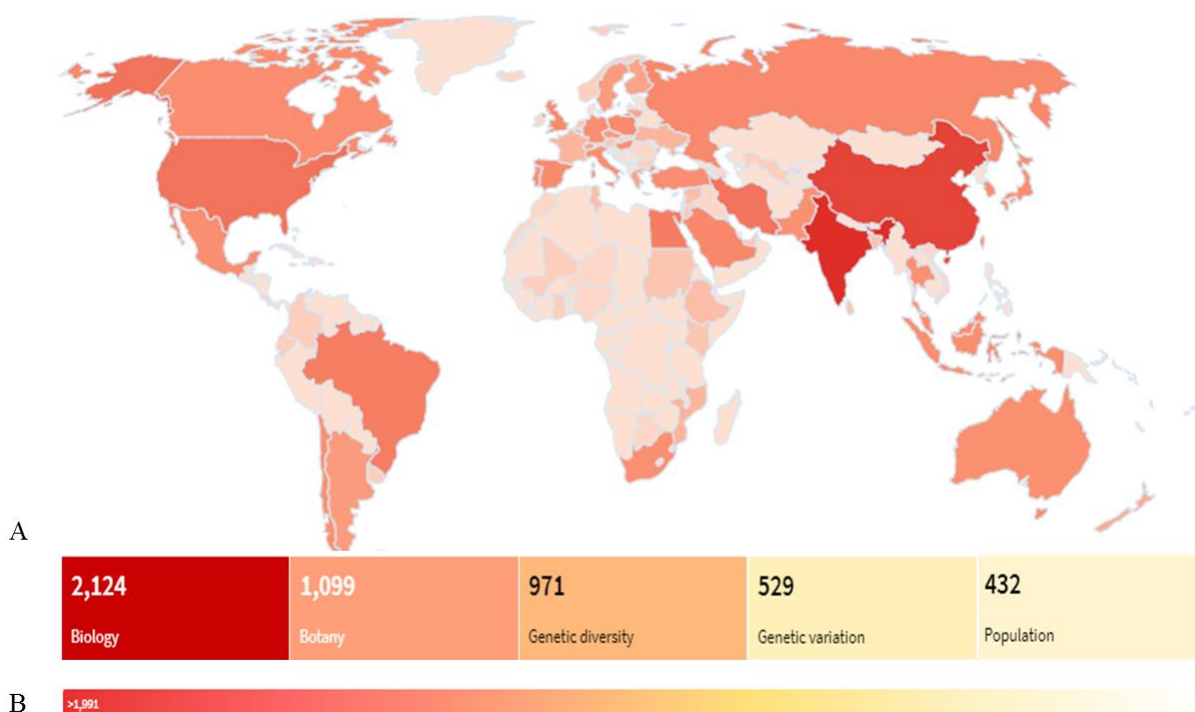


Figura 6. Países de origem (A) e principais áreas (B) de estudos com recursos genéticos vegetais utilizando marcadores ISSR. Fonte: Elaborado a partir da plataforma Lens.org (2021)¹.

Entre as várias aplicações em RGV, os marcadores ISSR podem gerar impressões digitais de DNA (*DNA fingerprint*) que permitem a identificação precisa de genótipos, informação importante no melhoramento de plantas e necessária para proteger a propriedade intelectual dos criadores de novas variedades, bem como garantir pureza para os consumidores (BORNET; BRANCHARD, 2012; HUANG *et al.*, 2014).

Para espécies do gênero *Thymus* spp. o uso de marcadores ISSR resultou na obtenção de polimorfismo suficiente e perfis reproduzíveis de DNA para detecção genética da variação e relações entre quatorze acessos (YOUSEFI *et al.*, 2015); em *Panax stipuleanatus* Tsai

¹ Opções de busca – Termo: (issr AND plant); Campos de pesquisa: título, palavras-chave, resumo e área de estudo; Tipo de publicação: artigos de periódicos.

(planta medicinal) para estimar a diversidade genética em populações selvagens (TRIEU *et al.*, 2016); em variedades de *Plectranthus* (cultivares ornamentais) protegidas por patente, foram utilizados para determinar a comercialização ilegal de cultivares (RICE *et al.*, 2010).

Marcadores ISSR também podem ser aplicados para análise filogenética de RGV. Em espécies do gênero *Myrcia* (*M. laruotteana*, *M. selloi*, *M. lajeana* e *M. tomentosa*) esses marcadores forneceram informações para uma nova classificação taxonômica das espécies (LIMA *et al.*, 2015). ISSR foram usados também para caracterizar os genomas de cultivares de cereais economicamente importantes (cevada, milho, arroz, sorgo – branco e vermelho e trigo) e abordar suas relações filogenéticas (RABEY; ALSHUBAILY; AL-OTAIBI, 2015).

Marcadores ISSR têm sido utilizados para estimar a diversidade genética de espécies que se tem pouco conhecimento sobre essa informação, como *Myrcia ovata* Cambess, planta medicinal e aromática (WHITE *et al.*, 2018). É utilizado também em espécies que estão em processo e domesticação ou inseridas em programas de melhoramento, como *Jatropha curcas* L. em que marcadores ISSR foram utilizados para estimar a variabilidade genética de dez híbridos obtidos por cruzamentos dialélicos (RIBEIRO *et al.*, 2017).

Esses marcadores são também ferramentas valiosas para a criação e gestão dos bancos ativos de germoplasma, visto que podem ser úteis para caracterizar a diversidade genética dos acessos (SOUZA *et al.*, 2015), avaliar a necessidade de ampliar o número de acessos (SAIFERT *et al.*, 2020) e identificar duplicatas (DAS *et al.*, 2020).

Em *H. speciosa*, marcadores ISSR foram empregados para avaliar a diversidade genética de acessos em bancos ativos de germoplasma (SANTOS *et al.*, 2017), em população natural (SOARES *et al.*, 2017; NUNES *et al.*, 2021) e entre populações naturais (FAJARDO *et al.*, 2018) com sucesso.

2.4 Single Nucleotide Polymorphism (SNP)

Genomas completos foram sequenciados, incluindo o de humanos e de plantas, e esses avanços em tecnologia possibilitaram novas descobertas, entre elas, polimorfismos de nucleotídeo único (SNPs), alteração de um único nucleotídeo (A, T, C ou G), ou seja, a mutação de um nucleotídeo na sequência do genoma de indivíduos de uma população (RAFALSKI, 2002).

SNPs são categorizados com base na substituição de nucleotídeos como transições, que são trocas entre purinas (A ↔ G) ou entre pirimidinas (C ↔ T). Com menor frequência como transversões, quando há troca de uma purina por uma pirimidina ou vice-versa (C/T ↔ A/G) (UDOH; OBASEOJEI; UZOEBU, 2021).

Uma vez que o DNA é composto por quatro bases nitrogenadas (A, C, T e G), em teoria é possível a existência de quatro alelos diferentes para cada nucleotídeo em um loco SNP. Porém, o que tem sido observado com maior frequência são apenas dois estados alternativos, por esse motivo os SNPs são considerados marcadores bialélicos e estima-se que sua taxa de mutação seja da ordem de 10^{-8} a 10^{-9} (FREELAND, 2005; SYVANEN, 2001).

Os SNPs são marcadores codominantes, portanto, permitem identificar no genótipo heterozigoto os dois alelos. São considerados menos informativos do ponto de vista genético quando comparados a marcadores microsatélites por serem bialélicos. Entretanto, esse fato é compensado porque são abundantes no genoma. SNPs e podem ocorrer em regiões codificadoras (éxons) ou não-codificadoras (íntrons e regiões intergênicas), geralmente são encontrados em espaços intergênicos sem função determinada (FREELAND, 2005).

Esses marcadores podem ser utilizados em diversas áreas de estudo, tais como estudos evolutivos, filogenéticos, ecológicos, no melhoramento e no mapeamento genético. Especificamente em RGV, estudos foram desenvolvidos em vários países e as principais áreas de aplicação são Biologia, Genética, População, Gene e para identificar loco de traço quantitativo (Figura 7).

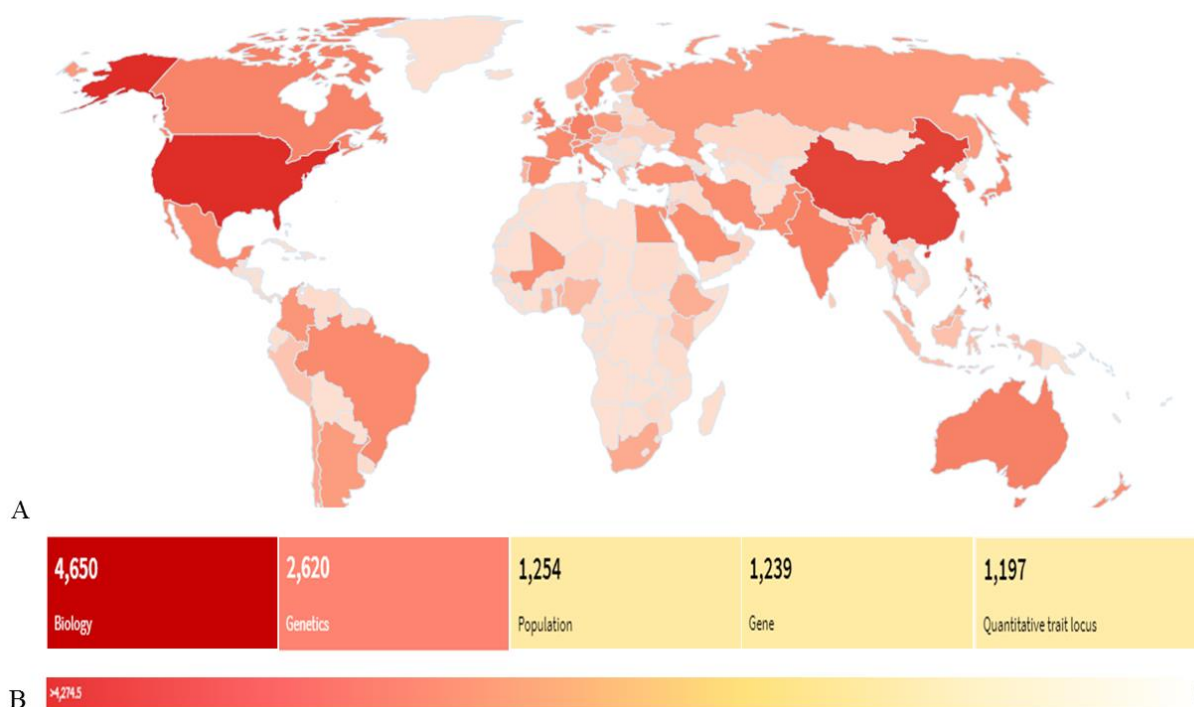


Figura 7. Países (A) e principais áreas (B) de estudos com recursos genéticos vegetais utilizando marcadores SNP. Fonte: Elaborado a partir da plataforma Lens.org (2021)².

Técnicas e métodos diferentes foram adotados para a descoberta e genotipagem de SNP no genoma das espécies de interesse. Para a identificação são exemplos de metodologias o sequenciamento de produtos da reação em cadeia da polimerase (PCR) e a identificação eletrônica de SNPs utilizando como base, por exemplo, bibliotecas disponíveis para a espécie em estudo como EST (*Expressed Sequence Tags*) ou bibliotecas genômicas (VAN OEVEREN; JANSSEN, 2009).

A genotipagem pode ocorrer empregando métodos baseados em hibridização, como a hibridização alelo específica, de sondas e em arranjos (SNP *array*); por métodos de ligação de oligonucleotídeo baseado em PCR; e mais recentes métodos baseados em sequenciamento de nova geração (*Next Generation Sequencing*, NGS) (GILAD; PRITCHARD; THORNTON, 2009; SYVANEN, 2001; TORKAMANEH; LAROCHE; BELZILE, 2016).

O NGS oferece várias abordagens que são capazes de realizar simultaneamente a descoberta de SNPs de todo o genoma e genotipagem em uma única etapa. Além disso, possibilita um grande número (milhares a milhões) de SNPs a um menor custo e pode ser aplicado mesmo em espécies com pouca ou nenhuma informação genética disponível em bancos públicos, visto que não é necessário definir previamente quais são os SNPs que serão genotipados ao longo do genoma (DAVEY *et al.*, 2011; TURCHETTO-ZOLET *et al.*, 2017).

Foram desenvolvidas várias plataformas NGS, como Roche 454 FLX Titanium (THUDI *et al.*, 2012), Illumina MiSeq e HiSeq2500 (BENTLEY *et al.*, 2008), Ion Torrent PGM (ROTHBERG *et al.*, 2011), que possibilitam novas aplicações. Normalmente os protocolos empregam enzimas de restrição específicas para reduzir a complexidade genômica do organismo de interesse; oligonucleotídeos de regiões conhecidas para ligar nas regiões de interesse ou o sequenciamento de genes candidatos para amplificar as regiões de interesse (TURCHETTO-ZOLET *et al.*, 2017).

Para a redução da complexidade genômica pode ser adotado o sequenciamento de DNA associado à restrição (*Restriction-site associated DNA sequencing* - RAD-seq), que

² Opções de busca – Termo: (snp AND plant); Campos de pesquisa: título, palavras-chave, resumo e área de estudo; Tipo de publicação: artigos de periódicos.

permite a seleção de um determinado número de locos RAD com base na escolha de uma enzima de restrição (BAIRD *et al.*, 2008). Para a identificação de marcadores SNPs, o DNA é digerido com uma enzima de restrição, ligados a adaptadores e submetidos à fragmentação aleatória (random shearing), e por fim sequenciados (DAVEY *et al.*, 2011). Outra abordagem é a genotipagem por sequenciamento (*Genotyping-by-Sequencing* - GBS), que é tecnicamente considerada simples, altamente multiplexada e que permite a construção de bibliotecas de representação reduzida (uma fração direcionada do genoma) usando enzimas de restrição, sondas de captura ou sequenciando o transcriptoma (DAVEY *et al.*, 2011; DESCHAMPS; LLACA; MAY, 2012).

A GBS possibilita simultaneamente descobrir polimorfismos e obter informações genóticas em toda a população de interesse. Assim, gera um grande número de SNPs para uso em análises genéticas, que é adequada para estudos populacionais, caracterização de germoplasma, reprodução e mapeamento de características em diversos organismos. É uma abordagem rápida e flexível que pode ser aplicada a várias espécies, conjuntos de germoplasma e em programas de melhoramento de plantas. Sobretudo é importante para novos conjuntos de germoplasma e espécies não caracterizadas (ELSHIRE *et al.*, 2011; NARUM *et al.*, 2013; POLAND; RIFE, 2012).

A abordagem de GBS inicial, assim como o RAD-seq utilizava uma única enzima de restrição para reduzir a complexidade do genoma, isso porque o uso desse tipo de enzima é fácil, rápido, extremamente específico, altamente reprodutível e pode atingir regiões importantes do genoma que são inacessíveis para abordagens de captura de sequência (ELSHIRE *et al.*, 2011). Posteriormente, foi desenvolvido um protocolo de GBS empregando duas enzimas de restrição e indicado para espécies que não dispõem de sequências do genoma de referência (POLAND *et al.*, 2012). Para a técnica de RAD-seq também ocorreram atualizações metodológicas e uma delas é o *double digest* RAD (ddRAD) que utiliza duas enzimas de restrição para digestão do DNA e não há necessidade da fragmentação aleatória como no protocolo original (PETERSON *et al.*, 2012).

No protocolo com duas enzimas de restrição em GBS é aplicado uma de corte raro e uma segunda enzima de corte comum, produzindo uma digestão com: i) fragmentos com um local de corte raro e um local de corte comum; ou ii) fragmentos com dois locais de corte comuns. Em estudo com cevada e trigo foram utilizadas as enzimas *PstI* (CTGCAG) – corte raro e *MspI* (CCGG) – corte comum, sendo demonstrada a robustez da GBS em espécies com genomas grandes, complexos e até poliploides (POLAND *et al.*, 2012).

Em linhas gerais, as etapas dessa abordagem envolvem o isolamento do DNA genômico de todas as amostras, fragmentação com enzimas de restrição e ligação dos adaptadores; amplificação por PCR, sequenciamento, análise das sequências com ou sem suporte de um genoma de referência e identificação dos SNPs (DAVEY *et al.*, 2011; POLAND; RIFE, 2012).

Por todas as características mencionadas, a GBS, diferente de outras tecnologias de genotipagem de alta densidade, tornou-se uma alternativa para a descoberta e genotipagem de SNPs em espécies sem genoma de referência, como a maioria das árvores tropicais. Neste caso a GBS é vantajosa por permitir rastrear milhares de polimorfismos em todo o genoma que estão sujeitos a processos evolutivos (variação na deriva, seleção, recombinação, mutação) e consequentemente para a variação genética. Além disso, por meio desses estudos é possível ter uma visão mais completa da base genômica para a adaptação, o que pode levar a genes candidatos para estudos funcionais adicionais (HOLLIDAY *et al.*, 2017; LEBEDEV *et al.*, 2020; NARUM *et al.*, 2013).

Vários estudos foram desenvolvidos com populações florestais e nestes demonstra-se o uso de GBS para descobrir e genotipar SNPs utilizados em estudos de diversidade e estrutura genética (DÍAZ *et al.*, 2021; ZHU *et al.*, 2019), filogenéticos (GUAJARDO *et al.*, 2020) e no melhoramento para seleção genômica, por exemplo em *Hevea brasiliensis*

(seringueira) (SOUZA *et al.*, 2019) e *Castanea dentata* (castanheiro americano) (WESTBROOK *et al.*, 2020).

Tais marcadores podem contribuir para estudos de genômica populacional, porque permitem a identificação de locos neutros (afetados de forma similar pela história demográfica e evolutiva das populações) e a identificação de locos *outliers* que estão supostamente sob seleção e apresentam comportamento diferenciado com relação à variação genética (ALLENDORF; HOHENLOHE; LUIKART, 2010). Embora muito informativos, os marcadores SNPs não foram usados até o momento em estudos genéticos com *H. speciosa*, seja empregando GBS ou outra tecnologia.

2.5 Modelos de distribuição de espécies (SDM)

Modelos de distribuição de espécies (SDM) são um conjunto de técnicas que permitem modelar os requisitos ambientais das espécies e o mapeamento de suas distribuições no espaço e no tempo, empregando métodos estatísticos e dados geoespaciais (ELITH; FRANKLIN, 2013). Portanto, são modelos empíricos desenvolvidos a partir da localização de espécies (abundância, ocorrência) e as variáveis ambientais que provavelmente influenciam a distribuição das espécies. SDM podem fornecer informações ecológicas e evolutivas e prever distribuições entre as paisagens (GUISAN; ZIMMERMANN, 2000).

À medida que foram surgindo estudos com SDM, diferentes terminologias foram adotadas, como mapeamento de vegetação preditiva (FRANKLIN, 1995), modelagem de distribuição de habitat preditiva (GUISAN; ZIMMERMANN, 2000), modelagem de adequação de habitat (HIRZEL *et al.*, 2006), modelagem de envelope bioclimático (PEARSON; DAWSON, 2003), mapeamento do índice de adequação de habitat (HAGHI VAYGHAN *et al.*, 2013) e modelos de nicho ecológico (*Ecological Niche Models* - ENMs) (VAZ; NABOUT, 2016).

Nos estudos mais recentes, são adotados SDM e ENM e os autores tratam como sinônimos (ZURELL *et al.*, 2020), enquanto outros defendem que ENM tem maior enfoque na estimativa de parâmetros de nichos ecológicos fundamentais e o SDM é mais focado em distribuições geográficas de espécies (FENG *et al.*, 2019; PETERSON; SOBERÓN, 2012), termo adotado no presente estudo.

O número de estudos empregando SDM para diferentes espécies aumentou consideravelmente, especificamente em plantas; considerando artigos publicados em periódicos disponíveis na base Lens.org, foram mais de seis mil publicações nas últimas duas décadas (Figura 8). E estão concentradas principalmente nas áreas de Ecologia e Biologia (Figura 9).

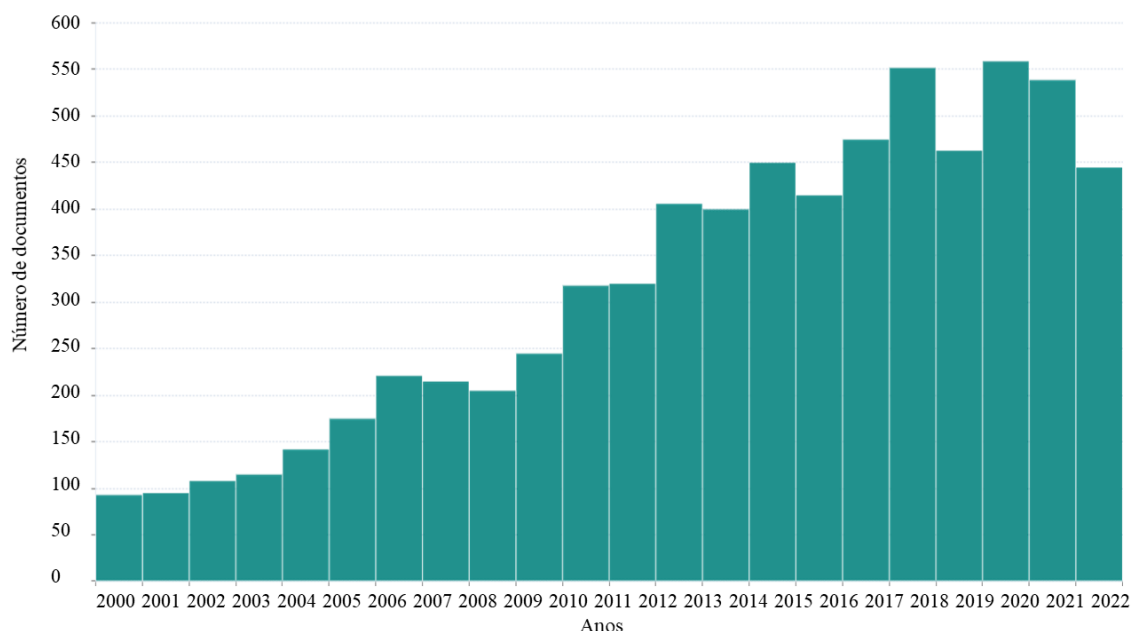


Figura 8. Estudos com plantas empregando modelos de distribuição de espécies (SDM). Fonte: Elaborado a partir da plataforma Lens.org (2021)³.

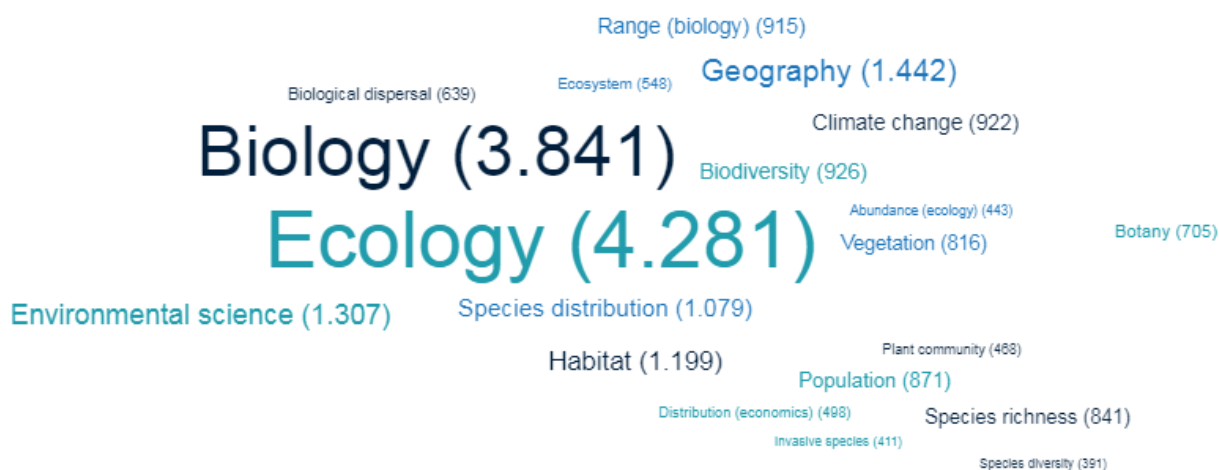


Figura 9. Nuvem de palavras com as principais áreas dos estudos com plantas que empregam modelos de distribuição de espécies (SDM). Fonte: Elaborado a partir da plataforma Lens.org (2021)⁴.

SDM tem uma ampla gama de aplicações incluindo a avaliação dos impactos das mudanças climáticas na biodiversidade (KOLANOWSKA; JAKUBSKA-BUSSE, 2020; PORFIRIO *et al.*, 2014), identificação de áreas prioritárias para conservação (LIU *et al.*, 2013; MORADI *et al.*, 2019), avaliação dos efeitos do microclima, luz e condições do solo na distribuição de espécies (GREISER *et al.*, 2020), gestão de espécies invasoras (CHADIN *et al.*, 2017; LAMSAL *et al.*, 2018), distribuição potencial de patógenos (MARTÍNEZ-MINAYA *et al.*, 2018), localização de novas populações de espécies pouco conhecidas e de interesse de conservação (SOUSA-SILVA *et al.*, 2014; WILLIAMS *et al.*, 2009), entre outros.

³ Opções de busca – Termo: (species distribution models AND plant); Campos de pesquisa: título, palavras-chave, resumo e área de estudo; Tipo de publicação: artigos de periódicos.

⁴ Opções de busca – Termo: (species distribution models AND plant); Campos de pesquisa: título, palavras-chave, resumo e área de estudo; Tipo de publicação: artigos de periódicos.

Para padronizar a construção dos modelos quanto à transparência e reprodutibilidade, alguns estudos foram desenvolvidos sugerindo protocolos (GRIMM *et al.*, 2006, 2010; ZURELL *et al.*, 2020) e uma lista de verificação (FENG *et al.*, 2019) com os elementos essenciais necessários para garantir a reprodutibilidade do SDM, e mais recente foi desenvolvida uma ferramenta (<https://odmap.wsl.ch/>) que possibilita a documentação de dados e cataloga vários aspectos do processo de modelagem (FITZPATRICK; LACHMUTH; HAYDT, 2021; ZURELL *et al.*, 2020). Assim, a elaboração de SDM envolve cinco etapas principais: Visão Geral, Dados, Modelo, Avaliação e Predição (*Overview, Data, Model, Assessment, Prediction*, ODMAP).

A etapa de visão geral consiste em um resumo contendo as hipóteses e pressupostos do modelo, algoritmos e software utilizado; etapa de dados consiste na descrição detalhada dos dados (ocorrência e condições ambientais) utilizados e do seu processamento; etapa modelo consiste em informações que possibilitem repetir a construção do modelo (colinearidade das variáveis; configurações de modelo e análise de não independência de dados); etapa avaliação consiste na aplicação de estatísticas de desempenho e verificações de plausibilidade em relação ao conhecimento biológico disponível, e etapa predição avalia a probabilidades de ocorrência contínua ou presença potencial e quantificação das incertezas no modelo predito (ZURELL *et al.*, 2020).

Os dados de ocorrência das espécies e dados ambientais utilizados para gerar SDM podem ser oriundos de pesquisas de campo próprias e/ou serem acessadas em bancos de dados públicos (Tabela 1), como o *Global Biodiversity Information Facility* (GBIF), que é um repositório global com dados de ocorrências para várias espécies (GBIF, 2021); o *WorldClim* - banco global de dados climáticos e meteorológicos com alta resolução espacial (WORLDCLIM, 2020). Os softwares e algoritmos utilizados são vários e podem ser escolhidos considerando os objetivos do modelo e conhecimento do pesquisador.

Tabela 1. Exemplos de bases de dados com informações sobre a ocorrência de espécies, variáveis ambientais e principais softwares e algoritmos utilizados para gerar os modelos de distribuição de espécies.

	Base de dados / Software	Descrição / Algoritmos	Site
Ocorrência	<i>Global Biodiversity Information Facility</i> - GBIF	Espécies e grupos de plantas, animais e microrganismos	http://data.gbif.org/welcome.htm
	NatureServe	Plantas, animais e ecossistemas na América do Norte e do Sul	http://www.natureserve.org/
Ambiente	WorldClim	Superfícies climáticas interpoladas para o globo com resolução de 1 km	http://www.worldclim.org/
	<i>Earth Resources Observation and Science</i> (EROS) Center	Elevação e variáveis relacionadas para o globo com resolução de 1 km	http://eros.usgs.gov/#/FindData/ProductsandDataAvailable/ElevationProducts

Gerar modelos	PNUMA	Tipos de solo global	http://www.grid.unep.ch/data/data.php?category=lithosphere
	<i>MaxEnt</i>	Entropia Máxima	http://www.cs.princeton.edu/~schapire/maxent/
	R-Project (pacotes para geração e avaliação dos modelos)	Bioclim; Distância de Mahalanobis; Domain (Distância/ Domínio de Gower); <i>Support Vector Machine</i> ; e Entropia Máxima	http://www.r-project.org/
	DIVA-GIS	Envelope ambiental	http://www.diva-gis.org
	ArcGIS com SDMtoolbox (59 ferramentas GIS baseadas em <i>script Python</i> desenvolvidas e compiladas em uma única interface.	Entropia Máxima	http://www.sdmtoolbox.org

Em se tratando de algoritmos, estes podem ser baseados apenas em dados de presença, que permitem então o uso de dados para os quais não estão disponíveis ou são insuficientes as informações de ausências. Baseiam-se na definição de envelopes ambientais, espaço climático definido como base correlações estatísticas entre dados de ocorrência de espécies e dados ambientais espaciais em torno dos locais onde as espécies ocorrem (HIRZEL *et al.*, 2002; LAHOZ-MONFORT; GUILLERA-ARROITA; WINTLE, 2014). Outro grupo de algoritmos requerem dados presença-ausência para gerar funções estatísticas que permitem a geração dos modelos com as distribuições de presença e ausência de espécies (LAWSON *et al.*, 2014).

Entre os algoritmos usados com frequência estão Bioclim (NIX, 1986); Domain (Distância/Domínio de Gower) (CARPENTER; GILLISON; WINTER, 1993) e Distância de Mahalanobis (FARBER; KADMON, 2003), que são baseados apenas em dados de presença e os algoritmos de aprendizado de máquina, Entropia máxima (*MaxEnt*) (PHILLIPS; DUDÍK, 2008) baseado em dados de presença-fundo e Máquina de Vetores de Suporte (*Support Vector Machine* - SVM) (TAX; DUIN, 2004) baseado em dados de presença-ausência.

Bioclim (envelope ambiental) foi o primeiro método utilizado para gerar SDM, no qual o algoritmo identifica locais que possuem valores que se enquadram na faixa daqueles observados para o conjunto de dados de ocorrência. A pontuação mais baixa em valores ambientais para uma célula de grade é mapeada e pode ser nula (fora da faixa de valores observada), faixa zero (baixo) ou máxima (muito alto). O modelo é calculado desconsiderando 5% dos valores mais baixos e mais altos de cada índice climático para reduzir o impacto de *outliers* (BOOTH *et al.*, 2014; NIX, 1986).

Para modelos gerados com Domain (Distância/Domínio de Gower), o algoritmo calcula a estatística de distância de Gower (G), medida entre os valores das variáveis ambientais de cada célula e de cada ponto de ocorrência. Desse modo, a distância entre o ponto A e a célula da grade B para uma única variável climática k é calculada como a diferença absoluta entre A e B dividida pelo intervalo de k em todos os pontos. A estatística de similaridade do domínio é calculada como $100 \times (1-G)$ e a semelhança máxima entre uma célula da grade e qualquer ponto é mapeada (CARPENTER; GILLISON; WINTER, 1993).

A Distância de Mahalanobis classifica locais potenciais por sua distância de Mahalanobis a um vetor que expressa as condições ambientais médias de todos os registros no

espaço ambiental para as espécies-alvo. Este algoritmo se baseia na média multivariada e na matriz de covariância e produz um envelope elíptico que explica explicitamente as possíveis correlações entre as variáveis ambientais (FARBER; KADMON, 2003).

A Entropia máxima (*MaxEnt*) é amplamente conhecida e utilizada para gerar SDM, consiste na aplicação de uma técnica de aprendizado de máquina que usa os dados de presença-*background* das camadas ambientais no processo de modelagem. Ao encontrar a distribuição da entropia máxima, permite estimar a probabilidade de uma espécie estar presente (PHILLIPS; DUDÍK, 2008).

A máquina de vetores de suporte (SVM) consiste em um algoritmo de aprendizado supervisionado que classifica os dados em grupos e tem configurações que controlam a complexidade do ajuste. Mapeia amostras para um espaço de alta dimensão definido por variáveis preditoras e encontra o hiperplano que melhor separa as amostras. Assim, usando projeções de hiperplanos, a SVM obtém resultados de predição ao determinar quais amostras pertencem a determinada categoria (CHA *et al.*, 2021; TAX; DUIN, 2004).

A modelagem tem como saídas mais comuns: i) resultados binários - os locais são classificados como parte da distribuição da espécie (1) ou fora de sua distribuição (0); ii) resultados contínuos em que cada local recebe uma probabilidade de fazer parte da distribuição da espécie ou adequação para a espécie (LIU; WHITE; NEWELL, 2011). SDM pode ser avaliado quanto a capacidade de discriminação, ou seja, a capacidade de um modelo de distinguir entre locais onde a espécie-alvo foi detectada (locais de presença) e locais onde está ausente (locais de ausência) e a confiabilidade que se refere à concordância entre as probabilidades de ocorrência previstas e as proporções observadas de locais ocupados pela espécie (PEARCE; FERRIER, 2000).

Na avaliação do desempenho do SDM são empregadas métricas estatísticas, como por exemplo, a curva de análise de características operacionais do receptor (*Receiver Operating Characteristic* - ROC), kappa e a estatística de habilidade verdadeira (*True Skill Statistic* - TSS).

A curva ROC permite avaliar o desempenho de métodos de classificação binários com ordem de classificação ou valores de saída contínuos. Essa métrica avalia o desempenho de modelos e produz um mapa de probabilidade que apresenta a sequência em que o modelo seleciona células da grade (FIELDING; BELL, 1997; JIMÉNEZ-VALVERDE, 2012).

A métrica kappa é amplamente utilizada para avaliar a precisão de modelos baseados em presença-ausência, corrige a precisão geral das previsões do modelo pela precisão, que se espera que ocorra ao acaso. Kappa varia de -1 (valores de zero ou menos indicam um desempenho não melhor do que aleatório) a +1 (indica concordância perfeita) (COHEN, 1960).

A TSS compara o número de previsões corretas, menos aquelas atribuídas a suposições aleatórias de um conjunto hipotético de previsões perfeitas, e leva em consideração erros de omissão (falsos negativos; uma espécie é erroneamente considerada ausente) e comissão (falsos positivos; uma espécie é erroneamente considerada como estando presente). TSS varia de -1 a +1, onde +1 indica concordância perfeita e valores de zero ou menos indicam um desempenho não melhor do que aleatório (ALLOUCHE; TSOAR; KADMON, 2006; RONDININI *et al.*, 2006).

Com o objetivo de reduzir as incertezas nos modelos preditos, Araújo e New (2007) sugerem uma abordagem, na qual são construídos vários modelos usando diferentes algoritmos que posteriormente são combinados para produzir uma previsão de consenso (*ensemble modelling*). A modelagem de consenso tem sido amplamente aplicada, especificamente em plantas, essa abordagem foi utilizada com sucesso para a modelagem de 107 espécies raras a muito raras na Suíça, que são espécies difíceis de modelar devido ao número reduzido de dados de ocorrência e muitas variáveis preditoras (BREINER *et al.*, 2015).

A modelagem de consenso também se mostrou adequada para estudos com espécies invasoras na China, combinando modelos preditos pelos algoritmos random forest, modelo de aumento de gradiente, rede neural artificial e análise discriminante flexível (FANG *et al.*, 2021). E em parques nacionais nos Estados Unidos da América pela combinação de modelos preditos por regressão logística, árvores de regressão impulsionadas, random forest, splines de regressão adaptativa multivariada e *MaxEnt* (STOHLGREN *et al.*, 2010).

Estudos empregando SDM em *H. speciosa* foram realizados empregando o algoritmo *MaxEnt* para avaliar o impacto das mudanças climáticas globais na distribuição geográfica potencial, produtividade e valor da produção de *H. speciosa* no Brasil (NABOUT *et al.*, 2016) e para estimar a distribuição potencial atual e passada de *H. speciosa* com base na modelagem de consenso de modelos preditos por dez algoritmos (apenas presença, presença-background e presença-ausência) (COLLEVATTI *et al.*, 2018).

A combinação de múltiplas ferramentas como os marcadores ISSR, os modelos de distribuição de espécies e a genômica de populações possibilitam obter maior número de evidências sobre as espécies e podem contribuir para um melhor entendimento do estado atual das populações naturais em termos de áreas com maior adequabilidade ambiental - portanto prioritárias para estratégias de conservação; diversidade e estrutura genética - informações importantes para o melhoramento; que poderão nortear a gestão desse recurso visando sua conservação e uso sustentável.

3. REFERÊNCIAS BIBLIOGRÁFICAS

- AGUILAR, A.; TWARDOWSKI, T.; WOHLGEMUTH, R. Bioeconomy for Sustainable Development. **Biotechnology Journal**, [s. l.], v. 14, n. 8, p. 1800638, 2019.
- AGUILAR, A.; WOHLGEMUTH, R.; TWARDOWSKI, T. Perspectives on bioeconomy. **New Biotechnology**, [s. l.], v. 40, p. 181–184, 2018.
- ALFENAS, A. C. *et al.* **Eletroforese de proteínas e isoenzimas de fungos e essências florestais**. 1. ed. Viçosa: SIF, 1991. v. 1.
- ALLEN, B. *et al.* **International review of Bio-economy Strategies with a focus on waste resources**. London: Institute for European Environmental Policy, 2015.
- ALLENDORF, F. W.; HOHENLOHE, P. A.; LUIKART, G. Genomics and the future of conservation genetics. **Nature Reviews Genetics**, [s. l.], v. 11, n. 10, p. 697–709, 2010.
- ALLENDORF, F. W.; LUIKART, G. Conservation and the genetics of populations. **Mammalia**, [s. l.], v. 7, n. 4, p. 189–197, 2007.
- ALLOUCHE, O.; TSOAR, A.; KADMON, R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). **Journal of Applied Ecology**, [s. l.], v. 43, n. 6, p. 1223–1232, 2006.
- ALMEIDA, E. S. da S. *et al.* Fenologia de *Hancornia speciosa* Gomes (Apocynaceae) em Montes Claros de Goiás, Brasil. **Enciclopédia Biosfera**, [s. l.], v. 15, n. 27, p. 8–18, 2018.
- ALMEIDA, L. M. *et al.* *Hancornia speciosa* latex for biomedical applications: physical and chemical properties, biocompatibility assessment and angiogenic activity. **Journal of Materials Science: Materials in Medicine**, [s. l.], v. 25, n. 9, p. 2153–2162, 2014.
- ALMEIDA, L. M. *et al.* Toxicity of silver nanoparticles released by *Hancornia speciosa* (Mangabeira) biomembrane. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, [s. l.], v. 210, p. 329–334, 2019.
- ARAÚJO, M. B.; NEW, M. Ensemble forecasting of species distributions. **Trends in Ecology & Evolution**, [s. l.], v. 22, n. 1, p. 42–47, 2007.
- ARAVANOPOULOS, F. A. (Phil). Conservation and Monitoring of Tree Genetic Resources in Temperate Forests. **Current Forestry Reports**, [s. l.], v. 2, n. 2, p. 119–129, 2016.
- BAIRD, N. A. *et al.* Rapid SNP discovery and genetic mapping using sequenced RAD markers. **Plos One**, São Francisco, v. 3, n. 10, p. e3376, 2008.
- BARCACCIA, G. *et al.* Inheritance and mapping of 2n-egg production in diploid alfalfa. **Genome**, [s. l.], v. 43, n. 3, p. 528–537, 2011.
- BENTLEY, D. R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. **Nature**, [s. l.], v. 456, n. 7218, p. 53–59, 2008.
- BHANDARI, H. R. *et al.* Assessment of Genetic Diversity in Crop Plants - An Overview. **Advances in Plants & Agriculture Research**, [s. l.], v. 7, n. 3, p. 279–286, 2017.

BONETE, J. M. *et al.* Tissue reaction and anti-biofilm action of new biomaterial composed of latex from *Hancornia speciosa* Gomes and silver nanoparticles. **Anais da Academia Brasileira de Ciências**, [s. l.], v. 92, n. 4, p. 1–14, 2020.

BOOTH, T. H. *et al.* bioclim: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. **Diversity and Distributions**, [s. l.], v. 20, n. 1, p. 1–9, 2014.

BORNET, B.; BRANCHARD, M. Nonanchored Inter Simple Sequence Repeat (ISSR) Markers: Reproducible and Specific Tools for Genome Fingerprinting. **Plant Molecular Biology Reporter**, [s. l.], v. 19, n. 3, p. 209–215, 2001.

BORNET, B.; BRANCHARD, M. Nonanchored Inter Simple Sequence Repeat (ISSR) markers: Reproducible and specific tools for genome fingerprinting. **Plant Molecular Biology Reporter**, [s. l.], v. 9, p. 209–215, 2012.

BRASIL. **Portaria Nº 313, de 30 de Dezembro de 2019**. Brasília, 2019. Disponível em: <https://www.in.gov.br/web/dou/-/portaria-n-313-de-30-de-dezembro-de-2019-236263039#:~:text=%2D65%2C%20resolve%3A-,Art.,19%20de%20dezembro%20de%202019.&text=3%C2%BA%20Esta%20Portaria%20entra%20em%20vigor%20na%20data%20de%20sua%20publica%C3%A7%C3%A3o>. Acesso em: 10 set. 2021.

BREINER, F. T. *et al.* Overcoming limitations of modelling rare species by using ensembles of small models. **Methods in Ecology and Evolution**, [s. l.], v. 6, n. 10, p. 1210–1218, 2015.

BRUFORD, M. W. *et al.* Monitoring Changes in Genetic Diversity. In: WALTERS, M.; SCHOLLES, R. J. (org.). **The GEO Handbook on Biodiversity Observation Networks**. 1. ed. Gewerbestrasse: Springer Nature, 2017. p. 107–128. *E-book*.

CAMPOS, L. Z. *et al.* Use of local ecological knowledge as phenology indicator in native food species in the semiarid region of Northeast Brazil. **Ecological Indicators**, [s. l.], v. 95, p. 75–84, 2018.

CARPENTER, G.; GILLISON, A. N.; WINTER, J. DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. **Biodiversity and Conservation**, [s. l.], v. 2, n. 6, p. 667–680, 1993.

CHA, Y. K. *et al.* An interpretable machine learning method for supporting ecosystem management: Application to species distribution models of freshwater macroinvertebrates. **Journal of Environmental Management**, [s. l.], v. 291, p. 112719, 2021.

CHADIN, I. *et al.* Distribution of the invasive plant species *Heracleum sosnowskyi* manden. in the komi republic (Russia). **PhytoKeys**, [s. l.], v. 77, n. 1, p. 71–80, 2017.

CHARLESWORTH, B. Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. **Nature Reviews Genetics**, [s. l.], v. 10, n. 3, p. 195–205, 2009.

COHEN, J. A Coefficient of Agreement for Nominal Scales. **Educational and psychological measurement**, v. 20, n. 1, p. 37–46, 1960.

- COLLEVATTI, R. G. *et al.* Unravelling the genetic differentiation among varieties of the Neotropical savanna tree *Hancornia speciosa* Gomes. **Annals of botany**, [s. l.], v. 122, n. 6, p. 973–984, 2018.
- CORADIN, L.; CAMILLO, J.; PAREYN, F. G. C. **Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial: Plantas para o Futuro - Região Nordeste**. 1. ed. Brasília: Ministério do Meio Ambiente, 2018. v. 1.
- CORADIN, L.; SIMINSKI, A.; REIS, A. **Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial: Plantas para o Futuro - Região Sul**. 1. ed. Brasília: MMA, 2011. v. 1.
- DARRAULT, R. O.; SCHLINDWEIN, C. Limited Fruit Production in *Hancornia speciosa* (Apocynaceae) and Pollination by Nocturnal and Diurnal Insects1. **Biotropica**, [s. l.], v. 37, n. 3, p. 381–388, 2005.
- DAS, A. *et al.* Identification of Duplicates in Ginger Germplasm Collection from Odisha Using Morphological and Molecular Characterization. **Proceedings of the National Academy of Sciences, India Section B: Biological Sciences**, [s. l.], v. 90, n. 5, p. 1057–1066, 2020.
- DAVEY, J. W. *et al.* Genome-wide genetic marker discovery and genotyping using next-generation sequencing. **Nature Reviews Genetics**, [s. l.], v. 12, n. 7, p. 499–510, 2011.
- DESCHAMPS, S.; LLACA, V.; MAY, G. D. Genotyping-by-Sequencing in Plants. **Biology**, [s. l.], v. 1, n. 3, p. 460–483, 2012.
- DÍAZ, B. G. *et al.* Genome-wide SNP analysis to assess the genetic population structure and diversity of *Acrocomia* species. **Plos One**, São Francisco, v. 16, n. 7, p. e0241025, 2021.
- ELITH, J.; FRANKLIN, J. Species Distribution Modeling. **Encyclopedia of Biodiversity: Second Edition**, [s. l.], p. 692–705, 2013.
- ELSHIRE, R. J. *et al.* A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. **Plos One**, São Francisco, v. 6, n. 5, p. e19379, 2011.
- EUROPEAN UNION. **Innovating for Sustainable Growth - A Bioeconomy for Europe**. Luxembourg: [s. n.], 2012.
- FAJARDO, C. G. *et al.* Genetic diversity in natural populations of *Hancornia speciosa* gomes: Implications for conservation of genetic resources. **Ciencia e Agrotecnologia**, Lavras, v. 42, n. 6, p. 623–630, 2018.
- FANG, Y. *et al.* Predicting the invasive trend of exotic plants in China based on the ensemble model under climate change: A case for three invasive plants of Asteraceae. **Science of The Total Environment**, [s. l.], v. 756, p. 143841, 2021.
- FAO. **Global Forest Resources Assessment 2020**. Roma: FAO, 2020. Disponível em: <http://www.fao.org/documents/card/en/c/ca9825en/>. Acesso em: 10 set. 2021.
- FARBER, O.; KADMON, R. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. **Ecological Modelling**, [s. l.], v. 160, n. 1–2, p. 115–130, 2003.

FENG, X. *et al.* A checklist for maximizing reproducibility of ecological niche models. **Nature Ecology and Evolution**, [s. l.], v. 3, n. 10, p. 1382–1395, 2019.

FIELDING, A. H.; BELL, J. F. A review of methods for the assessment of prediction errors in conservation presence/absence models. **Environmental Conservation**, [s. l.], v. 24, n. 1, p. 38–49, 1997.

FITZPATRICK, M. C.; LACHMUTH, S.; HAYDT, N. T. The ODMAP protocol: a new tool for standardized reporting that could revolutionize species distribution modeling. **Ecography**, [s. l.], v. 44, n. 7, p. 1067–1070, 2021.

FLORA DO BRASIL. **Hancornia speciosa** Gomes. [S. l.], 2021. Disponível em: <http://floradobrasil.jbrj.gov.br/reflora/listaBrasil/FichaPublicaTaxonUC/FichaPublicaTaxonUC.do?id=FB15558>. Acesso em: 15 set. 2021.

FRANKLIN, J. Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. **Progress in Physical Geography**, [s. l.], v. 19, n. 4, p. 474–499, 1995.

FREELAND, J. R. **Molecular Ecology**. Nova Jersey: John Wiley & Sons, Ltd., 2005. *E-book*.

FRISVOLD, G. B. *et al.* Understanding the U.S. Bioeconomy: A New Definition and Landscape. **Sustainability**, [s. l.], v. 13, n. 4, p. 1627, 2021.

GBIF. **Free and open access to biodiversity data**. [S. l.], 2021. Disponível em: <https://www.gbif.org/>. Acesso em: 8 abr. 2021.

GILAD, Y.; PRITCHARD, J. K.; THORNTON, K. Characterizing natural variation using next-generation sequencing technologies. **Trends in Genetics**, [s. l.], v. 25, n. 10, p. 463–471, 2009.

GOVINDARAJ, M.; VETRIVENTHAN, M.; SRINIVASAN, M. Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. **Genetics Research International**, [s. l.], v. 2015, p. 1-14, 2015.

GOWAYED, S. M. H.; ABD EL-MONEIM, D. Detection of genetic divergence among some wheat (*Triticum aestivum* L.) genotypes using molecular and biochemical indicators under salinity stress. **Plos One**, São Francisco, v. 16, n. 3, p. e0248890, 2021.

GRAUDAL, L. *et al.* Global to local genetic diversity indicators of evolutionary potential in tree species within and outside forests. **Forest Ecology and Management**, [s. l.], v. 333, p. 35–51, 2014.

GRAUDAL, L. *et al.* **Indicators of the genetic diversity of trees-state, pressure, benefit and response**. Roma: [s. n.], 2020. Disponível em: <https://doi.org/10.4060/cb2492en>. Acesso em: 9 set. 2021.

GREISER, C. *et al.* Climate limitation at the cold edge: contrasting perspectives from species distribution modelling and a transplant experiment. **Ecography**, [s. l.], v. 43, n. 5, p. 637–647, 2020.

GRIMM, V. *et al.* A standard protocol for describing individual-based and agent-based models. **Ecological Modelling**, [s. l.], v. 198, n. 1–2, p. 115–126, 2006.

GRIMM, V. *et al.* The ODD protocol: A review and first update. **Ecological Modelling**, [s. l.], v. 221, n. 23, p. 2760–2768, 2010.

GUAJARDO, V. *et al.* Genome-wide SNP identification in *Prunus* rootstocks germplasm collections using Genotyping-by-Sequencing: phylogenetic analysis, distribution of SNPs and prediction of their effect on gene function. **Scientific Reports**, Londres, v. 10, n. 1, p. 1–14, 2020.

GUISAN, A.; ZIMMERMANN, N. E. Predictive habitat distribution models in ecology. **Ecological Modelling**, [s. l.], v. 135, n. 2–3, p. 147–186, 2000.

HARTL, D. L.; CLARK, A. G. **Princípios de genética de populações**. 4. ed. Porto Alegre: Artmed, 2010.

HAGHI VAYGHAN, A. *et al.* Suitability indices and habitat suitability index model of Caspian kutum (*Rutilus frisii kutum*) in the southern Caspian Sea. **Aquatic Ecology**, [s. l.], v. 47, n. 4, p. 441–451, 2013.

HEIDARI, P. *et al.* Phenotypic variability of *Pyrus boissieriana* Buhse: Implications for conservation and breeding. **Scientia Horticulturae**, [s. l.], v. 247, p. 1–8, 2019.

HIRZEL, A. H. *et al.* Ecological-Niche factor analysis: How to compute habitat-suitability maps without absence data?. **Ecology - Ecology Society of America**, [s. l.], v. 83, n. 7, p. 2027–2036, 2002.

HIRZEL, A. H. *et al.* Evaluating the ability of habitat suitability models to predict species presences. **Ecological Modelling**, [s. l.], v. 199, n. 2, p. 142–152, 2006.

HOLLIDAY, J. A. *et al.* Advances in ecological genomics in forest trees and applications to genetic resources conservation and breeding. **Molecular Ecology**, [s. l.], v. 26, n. 3, p. 706–717, 2017.

HUANG, L. *et al.* Genetic stability and DNA fingerprinting of the *Hemarthria compressa* cultivar “Guangyi.” **Biochemical Systematics and Ecology**, [s. l.], v. 55, p. 310–316, 2014.

HUANG, Y.-F. *et al.* Using Genotyping-By-Sequencing (GBS) for Genomic Discovery in Cultivated Oat. **Plos One**, São Francisco, v. 9, n. 7, p. e102448, 2014.

IBGE. **Produção da Extração Vegetal e da Silvicultura | PEVS 2019**. [S. l.], 2019. Disponível em: <https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9105-producao-da-extracao-vegetal-e-da-silvicultura.html?=&t=sobre>. Acesso em: 12 set. 2021.

ISSA, I.; DELBRÜCK, S.; HAMM, U. Bioeconomy from experts’ perspectives – Results of a global expert survey. **Plos One**, São Francisco, v. 14, n. 5, p. e0215917, 2019.

JEDRZEJCZYK, I.; REWERS, M. Genome size and ISSR markers for *Mentha* L. (Lamiaceae) genetic diversity assessment and species identification. **Industrial Crops and Products**, [s. l.], v. 120, p. 171–179, 2018.

JEDRZEJCZYK, I.; REWERS, M. Identification and genetic diversity analysis of edible and medicinal *Malva* species using flow cytometry and ISSR molecular markers. **Agronomy**, [s. l.], v. 10, n. 5, 2020.

- JIMÉNEZ-VALVERDE, A. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. **Global Ecology and Biogeography**, [s. l.], v. 21, n. 4, p. 498–507, 2012.
- KOLANOWSKA, M.; JAKUBSKA-BUSSE, A. Is the lady's-slipper orchid (*Cypripedium calceolus*) likely to shortly become extinct in Europe? - Insights based on ecological niche modelling. **Plos One**, São Francisco, v. 15, n. 1, p. 1–21, 2020.
- KUMAR, M. *et al.* Biochemical and molecular markers for characterization of chrysanthemum germplasm: A review. **Journal of Pharmacognosy and Phytochemistry**, [s. l.], v. 7, n. 5, p. 2641–2652, 2018.
- LAHOZ-MONFORT, J. J.; GUILLERA-ARROITA, G.; WINTLE, B. A. Imperfect detection impacts the performance of species distribution models. **Global Ecology and Biogeography**, [s. l.], v. 23, n. 4, p. 504–515, 2014.
- LAMSAL, P. *et al.* Invasive alien plant species dynamics in the Himalayan region under climate change. **Ambio**, [s. l.], v. 47, n. 6, p. 697–710, 2018.
- LAWSON, C. R. *et al.* Prevalence, thresholds and the performance of presence–absence models. **Methods in Ecology and Evolution**, [s. l.], v. 5, n. 1, p. 54–64, 2014.
- LEBEDEV, V. G. *et al.* Genomic Selection for Forest Tree Improvement: Methods, Achievements and Perspectives. **Forests**, [s. l.], v. 11, n. 11, p. 1190, 2020.
- LEDERMAN, I. E. *et al.* **Mangaba (*Hancornia speciosa* Gomes)**. 1. ed. Jaboticabal: FUNEP, 2000. v. 1.
- LENS.ORG. **The Lens - Free & Open Patent and Scholarly Search**. [S. l.], 2021. Disponível em: <https://www.lens.org/>. Acesso em: 18 set. 2021.
- LEWANDOWSKI, I. **Bioeconomy: Shaping the Transition to a Sustainable, Biobased Economy**, [s. l.], p. 356, 2018.
- LIMA, D. F. *et al.* Species boundaries inferred from ISSR markers in the *Myrcia laruotteana* complex (Myrtaceae). **Plant Systematics and Evolution**, [s. l.], v. 301, n. 1, p. 353–363, 2015.
- LIU, C. *et al.* Species distribution modelling for conservation planning in Victoria, Australia. **Ecological Modelling**, [s. l.], v. 249, n. 2011, p. 68–74, 2013.
- LIU, C.; WHITE, M.; NEWELL, G. Measuring and comparing the accuracy of species distribution models with presence–absence data. **Ecography**, [s. l.], v. 34, n. 2, p. 232–243, 2011.
- LÚCIO, D. de M. *et al.* Biometry of fruits and seeds of *Hancornia speciosa* Gomes (Apocynaceae) in natural populations. **Agriculture and Forestry**, [s. l.], v. 65, n. 1, p. 99–110, 2019.
- LUZ, G. de A. *et al.* Genetic diversity and population structure of mangabeira (*Hancornia speciosa*) estimated using ISSR markers. **Acta Scientiarum. Biological Sciences**, [s. l.], v. 42, p. e53606, 2020.

MARTÍNEZ-MINAYA, J. *et al.* Spatial and climatic factors associated with the geographical distribution of citrus black spot disease in South Africa. A Bayesian latent Gaussian model approach. **European Journal of Plant Pathology**, [s. l.], v. 151, n. 4, p. 991-1007, 2018.

MCTIC. **Estratégia Nacional de Ciência, Tecnologia e Inovação 2016|2022**. Brasília: Ministério da Ciência, Tecnologia, Inovações e Comunicações, 2016.

MIEDANER, T.; KORZUN, V.; BAUER, E. Genomics-Based Hybrid Rye Breeding. In: MIEDANER, T.; KORZUN, V. (org.). **Applications of Genetic and Genomic Research in Cereals**. 1. ed. Amsterdam: Elsevier, 2019. p. 329–348. v. 1.

MIRANDA, E. C.; RUIZ-CABELLO, M. V. C.; HURTADO, M. C. Food biopharmaceuticals as part of a sustainable bioeconomy: Edible vaccines case study. **New Biotechnology**, [s. l.], v. 59, p. 74–79, 2020.

MMA. **Biodiversidade Brasileira**. [S. l.], 2021. Disponível em: <https://antigo.mma.gov.br/biodiversidade/biodiversidade-brasileira.html>. Acesso em: 12 set. 2021.

MONACHINO, J. A Revisión de *Hancornia* (Apocynaceae). **Lilloa**, [s. l.], v. 11, p. 19–48, 1945.

MONDINI, L.; NOORANI, A.; PAGNOTTA, M. A. Assessing Plant Genetic Diversity by Molecular Tools. **Diversity**, [s. l.], v. 1, n. 1, p. 19–35, 2009.

MORADI, S. *et al.* Identifying high-priority conservation areas for avian biodiversity using species distribution modeling. **Ecological Indicators**, [s. l.], v. 97, n. 2018, p. 159–164, 2019.

MORI, A. S.; LERTZMAN, K. P.; GUSTAFSSON, L. Biodiversity and ecosystem services in forest ecosystems: a research agenda for applied forest ecology. **Journal of Applied Ecology**, [s. l.], v. 54, n. 1, p. 12–27, 2017.

MURRAY, B. G. Plant Diversity, Conservation and Use. In: THOMAS, B.; MURRAY, B. G.; MURPHY, D. J. (org.). **Encyclopedia of Applied Plant Science: Breeding Genetics and Biotechnology**. 2. ed. Amsterdam: Elsevier, 2017. p. 289–308.

NABOUT, J. C. *et al.* The Impact of Global Climate Change on the Geographic Distribution and Sustainable Harvest of *Hancornia speciosa* Gomes (Apocynaceae) in Brazil. **Environmental Management**, [s. l.], v. 57, n. 4, p. 814–821, 2016.

NARUM, S. R. *et al.* Genotyping-by-sequencing in ecological and conservation genomics. **Molecular ecology**, [s. l.], v. 22, n. 11, p. 2847, 2013.

NEI, M. Analysis of Gene Diversity in Subdivided Populations. **Proceedings of the National Academy of Sciences**, [s. l.], v. 70, n. 12, p. 3321–3323, 1973.

NG, W. L.; TAN, S. G. Inter-Simple Sequence Repeat (ISSR) markers: Are we doing it right? **ASM Science Journal**, [s. l.], v. 9, n. 1, p. 30–39, 2015.

NIX, H. A. A biogeographic analysis of Australian elapid snakes. In: **Atlas of elapid snakes of Australia**. 7. ed. Canberra: Australian Government Publishing Service, 1986. p. 4–15.

NOOKARAJU, A.; AGRAWAL, D. C. Genetic homogeneity of in vitro raised plants of grapevine cv. Crimson Seedless revealed by ISSR and microsatellite markers. **South African Journal of Botany**, [s. l.], v. 78, p. 302–306, 2012.

NUNES, V. V. *et al.* Geno-phenotypic diversity in a natural population of *Hancornia speciosa* Gomes: implications for conservation and improvement. **Genetic Resources and Crop Evolution**, [s. l.], v. 68, n. 7, p. 2869–2882, 2021.

OECD-FAO. OECD-FAO Agricultural Outlook 2020-2029. 2020. Disponível em: <https://doi.org/10.1787/1112c23b-en>. Acesso em: 10 set. 2021.

OLIVEIRA, K. S. de; ALOUFA, M. A. I. Extrativismo e geração de renda da mangaba (*Hancornia speciosa* Gomes) em região litorânea do Rio Grande do Norte. **Revista Brasileira de Agroecologia**, [s. l.], v. 15, n. 5, p. 291–304, 2020.

PATRA, B. C. *et al.* Rice. In: SINGH, M.; UPADHYAYA, H. D. (org.). **Genetic and Genomic Resources for Grain Cereals Improvement**. 1. ed. Amsterdam: Elsevier, 2016. v. 1, p. 1–80.

PEARCE, J.; FERRIER, S. Evaluating the predictive performance of habitat models developed using logistic regression. **Ecological Modelling**, [s. l.], v. 133, n. 0000, p. 225–245, 2000. Disponível em: www.elsevier.com/locate/ecolmodel. Acesso em: 9 set. 2021.

PEARSON, R. G.; DAWSON, T. P. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? **Global Ecology and Biogeography**, [s. l.], v. 12, n. 5, p. 361–371, 2003.

PEREIRA, A. V. P. *et al.* *Hancornia speciosa*: mangaba. In: VIEIRA, R. F.; CAMILLO, J.; CORADIN, L. (org.). **Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial: Plantas para o Futuro - Região Centro-Oeste**. 1. ed. Brasília: MMA, 2016. p. 237–256. v. 1.

PETERSON, A. T.; SOBERÓN, J. Species Distribution Modeling and Ecological Niche Modeling: Getting the Concepts Right. **Brazilian Journal of Nature Conservation Essays & Perspectives Natureza & Conservação**, [s. l.], v. 10, n. 2, p. 102–107, 2012.

PETERSON, B. K. *et al.* Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. **Plos One**, São Francisco, v. 7, n. 5, p. e37135, 2012.

PHILLIPS, S. J.; DUDÍK, M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. **Ecography**, [s. l.], v. 31, n. 2, p. 161–175, 2008.

PINHEIRO, E. *et al.* **O latex e a borracha da mangabeira**. Belém: Embrapa Amazônia Oriental. 28p. 2004. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/406036/o-latex-e-a-borracha-da-mangabeira>. Acesso em: 4 set. 2021.

PLANTS OF THE WORLD ONLINE. *Hancornia speciosa* Gomes. Disponível em: <http://www.plantsoftheworldonline.org/taxon/urn:lsid:ipni.org:names:328991-2>. Acesso em: 26 out. 2021.

POLAND, J. A. *et al.* Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. **Plos One**, São Francisco, v. 7, n. 2, p. e32253, 2012.

POLAND, J. A.; RIFE, T. W. Genotyping-by-Sequencing for Plant Breeding and Genetics. **The Plant Genome**, [s. l.], v. 5, n. 3, 2012.

PORFIRIO, L. L. *et al.* Improving the Use of Species Distribution Models in Conservation Planning and Management under Climate Change. **Plos One**, São Francisco, v. 9, n. 11, p. e113749, 2014.

PORTAL UFS. **Pesquisadores desenvolvem bioplástico a partir de resíduos de cana-de-açúcar e sementes de mangaba**. 2021. Disponível em: <https://www.ufs.br/conteudo/67983-pesquisadores-desenvolvem-bioplastico-a-partir-de-residuos-de-cana-de-acucar-e-sementes-de-mangaba>. Acesso em: 12 set. 2021.

RABEY, H. A. el; ALSHUBAILY, F.; AL-OTAIBI, K. M. Phylogenetic relationships of some economically important cereal plants based on genome characterization using molecular markers. **Firenze University Press**, [s. l.], v. 68, n. 3, p. 225–232, 2015.

RAFALSKI, A. Applications of single nucleotide polymorphisms in crop genetics. **Current Opinion in Plant Biology**, [s. l.], v. 5, n. 2, p. 94–100, 2002.

RAO, V. R.; HODGKIN, T. Genetic diversity and conservation and utilization of plant genetic resources. **Plant Cell, Tissue and Organ Culture**, [s. l.], v. 68, n. 1, p. 1–19, 2002.

RIBEIRO, D. O. *et al.* Genetic variability in *Jatropha curcas* L. from diallel crossing. **Genetics and Molecular Research**, [s. l.], v. 16, n. 2, p. 16029651, 2017.

RICE, L. J. *et al.* DNA fingerprinting of *Plectranthus* plants for protection of cultivar registration. **South African Journal of Botany**, [s. l.], v. 76, n. 2, p. 401, 2010.

RONDININI, C. *et al.* Tradeoffs of different types of species occurrence data for use in systematic conservation planning. **Ecology Letters**, [s. l.], v. 9, n. 10, p. 1136–1145, 2006.

ROTHBERG, J. M. *et al.* An integrated semiconductor device enabling non-optical genome sequencing. **Nature**, [s. l.], v. 475, n. 7356, p. 348–352, 2011.

SAIFERT, L. *et al.* Evaluation of the genetic diversity in the feijoa accessions maintained at Santa Catarina, Brazil. **Crop Science**, [s. l.], v. 60, n. 1, p. 345–356, 2020.

SANTOS, P. S. *et al.* Genetic diversity and the quality of Mangabeira tree fruits (*Hancornia speciosa* Gomes – Apocynaceae), a native species from Brazil. **Scientia Horticulturae**, [s. l.], v. 226, p. 372–378, 2017.

SANTOS, R. M. *et al.* Pyrolysis of mangaba seed: Production and characterization of bio-oil. **Bioresource Technology**, [s. l.], v. 196, p. 43–48, 2015.

SANTOS, E. A.; SOUZA, R. M. Territorialidade das catadoras de mangaba no litoral sul de Sergipe. **Geosaberes Revista Estudos Geoeducacionais**, [s. l.], v. 6, n. 3, p. 629–642, 2015.

SASSON, A.; MALPICA, C. Bioeconomy in Latin America. **New Biotechnology**, [s. l.], v. 40, p. 40–45, 2018.

SILVA JÚNIOR, J. F. da *et al.* *Hancornia speciosa*: mangaba. In: CORADIN, L.; CAMILLO, J.; PAREYN, F. G. C. (org.). **Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial - Plantas para o Futuro: Região Nordeste**. 1. ed. Brasília: Ministério do Meio Ambiente, 2018. p. 177–192. v. 1.

- SOARES, F. S. *et al.* Diversity and genetic structure of mangabeira (*Hancornia speciosa* Gomes), a fruit species from Cerrado. **Semina: Ciências Agrárias**, [s. l.], v. 38, n. 4, p. 2479-2488, 2017.
- SOUSA, A. *et al.* ISSR markers in wild species of *Passiflora* L. (Passifloraceae) as a tool for taxon selection in ornamental breeding. **Genetics and Molecular Research**, [s. l.], v. 14, n. 4, p. 18534–18545, 2015.
- SOUSA-SILVA, R. *et al.* Improving the assessment and reporting on rare and endangered species through species distribution models. **Global Ecology and Conservation**, [s. l.], v. 2, p. 226–237, 2014.
- SOUZA, L. M. *et al.* Genomic Selection in Rubber Tree Breeding: A Comparison of Models and Methods for Managing G×E Interactions. **Frontiers in Plant Science**, [s. l.], v. 10, p. 1353, 2019.
- STOHLGREN, T. J. *et al.* Ensemble Habitat Mapping of Invasive Plant Species. **Risk Analysis**, [s. l.], v. 30, n. 2, p. 224–235, 2010.
- SYVANEN, A.-C. Accessing genetic variation: genotyping single nucleotide polymorphisms. **Nature Reviews Genetics**, [s. l.], v. 2, n. 12, p. 930–943, 2001.
- TAX, D. M. J.; DUIN, R. P. W. Support Vector Data Description. **Machine Learning**, [s. l.], v. 54, p. 45–66, 2004.
- THAKUR, J. *et al.* Genetic Homogeneity Revealed Using SCoT, ISSR and RAPD Markers in Micropropagated *Pittosporum eriocarpum* Royle- An Endemic and Endangered Medicinal Plant. **Plos One**, São Francisco, v. 11, n. 7, p. e0159050, 2016.
- THUDI, M. *et al.* Current state-of-art of sequencing technologies for plant genomics research. **Briefings in Functional Genomics**, [s. l.], v. 11, n. 1, p. 3–11, 2012.
- TIKENDRA, L.; AMOM, T.; NONGDAM, P. Molecular genetic homogeneity assessment of micropropagated *Dendrobium moschatum* Sw. - A rare medicinal orchid, using RAPD and ISSR markers. **Plant Gene**, [s. l.], v. 19, p. 100196, 2019.
- TOMAZI, R. *et al.* Hypoglycemic Activity of Aqueous Extract of Latex from *Hancornia speciosa* Gomes: A Study in Zebrafish and In Silico. **Pharmaceuticals**, [s. l.], v. 14, n. 9, p. 856, 2021.
- TORKAMANEH, D.; LAROCHE, J.; BELZILE, F. Genome-Wide SNP Calling from Genotyping by Sequencing (GBS) Data: A Comparison of Seven Pipelines and Two Sequencing Technologies. **Plos One**, São Francisco, v. 11, n. 8, p. e0161333, 2016.
- TRIEU, L. N. *et al.* Genetic diversity of *Panax stipuleanatus* Tsai in North Vietnam detected by inter simple sequence repeat (ISSR) markers. **Biotechnology & Biotechnological Equipment**, [s. l.], v. 30, n. 3, p. 506–511, 2016.
- TROPICOS. **Tropicos.com.org**. 2021. Disponível em: <https://www.tropicos.org/home>. Acesso em: 15 set. 2021.
- TURCHETTO-ZOLET, A. C. *et al.* **Marcadores Moleculares na Era genômica: Metodologias e Aplicações**. 1. ed. Ribeirão Preto: Sociedade Brasileira de Genética, 2017. v. 1.

UDOH, L. I.; OBASEOJEI, W. P.; UZOEBU, C. Single Nucleotide Polymorphisms: A Modern Tool to Screen Plants for Desirable Traits. In: ABDURAKHMONOV, I. Y. (org.). **Plant Breeding - Current and Future Views**. London: IntechOpen, 2021. p. 1–14. *E-book*. Disponível em: <https://www.intechopen.com/chapters/74230>. Acesso em: 30 set. 2021.

UNITED NATIONS. **Transforming our world: the 2030 Agenda for Sustainable Development**. 2021. Disponível em: <https://sdgs.un.org/2030agenda>. Acesso em: 4 set. 2021.

UNITED NATIONS. **World Population Prospects 2019: Highlights**. 2019. Disponível em: https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf. Acesso em: 25 set. 2021.

VAN OEVEREN, J.; JANSSEN, A. Mining SNPs from DNA sequence data; computational approaches to SNP discovery and analysis. **Methods in Molecular Biology**, [s. l.], v. 578, p. 73–91, 2009.

VAZ, Ú. L.; NABOUT, J. C. Using ecological niche models to predict the impact of global climate change on the geographical distribution and productivity of *Euterpe oleracea* Mart. (Arecaceae) in the Amazon. **Acta Botanica Brasilica**, [s. l.], v. 30, n. 2, p. 290–295, 2016.

VENCOVSKY, R.; CROSSA, J. Measurements of Representativeness Used in Genetic Resources Conservation and Plant Breeding. **Crop Science**, [s. l.], v. 43, n. 6, p. 1912–1921, 2003.

VIEIRA, R. F.; CAMILLO, J.; CORADIN, L. **Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial: Plantas para o Futuro - Região Centro-Oeste**. 1. ed. Brasília: MMA, 2016. v. 1.

VILAS BOAS, E. *et al.* *Hancornia speciosa*. In: LIMA, F. F.; LESCANO C. H.; OLIVEIRA, I. P. (eds). **Fruits of the Brazilian Cerrado**, Cham: Springer, 2021. p. 161–180.

WAMBUGU, P. W.; NDJIONDJOP, M. N.; HENRY, R. J. Role of genomics in promoting the utilization of plant genetic resources in genebanks. **Briefings in Functional Genomics**, [s. l.], v. 17, n. 3, p. 198–206, 2018.

WESTBROOK, J. W. *et al.* Optimizing genomic selection for blight resistance in American chestnut backcross populations: A trade-off with American chestnut ancestry implies resistance is polygenic. **Evolutionary Applications**, [s. l.], v. 13, n. 1, p. 31–47, 2020.

WHITE, L. A. S. *et al.* Genetic diversity of a native population of *Myrcia ovata* (Myrtaceae) using ISSR molecular markers. [s. l.], 2018.

WHITE HOUSE. **National Bioeconomy Blueprint**. Washington: The White House, 2012. Disponível em: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/national_bioeconomy_blueprint_april_2012.pdf. Acesso em: 25 set. 2021.

WILLIAMS, J. N. *et al.* Using species distribution models to predict new occurrences for rare plants. **Diversity and Distributions**, [s. l.], v. 15, n. 4, p. 565–576, 2009.

WORLDCLIM. **Global climate and weather data — WorldClim**. 2020. Disponível em: <https://www.worldclim.org/data/cmip6/cmip6climate.html> https://worldclim.org/data/index.html. Acesso em: 8 set. 2021.

YOUSEFI, V. *et al.* Molecular characterization of *Thymus* species using ISSR markers. **Journal Animal Plant Science**, [s. l.], v. 25, n. 4, p. 1087–1094, 2015.

ZHAO, L. *et al.* Assessment of the genetic diversity and genetic relationships of *Lilium* in China using ISSR markers. **Biochemical Systematics and Ecology**, [s. l.], v. 55, p. 184–189, 2014.

ZHU, S. *et al.* Genetic Diversity Analysis of Olive Germplasm (*Olea europaea* L.) With Genotyping-by-Sequencing Technology. **Frontiers in Genetics**, [s. l.], v. 10, n. JUL, p. 755, 2019.

ZIETKIEWICZ, E.; RAFALSKI, A.; LABUDA, D. Genome Fingerprinting by Simple Sequence Repeat (SSR)-Anchored Polymerase Chain Reaction Amplification. **Genomics**, [s. l.], v. 20, n. 2, p. 176–183, 1994.

ZURELL, D. *et al.* A standard protocol for reporting species distribution models. **Ecography**, [s. l.], v. 43, n. 9, p. 1261–1277, 2020.

4. ARTIGO 1

POTENCIAL FARMACOLÓGICO, ALIMENTÍCIO E DADOS MOLECULARES DE *Hancornia speciosa*: UMA REVISÃO SISTEMÁTICA

Periódico (aceito para publicação): Genetic Resources and Crop Evolution

RESUMO

Hancornia speciosa é uma árvore frutífera e medicinal nativa da América do Sul. No Brasil possui valor econômico atual ou potencial e é prioritária para o desenvolvimento de pesquisas. O objetivo do presente trabalho foi mapear o estado da arte por meio de artigos, patentes e dados moleculares da espécie visando elencar perspectivas para novas pesquisas. Avaliou-se a produção científica anual, estrutura intelectual, social e conceitual; patentes - números de depósitos por base, parte da planta utilizada, países de depósito, classificação internacional e cessionários; bem como dados moleculares disponíveis. O Brasil figura como o país com maior número de publicações (303) no período de 1992 a 2020, com uma taxa de crescimento percentual anual de 9,28, e produtos tecnológicos (29) desenvolvidos a partir de diferentes partes da planta. Os artigos e patentes foram desenvolvidos majoritariamente por pesquisadores afiliados a universidades públicas localizadas nas diferentes regiões do Brasil, que são também as principais cessionárias das patentes. Os dados moleculares foram sequências de nucleotídeos (164) e proteínas (236) que pertencem ao genoma cloroplastidial e estão relacionados à identificação da espécie como códigos de barra de DNA e proteínas que atuam na fotossíntese. A compilação e análise das informações científicas, tecnológicas e moleculares realizadas neste trabalho permitiram identificar as linhas de pesquisas desenvolvidas, as lacunas do conhecimento sobre a espécie e contribuíram para direcionar novas pesquisas.

Palavras-chave: biodiversidade, bioeconomia, etnofarmacologia, mangaba, mapeamento científico, patentes.

ABSTRACT

Título: Pharmacological, food potential, and molecular data of *Hancornia speciosa*: a systematic review

Hancornia speciosa is a fruit and medicinal tree species native to South America. In Brazil, it is considered of potential economic value and priority for research and development. In this work, the state-of-art map is presented, including articles, patents, and the species molecular, in order to identify perspectives for future research. The annual scientific production, intellectual, social, and conceptual structure were evaluated, along with the number of patent deposits, components of the plant used, countries of deposit, international classification and assignees, and the accessibility of available molecular data. Brazil has the most significant publications (306) between 1992 and 2020. Technological products (29) have been developed from different plant tissues. Most of the articles and patents were developed by researchers from public universities from different regions of Brazil. The molecular data are sequences of nucleotides (164) and proteins (236) of the chloroplast genome and are described to identify the specie as DNA barcodes and proteins involved in photosynthesis. The compilation and report of scientific, technological, and molecular information in the present review allowed the identification of new researching perspectives to be developed based on the gaps in knowledge regarding the species and perspectives for the definition of future research.

Key-words: biodiversity, bioeconomy, ethnopharmacology, mangaba, patents, scientific mapping.

4.1. Introdução

A América do Sul adotou princípios de bioeconomia com diferentes níveis de impacto socioeconômico, entre eles destaca-se a valorização dos recursos da biodiversidade na medicina (Sasson e Malpica 2018). Dada as limitações econômicas, há necessidade de remédios mais baratos e eficazes, que possam servir como alternativa e/ou como medicação complementar para o tratamento de algumas doenças, principalmente nas categorias de renda média-alta, média-baixa e baixa (Goyal e Ayeleso 2018).

Desde a antiguidade as plantas são fonte de produtos essenciais para a vida humana, incluindo alimentos e medicamentos, principalmente em comunidades em situação de vulnerabilidade socioeconômica. *Hancornia speciosa* Gomes (Apocynaceae) é uma espécie arbórea frutífera e medicinal nativa da América do Sul e está entre as inúmeras plantas utilizadas por comunidades tradicionais. Conhecida popularmente como ‘mangabeira’ (ou seja, coisa boa de comer, no dialeto indígena tupi-guarani), era usada no século passado para a extração de látex como matéria-prima para a produção de borracha. No entanto, tal atividade foi interrompida devido à baixa produtividade em comparação com *Hevea brasiliensis* L. (seringueira) (Vieira Neto et al. 2002; Bonete et al. 2020).

Suas populações estão amplamente distribuídas na região biogeográfica neotropical, especificamente no Noroeste e Centro-Oeste do Brasil, Paraguai, Peru e Bolívia (GBIF 2021; Tropicos 2021). A espécie é uma planta perene, semidecídua ou caducifólia, mudando suas folhas durante o período mais seco do ano. A fase de frutificação varia nas diferentes regiões do Brasil. No litoral do Nordeste existem duas fases de frutificação: inverno (maio a julho ou julho a outubro) e verão (janeiro a abril ou dezembro a abril) (Oliveira e Aloufa 2020) e em áreas do cerrado os frutos também podem ser coletados ao longo do ano. No Centro-Oeste é relatada apenas uma coleta por ano, de outubro a maio (Campos et al. 2018).

Os frutos são colhidos principalmente por extrativistas que os selecionam em populações naturais de comunidades locais. Em *H. speciosa*, a pesquisa sobre domesticação e

técnicas agronômicas para cultivo sustentável e comercial ainda está em seus estágios iniciais (Paula et al. 2019). Existem indícios de que o plantio comercial esteja se estabelecendo nas regiões Nordeste e Centro-Oeste (EMATER 2019), porém na literatura científica não há dados disponíveis.

As frutas são comercializadas *in natura* ou industrializadas como polpas para sucos, sorvetes, licores e geleias (Cardoso et al. 2014; Santos et al. 2017). Há grande interesse porque são ricas em potássio, ferro, zinco, ácido ascórbico, compostos fenólicos e carotenoides, como β -caroteno, β -criptoxantina, α -tocoferol e α -, β - e γ -tocotrienóis (Cardoso et al. 2014; De Lima et al. 2015a, b).

Essa planta também é tradicionalmente utilizada no tratamento de diversas doenças por comunidades do Brasil (Cercato et al. 2015). O látex possui ação anti-inflamatória (Marinho et al. 2011), angiogênica (Almeida et al. 2019) e osteogênica (Floriano et al. 2016) e as folhas possuem efeito antidiabético (Pereira et al. 2015), propriedades antioxidantes (Marques et al. 2015) que proporcionam a cicatrização de feridas e atividade anti-inflamatória (Geller et al. 2015). Registros etnobotânicos, químicos e farmacológicos de espécies de Apocynaceae Juss, destacam *H. speciosa* como a mais citada entre 78 espécies distribuídas em 27 gêneros no Brasil (Santos et al. 2013).

A espécie é conservada principalmente por meio de plantas vivas no campo em Bancos Ativos de Germoplasma, são 1.438 acessos distribuídos em oito bancos mantidos por instituições públicas, Embrapa Amapá (86 acessos), Embrapa Cerrados (15 acessos), Embrapa Meio Norte (39 acessos), Embrapa Tabuleiros Costeiros (271 acessos), Empresa Estadual de Pesquisa Agropecuária da Paraíba (540 acessos), Universidade Federal de Alagoas (20 acessos), Universidade Federal de Goiás (57 acessos) e Universidade Estadual de Goiás (400 acessos) (Da Silva Júnior et al. 2018a).

H. speciosa integra a lista de espécies da flora brasileira que possui valor econômico atual ou potencial, considerada espécie prioritária para o desenvolvimento de pesquisas (Coradin et al. 2018) e, apesar de apresentar importância bioeconômica, compõe a lista de espécies subutilizadas que contribuem para os Indicadores Nutricionais de Biodiversidade (FAO 2021). Portanto, é fundamental mapear estudos e lacunas existentes para fomentar e direcionar futuras pesquisas sobre a espécie. Nesse sentido, revisões sistemáticas empregando análises bibliométricas são úteis, pois fornecem análises objetivas e confiáveis (Aria e Cuccurullo 2017).

Uma das etapas para a realização de estudos bibliométricos é a obtenção do conjunto de metadados científicos que são fontes de informações bibliográficas (Zupic e Čater 2015) e estão disponíveis em bases de dados bibliográficos online. A bibliometria possibilita a mensuração e a avaliação dos resultados de pesquisas apresentados na forma escrita, nas mais diversas representações gráficas (Ball 2017) e são relevantes para a compreensão do cenário atual no qual se encontra o tema de interesse.

Além de metadados científicos, as informações técnicas sobre patentes disponíveis em bases como *World Intellectual Property Organization*, *European Patent Office*, *Espacenet-Latipat* e o Instituto Nacional de Propriedade Industrial, e dados moleculares disponíveis em bancos públicos como o *Genetic Sequence Database*, *European Molecular Biology Laboratory*, *DNA Data Bank of Japan*, *Universal Protein Resource* e o *Barcode of Life Data System*, auxiliam na completamentação das informações, podendo inclusive ser usados na avaliação do nível de domesticação da espécie.

Esta revisão se propõe a mapear (1) estudos científicos, (2) produtos tecnológicos patenteados e (3) dados moleculares em bancos públicos para *H. speciosa*, visando contribuir no direcionamento de pesquisas futuras.

4.2. Material e Métodos

O levantamento bibliográfico foi realizado em novembro de 2020 com as seguintes palavras-chave: “*Hancornia speciosa*” OR “mangaba” OR “mangaba tree”, que foram prospectadas no título, resumo e/ou palavras-chave dos artigos científicos; bem como em título e resumo das patentes.

Para os dados moleculares, as sequências de nucleotídeos e proteínas foram prospectadas em bancos públicos utilizando o termo “*Hancornia speciosa*”. Não foi delimitado um período a fim de se obter o máximo de informações possível (Figura 1).

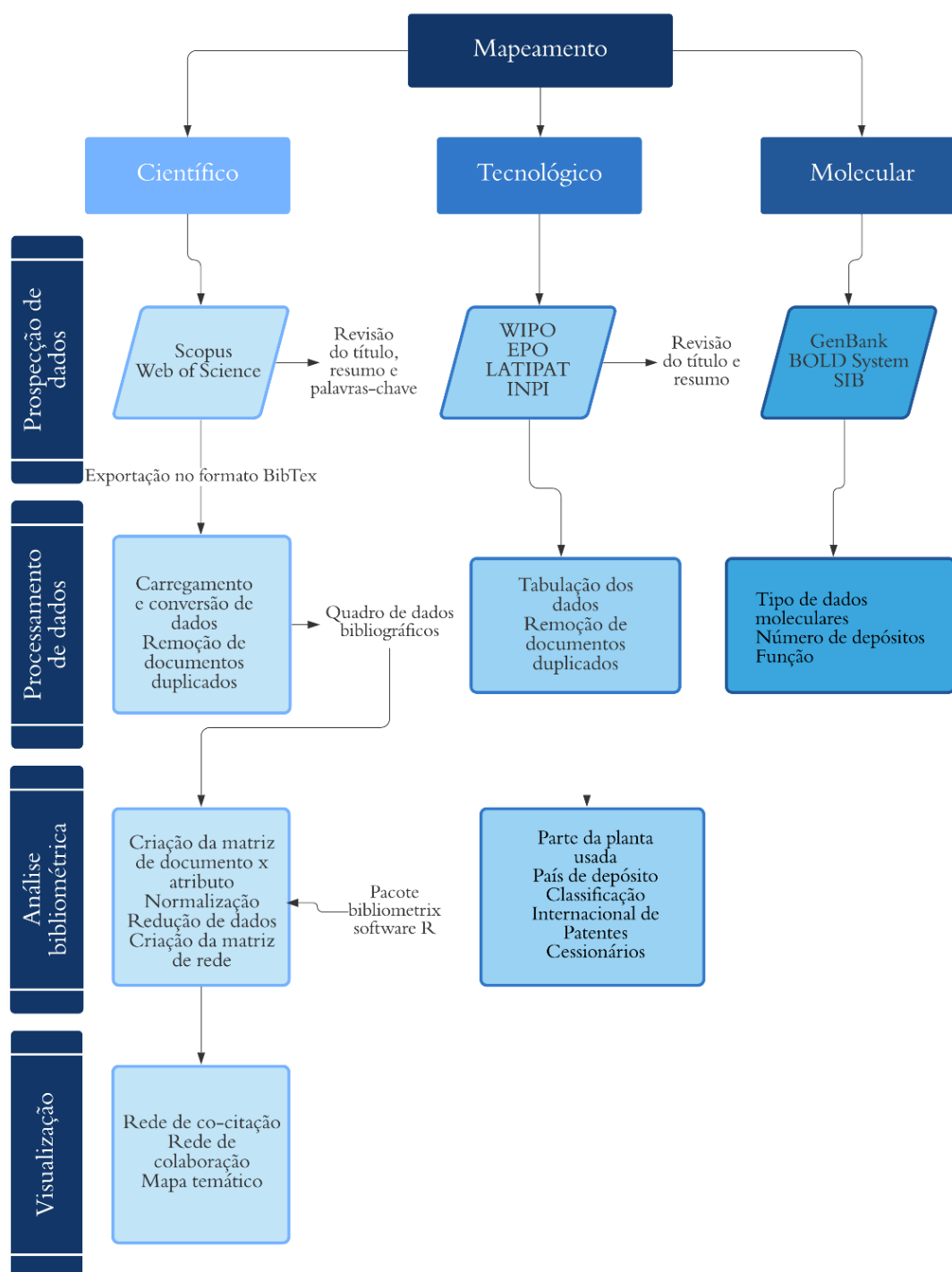


Figura 1. Fluxograma do processo de mapeamento científico, tecnológico e de dados moleculares para *Hancornia speciosa* Gomes.

4.2.1. Mapeamento científico

Os artigos foram prospectados nas bases científicas *Scopus* (<http://www.scopus.com>) e *Web of Science* (<http://www.webofknowledge.com>). Foram revisados pelo título e resumo e retirados aqueles que não se enquadraram na pesquisa. Metadados referentes às publicações científicas obtidos para cada termo nas duas bases foram exportados no formato BibTex, removidos os arquivos duplicados, combinados como um único conjunto de dados e analisados com o auxílio do pacote *bibliometrix* do software R v. 4.0.2 (Aria e Cuccurullo 2017).

A produção científica anual foi avaliada por meio do número de artigos sobre as espécies que foram publicados por ano, juntamente com o índice H, que mede a produção da pesquisa quanto ao impacto das citações, considerando também o conjunto das obras mais citadas e o número de citações que receberam (Aria et al. 2020).

A estrutura intelectual foi avaliada pela análise de cocitação que usa contagens de cocitação para construir medidas de similaridade entre os documentos, autores ou periódicos (Kang et al. 2020; McCain 1991). É possível medir a relação entre dois artigos com base no número de publicações em que aparecem citados concomitantemente, e isso permite estimar o impacto e a influência dos diferentes autores no campo analisado (Aria et al. 2020).

A estrutura social foi avaliada pelo mapa mundial de colaborações, para o qual foi considerado pelo menos uma colaboração entre autores de diferentes países (Aria et al. 2020).

A estrutura conceitual foi avaliada por meio de mapeamento temático, que permite a visualização de quatro tipos diferentes de temas com base na densidade e centralidade. A densidade é a força dos links internos entre todas as palavras-chave que são utilizadas para descrever o tema de pesquisa; a centralidade é a força dos links externos com outros temas, explorando assim o campo de palavras-chave dos autores (Aria e Cuccurullo 2017; Mumu et al. 2020).

4.2.2. Mapeamento tecnológico

As patentes foram prospectadas na *World Intellectual Property Organization* - WIPO (<https://www.wipo.int/portal/en/index.html>), no *European Patent Office* - EPO (<https://www.epo.org/>), *Espacenet-Latipat* (https://lp.espacenet.com/?locale=pt_LP) e Instituto Nacional da Propriedade Industrial - INPI (<https://www.gov.br/inpi/pt-br>).

Os documentos foram revisados e retirados aqueles que não se enquadraram na pesquisa ou que estavam duplicados. As variáveis avaliadas foram: parte da planta utilizada, países de depósito, classificação internacional e cessionários.

4.2.3 Mapeamento de dados moleculares

As sequências de ácidos nucleicos foram recuperadas de bancos públicos de dados moleculares: *Genetic Sequence Database* - *GenBank* (<http://www.ncbi.nlm.nih.gov/genbank>), *Barcode of Life Data System* - *BOLD System* (<http://www.barcodinglife.org>) e *Swiss Institute of Bioinformatics* - *SIB* (<https://www.sib.swiss>), e avaliadas quanto ao tipo de dado molecular e números de depósitos.

4.3. Resultados

4.3.1. Mapeamento científico

Foram obtidos 407 documentos para a espécie *H. speciosa* indexados nas bases de dados *Web of Science* (189) e *Scopus* (218). Após o processamento dos dados e exclusão dos duplicados, restaram para análise 303 documentos publicados entre 1992 e 2020. As publicações sobre *H. speciosa* apresentaram uma taxa de crescimento anual de 9,28%. Em 2009, foram 35 artigos e, nos anos subsequentes, o número de publicações por ano variou de 11 em 2013 a 28 em 2017.

Os documentos foram distribuídos em 129 fontes (periódicos), e um terço do total dos artigos foram publicados em sete periódicos, a saber, Revista Brasileira de Fruticultura (30), Ciência e Agrotecnologia (18), Pesquisa Agropecuária Brasileira (16), *Bioscience Journal* (12), Ciência e Tecnologia de Alimentos (8), Ciência Florestal (8) e *Journal of Ethnopharmacology* (8).

Os sete periódicos, responsáveis por um terço de todas as publicações, também estão entre aqueles com maior índice H, com valores igual a 9 para os periódicos Revista Brasileira de Fruticultura e Ciência e Agrotecnologia; 8 para a Pesquisa Agropecuária Brasileira e o *Journal of Ethnopharmacology*; 5 para *Bioscience Journal*; 4 para os periódicos Ciência Florestal e Ciência e Tecnologia de Alimentos.

Os artigos foram escritos por 971 autores, com valor médio de 0,319 documentos por autor. 99,69% dos documentos possuíam autoria múltipla e 0,31% autoria única. Em relação à colaboração dos autores, houve 3,13 autores por documento, 5,09 coautores por documento e um índice de colaboração igual a 3,16. A estrutura intelectual das publicações sobre *H. speciosa* foi avaliada pela rede de cocitação de documentos (Figura 2).

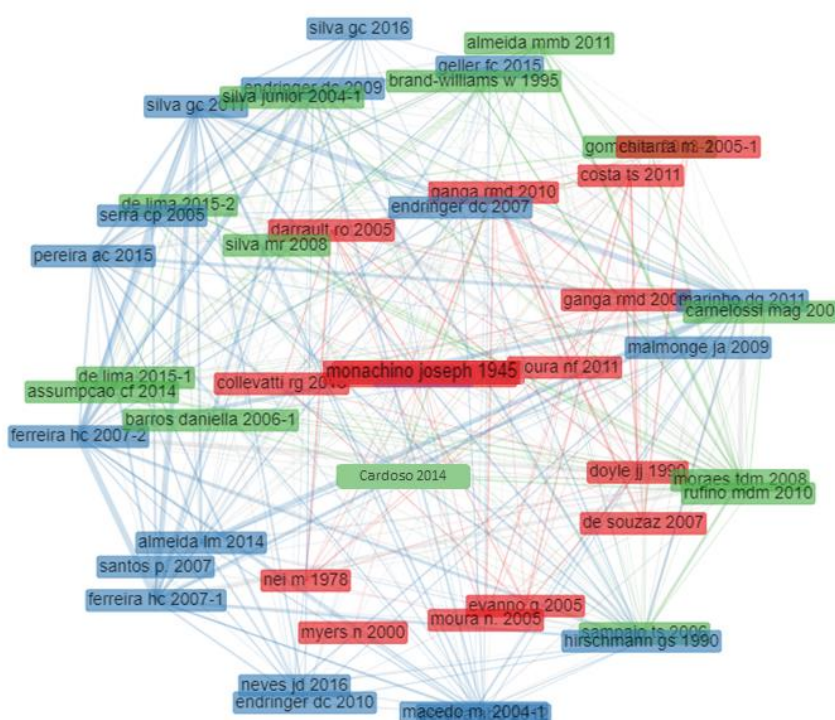


Figura 2. Rede de cocitação de documentos relacionados com *Hancornia speciosa* Gomes.

Nos estudos sobre *H. speciosa*, foram identificados 3 grupos distintos representados pelas cores azul, verde e vermelho. Artigos intimamente relacionados têm características semelhantes, desta forma um grupo com fortes relações de cocitação pode ser considerado uma área temática de pesquisa (Hjørland 2013).

A estrutura intelectual corresponde à base de conhecimento de um tópico ou campo de pesquisa. Em estudos com *H. speciosa* verificou-se a formação de três campos de pesquisa principais, sendo o primeiro (azul – Figura 2) formado por trabalhos que caracterizaram e investigaram o potencial biotecnológico do látex (Almeida et al. 2014; Floriano et al. 2016; Marinho et al. 2011) e folhas (Ferreira et al. 2007a, b; Pereira et al. 2015; Silva et al. 2011) de *H. speciosa*.

O segundo campo de pesquisa reúne trabalhos sobre o fruto, a mangaba (verde - Figura 2). Nestes trabalhos, a composição e as características da polpa (Cardoso et al. 2014; De Lima et al. 2015b), os compostos bioativos e o potencial antioxidante (De Lima et al. 2015a; Rufino et al. 2010) foram estudados.

Por fim, tem-se o terceiro campo de pesquisa (vermelho – Figura 2) formado por estudos voltados para a caracterização morfológica e diversidade genética da espécie, a exemplo, Monachino (1945) que, considerando diferenças morfológicas de folhas, frutos e flores, indicou a existências de seis variações botânicas; estudos sobre a variabilidade morfológica de árvores e frutos (Ganga et al. 2010); morfologia floral (Darrault e Schlindwein 2005) e diversidade e estrutura genética de populações naturais (Collevatti et al. 2016) e de acessos em Bancos de Germoplasma (Santos et al. 2017).

A estrutura social dos trabalhos sobre a espécie foi avaliada pelo mapa mundial de colaborações (Figura 3), para o qual considerou-se pelo menos uma colaboração entre autores de diferentes países. O Brasil figurou como o país com maior número de publicações, todas desenvolvidas por pesquisadores brasileiros. Em algumas publicações, ocorreu a colaboração entre o Brasil e países como Estados Unidos da América (5 publicações), Canadá (3 publicações), Alemanha (2 publicações), Portugal (2 publicações), Austrália (2 publicações), Itália (1 publicação), Espanha (1 publicação) e Colômbia (1 publicação).



Figura 3. Mapa mundial de colaboração para publicações sobre *Hancornia speciosa* Gomes.

Embora existam trabalhos sobre a espécie desenvolvidos pela colaboração internacional, verifica-se que estes são majoritariamente produzidos por pesquisadores de universidades do Brasil.

A estrutura social do trabalho com *H. speciosa* foi baseada em quatro comunidades científicas principais, sendo elas: 1) pesquisadores da Universidade Federal de Goiás, Universidade Estadual de Goiás, Pontifícia Universidade Católica de Goiás (PUC Goiás - instituição privada de ensino) e Instituto Federal de Educação, Ciência e Tecnologia de Goiás; 2) Universidade Federal de Sergipe que colabora com a Embrapa Tabuleiros Costeiros, Universidade Tiradentes (instituição privada de ensino) e Universidade Federal de Minas Gerais; 3) pesquisadores dessas duas comunidades (2 e 3) que desenvolveram atividades com pesquisadores da Universidade Estadual de Campinas e Universidade Estadual Paulista; 4) Universidade Federal de Lavras e Universidade Federal da Paraíba.

Nos trabalhos sobre *H. speciosa*, um total de 250 palavras-chave foram examinadas para criar o mapa temático (Figura 4), em que a frequência mínima do grupo foi 5 e o número de rótulos (para cada grupo) foi 1.

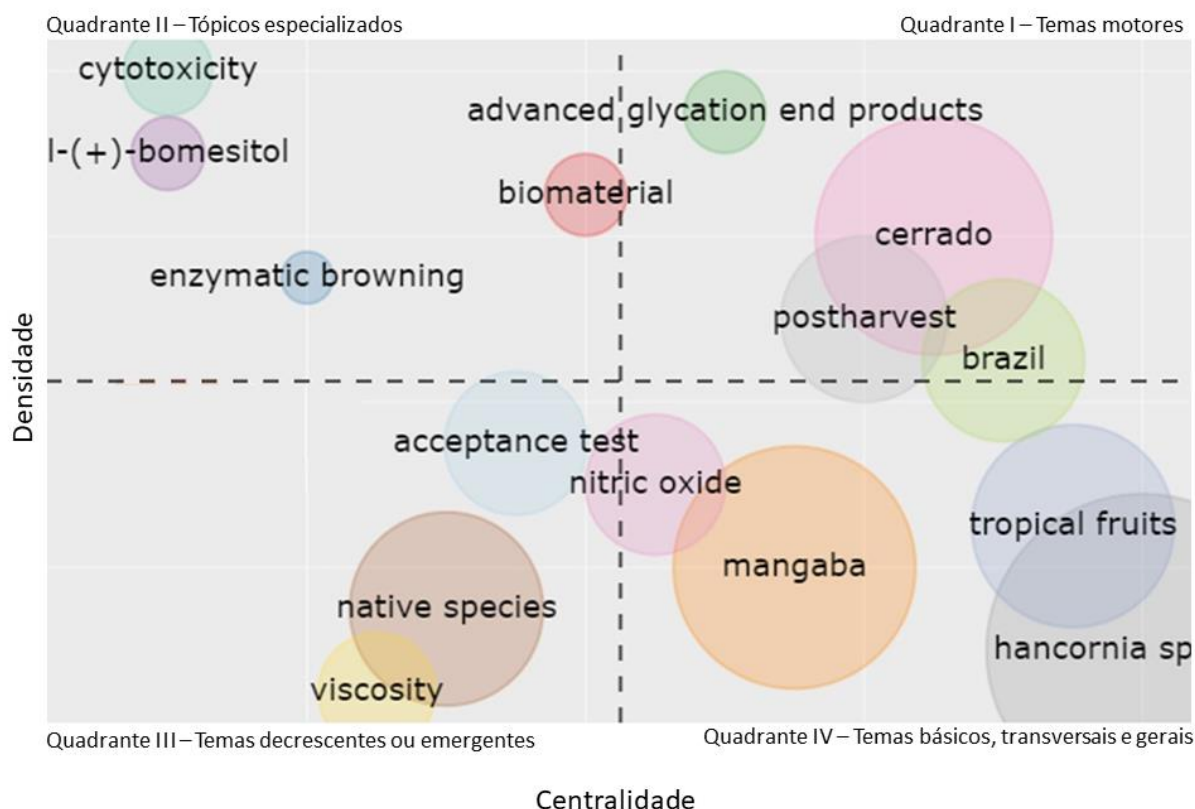


Figura 4. Mapa temático de publicações com *Hancornia speciosa* Gomes.

Ao considerar a centralidade e a densidade, o mapa temático se divide em quatro quadrantes: (i) - quadrante superior direito (alta centralidade, alta densidade) temas motores, são temas bem desenvolvidos e importantes para a estruturação de um campo de pesquisa; ii - quadrante superior esquerdo (alta centralidade, baixa densidade) temas especializados e caráter periférico com importância apenas marginal para o campo; iii - quadrante inferior esquerdo (baixa centralidade, baixa densidade) temas emergentes ou decrescentes, são desenvolvidos de maneira fraca e marginais; e iv - quadrante inferior direito (baixa centralidade, alta densidade) temas básicos, transversais e gerais, são importantes para um campo de pesquisa, mas não são bem desenvolvidos (Cobo et al. 2011).

Nos estudos com *H. speciosa*, os temas mais desenvolvidos na literatura (Figura 4) formaram 4 grupos: 1) Cerrado, que engloba palavras-chave relacionadas a diferentes estudos sobre a espécie, como diferentes variações botânicas, estudos etnobotânicos, etnofarmacologia e estudos com marcadores moleculares; 2) Pós-colheita - potencial alimentício de *H. speciosa*; 3) Brasil - regiões de ocorrência da espécie, frutas, propriedades físico-químicas, antioxidantes e vitaminas; e 4) Produtos finais da glicação avançada - reúne termos como xantina oxidase, hipoxantina, metil glioxal, glioxal e captura de espécies reativas de carbonila (*RCS-trapping*).

Os produtos finais da glicação avançada (AGEs) representam um conjunto de substâncias que contribuem diretamente para o desencadeamento e/ou agravamento de patologias associadas ao envelhecimento (Marques et al. 2018), e os termos agrupados nesse grupo estão presentes em publicações que investigam o potencial de extratos de *H. speciosa* para inibir o processo de glicação.

Os temas especializados e de caráter periférico (Figura 4) formaram os grupos: 1) citotoxicidade; 2) l-(+)-bomesitol; 3) escurecimento enzimático; e 4) biomaterial. Nos quais

foram agrupados termos relacionados a estudos fotoquímicos, atividade biológica e atividade citotóxica dos extratos (Santos et al. 2016; Ribeiro et al. 2012); e estudos citotóxicos com látex (Almeida et al. 2014; Ribeiro et al. 2016) e frutos (Assumpção et al. 2014) de *H. speciosa*.

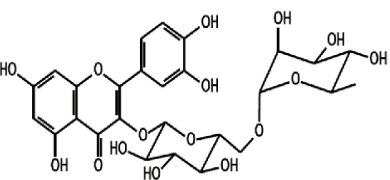
Foram agrupados também estudos que tratam sobre métodos (congelamento, secagem, pasteurização e liofilização) empregados na preservação de compostos nutricionais, bioativos e da atividade antioxidante da polpa (Paula et al. 2019; Zitha et al. 2020), que possui vida pós-colheita curta, característica que limita a comercialização da fruta *in natura*; estudos relacionados ao potencial do látex para uso como biomaterial na forma de biomembrana com atividades angiogênicas, anti-inflamatórias e antibacterianas (Almeida et al. 2019) e estudos relacionados à atividade antibiofilme, promissora para aplicações na área de regeneração de feridas cutâneas (Bonete et al. 2020).

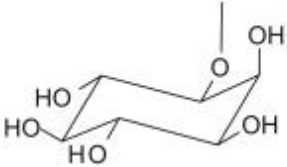
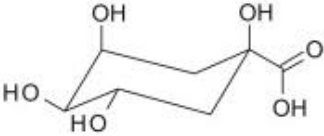
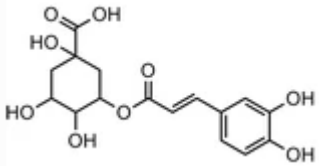
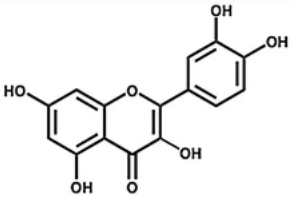
No quadrante inferior esquerdo estão os grupos teste de aceitação, viscosidade e espécies nativas que compõem os temas desenvolvidos de maneira fraca e marginais (Figura 4). Teste de aceitação e viscosidade agruparam temas relacionados à análise sensorial de produtos alimentícios desenvolvidos com *H. speciosa*, a exemplo, bebida láctea (Da Silva et al. 2015; Moura et al. 2016), sorvete (Santos e Silva, 2012) e geleia (Zitha et al. 2020). O grupo espécie nativa engloba temas variados, como germinação, sementes recalcitrantes, cultura de tecidos, conservação *in vitro*, conservação da diversidade.

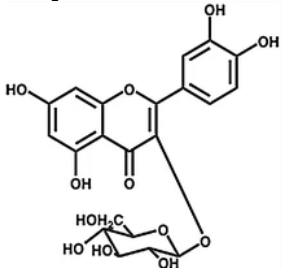
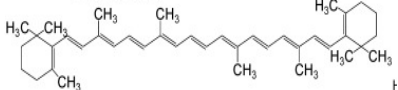
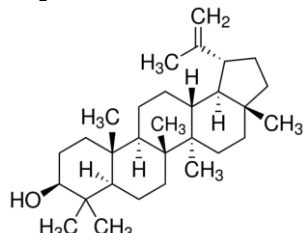
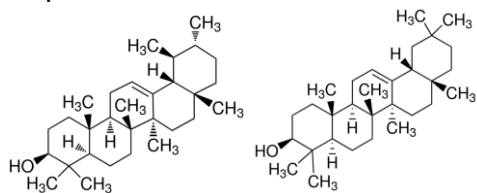
No mapa temático, os grupos óxido nítrico, mangaba, frutas tropicais e *Hancornia speciosa* (Figura 4) são temas importantes para o campo de pesquisa, mas não são bem desenvolvidos. O grupo óxido nítrico está relacionado aos estudos que investigam o efeito vasodilatador do extrato etanólico das folhas de *H. speciosa* por meio da ativação de um mecanismo dependente da produção de óxido nítrico (Ferreira et al. 2007a, b; Silva et al. 2011, 2016). Os demais grupos englobam termos variados que estão relacionados com populações tradicionais, extrativismo, etnofarmacologia, germinação, conservação, diversidade genética e outros compostos bioativos presentes na planta como a rutina, ácido quínico, ácido clorogênico e antocianinas.

H. speciosa possui muitos compostos bioativos (Quadro 1) e alguns deles possuem propriedades funcionais que explicam o potencial da espécie para uso no tratamento de doenças e o seu valor nutricional como alimento.

Quadro 1. Alguns dos compostos bioativos identificados em diferentes partes da planta *Hanconia speciosa* Gomes e suas propriedades funcionais.

Compostos bioativos	Propriedades funcionais	Parte da planta	Referência
Rutina 	Antidiabético - inibição da α -amilase e α -glucosidase Anticancer Anti-inflamatório Inibição da AChE e da BChE Anti-hipertensivo Inibidor promissor da protease principal e outros alvos proteicos do SARR-CoV2	Folhas Cascas Frutos	Silva et al. 2011; Perk et al. 2014; Pereira et al. 2015; Torres-Rêgo et al. 2016; Ghorbani 2017; Nikfarjam et al. 2017; Dos Santos et al. 2018; Al-Zahrani 2020; Das et al. 2020; Leite

			et al. 2020
l-(+)-bornesitol 	Inibidor da enzima de conversão da angiotensina Anti-hipertensivo Agente quimiopreventivo do câncer	Folhas	Endringer et al. 2009, 2014; Silva et al. 2011; Moreira et al. 2019
Ácido quínico 	Anti-inflamatório - inibição de NF-κB e da produção de TNF-α Inibição da α-glucosidase Quimiopreventivo Anti-hipertensivo	Folhas	Endringer et al. 2009; Pereira et al. 2011; Silva et al. 2011; Geller et al. 2015
Ácido cloregênico 	Anti-inflamatória	Frutos Folhas Látex	Torres-Rêgo et al. 2016
Catequina 	Anticâncer Antiobesidade Antidiabético - inibição da α-amilase e α-glucosidase Anticardiovascular Anti-infeccioso Hepatoprotetor Neuroprotetor Antioxidante Inibição da AChE e da BChE Promissor antiviral por impedir a formação do complexo S Proteína-ACE2 do SARS-CoV2	Folhas Casca Frutos Látex	Katalinić et al. 2004; Dos Santos et al. 2018; Isemura 2019; Leite et al. 2020; Jena et al. 2021

Isoquercetina 	Antioxidante Antidiabético - inibição da α-amilase e α-glucosidase	Folhas	Dos Santos et al. 2018; Leite et al. 2020
β-Caroteno 	Antioxidante, Antidiabético Antiobesidade	Frutos Folhas	Bailão et al. 2015; Dos Santos et al. 2018
Lupeol 	Antidiabético - inibição da α-glucosidase e da proteína tirosina fosfatase-1B	Folhas	Pereira et al. 2015
α / β-amirina 	Antidiabético	Folhas	Pereira et al. 2015

4.3.2 Mapeamento tecnológico

A prospecção de patentes baseadas em *H. speciosa* retornou um total de 112 documentos (WIPO - 33, EPO - 23, LATIPAT - 26 e INPE - 30). Com o refinamento dos dados e remoção de arquivos duplicados restaram 29 patentes para análise.

Os produtos tecnológicos possuem como origem geográfica o Brasil e para alguns também foram feitos depósitos em outros países, como Estados Unidos da América, Japão, Argentina, Canadá, China e em países que compõem a Organização Europeia de Patente e Organização Mundial da Propriedade Intelectual.

As patentes foram desenvolvidas a partir de diferentes partes da planta, sementes (Leite et al. 2018), extratos das folhas (Braga et al. 2009), látex (Barbosa et al. 2020; Neves et al. 2014, frutos (Barroso et al. 2018; Santana et al. 2016) e pela combinação de frutos e folhas (Bitencourt et al. 2014; Droppa-Almeida et al. 2019).

Produtos tecnológicos baseados em *H. speciosa* dividem-se em produtos alimentícios (55,17%), reagente para produção de compostos químicos (24,13%), medicamentos (17,24%) e cosméticos (3,44%). De acordo com a Classificação Internacional de Patentes (WIPO 2020), estes produtos estão inseridos na área de Necessidades Humanas; Operações, Transporte, Separação, Mistura e Química, Metalurgia (Tabela 1).

Tabela 1. Classificação Internacional de Patentes para produtos tecnológicos desenvolvidas com *Hancornia speciosa* Gomes.

Classificação Internacional de Patentes	Símbolo
Necessidades Humanas	Agricultura; silvicultura; criação animal; caça; armadilhas; pescaria
	A01N65/08, A01P1/00
	Cozimento; equipamento para fazer ou processar massas; massas para assar
	A21D 13/08
Necessidades Humanas	Alimentos ou gêneros alimentícios; seu tratamento não coberto por outras classes
	A23G 3/48, A23G3/54, A23B7/024, A23L 2/00, A23L 3/00, A23L 19/00, A23L 1/064, A23B 7/00, 23B7/0205
	Ciências médicas ou veterinárias; higiene
	A61L 27/00, A61K 36/24, A61K, 36/00, A61K127/00, A61K131/00, A61K 135/00, A61K7/48, A61K38/16, A61P31/04, A61P31/04, A61L 27/54, A61P 19/08
Executando operações; transportando; separando; mistura	Processos físicos ou químicos, ou aparelhos em geral
	B01D 71/24
	Pulverização ou atomização em geral; aplicação de líquidos ou outros materiais fluentes em superfícies em geral
	B05D 5/00
Executando operações; transportando; separando; mistura	Eliminação de resíduos sólidos; recuperação de solo contaminado
	B09B 3/00
	Trabalho de plásticos; trabalho de substâncias em um estado plástico em geral
	B29C 35/00
Química; Metalurgia	Tratamento de água, águas residuais, esgoto ou lama
	C02F 1/00, C02F 101/00
	Química orgânica
	C07K14/00
	Compostos macromoleculares orgânicos; sua preparação ou processamento químico; composições baseadas nisso
	C08B 30/14, C08K 5/053, C08L 3/02
Química; Metalurgia	Bioquímica; Cerveja; aguardente; vinho; vinagre; microbiologia; enzimologia; mutação ou engenharia genética
	C12P 1/00, C12R 1/00, C12N 9/00, C12C 5/00, C12C 12/00, C12G 3/04

As instituições cessionárias de patentes desenvolvidas com *H. speciosa* são em sua maioria instituições de ensino superior públicas que estão distribuídas em diferentes regiões do país. A maior parte delas está localizada na região Nordeste. Sergipe é o estado brasileiro com o maior número de patentes que foram desenvolvidas por pesquisadores vinculados à Secretaria de Estado da Inclusão, Assistência e do Desenvolvimento Social (9 patentes), Universidade Federal de Sergipe (4 patentes) e Instituto de Tecnologia e Pesquisa da Universidade Tiradentes (4 patentes).

As demais instituições cessionárias localizadas na região Nordeste são a Universidade Federal de Campina Grande e a Universidade Federal da Paraíba (2 patentes), a Universidade Federal do Rio Grande do Norte, Universidade Federal de Pernambuco e a Universidade Federal do Ceará (1 patente).

Na região Sudeste, são instituições cessionárias a Universidade Federal de Minas Gerais e Universidade Estadual de Campinas com 1 patente cada e a empresa Pele Nova Biotecnologia, localizada no estado de São Paulo com 3 patentes.

4.3.3 Mapeamento de dados moleculares

A busca por depósitos de sequências de *H. speciosa* em bancos públicos de dados moleculares resultou em um total de 164 sequências de nucleotídeos e 236 de proteínas, obtidas no *Genetic Sequence Database* (GenBank). Esse banco faz parte do *National Center for Biotechnology Information* (NCBI).

No *Barcode of Life Data System (BOLD System)* foram recuperadas duas sequências de nucleotídeos dos genes da subunidade maior da Rubisco (ribulose-1,5-bisfosfato) - *rbcL* e maturase k - *matK* (Figura 5).



Figura 5. Códigos de barras de DNA *rbcL* (A) e *matK* (B) depositados no *Barcode of Life Data System* para a espécie *Hancornia speciosa* Gomes. Fonte: *BOLD System*, 2020.

As sequências depositadas para *H. speciosa* são do genoma cloroplastidial (cpDNA) e estão relacionadas a proteínas que desempenham funções na fotossíntese, metabolismo, transcrição, proteínas ribossômicas, controle de qualidade de proteínas e montagem, inserção de membrana. Para algumas delas foram verificados depósitos da estrutura tridimensional no banco de dados do *Swiss Institute of Bioinformatics* (SIB) (Figura 6).

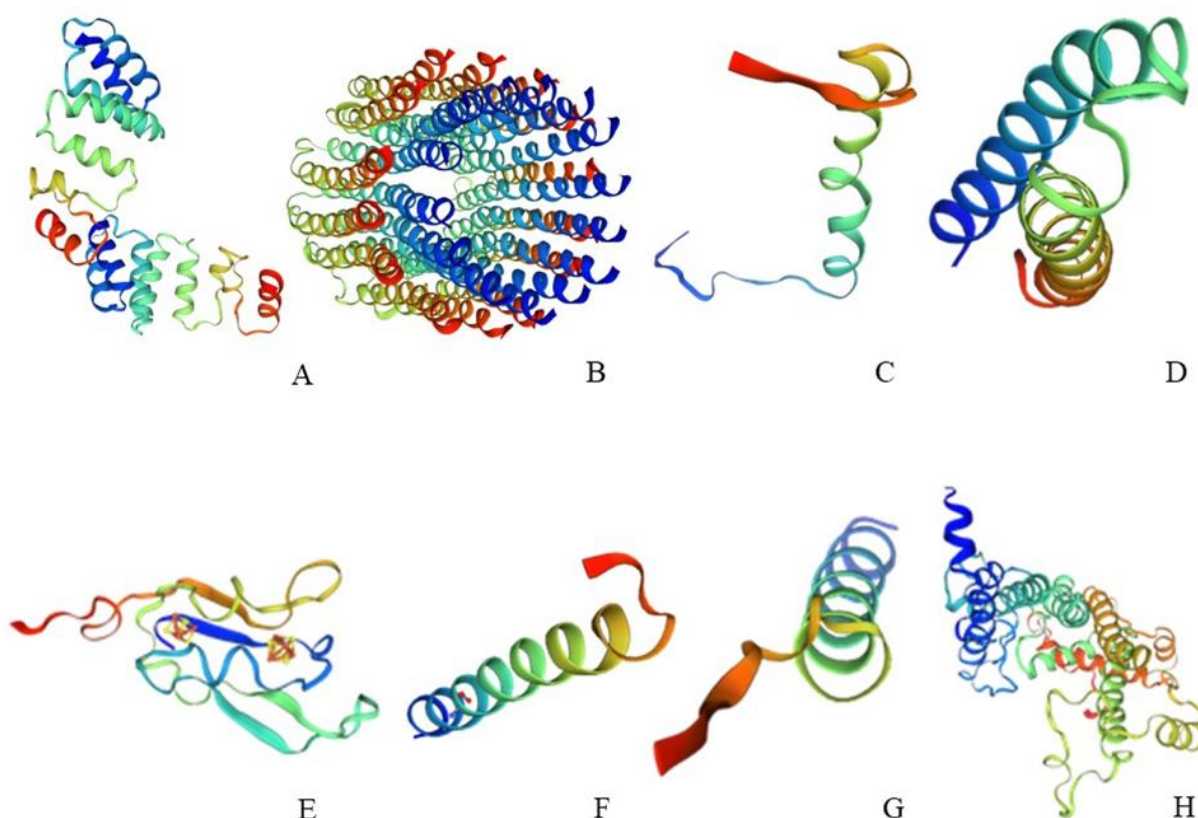


Figura 6. Estrutura tridimensional de proteínas modeladas por homologia para *Hancornia speciosa* Gomes. A. Proteína de montagem do fotossistema I; B. Subunidade c da ATP sintase. C. Subunidade 9 do centro de reação do fotossistema I; D. Proteína Z do centro de reação; E. Centro ferro-enzofre do fotossistema I; F. Proteína T do centro de reação do fotossistema II; G. Subunidade 6 do citocromo b6-f; H. Proteína D do fotossistema II. Fonte: SIB, 2020.

4.4. Discussão

A bioeconomia surge como parte da solução de desafios globais, como alimentar uma população crescente, mitigar os impactos ambientais e possibilitar o desenvolvimento econômico. Em países megadiversos, as estratégias e políticas de bioeconomia são pautadas no uso da biodiversidade e, no Brasil, *H. speciosa* é um importante ativo para a bioeconomia por possibilitar o fornecimento de alimentos com qualidade nutricional e compostos bioativos para medicamentos. Além disso, a espécie se desenvolve em solos arenosos ou areno-argilosos, ácidos, pobres em nutrientes e matéria orgânica, com baixa retenção de água, acidificados e pedregosos, considerados indesejados para fins agrícolas (Pereira et al. 2016; Da Silva Júnior et al. 2018a).

O número de publicações científicas desenvolvidas com *H. speciosa* vem crescendo e a tendência é que continue aumentando, tendo em vista que a espécie contribui para os Indicadores Nutricionais de Biodiversidade de acordo com a Rede Internacional de Sistemas de Dados de Alimentos (INFOODS/FAO) e faz parte da flora brasileira com valor econômico atual ou potencial (Coradin et al. 2018; FAO 2021).

As publicações sobre *H. speciosa* são majoritariamente desenvolvidas por instituições de pesquisas localizadas no Brasil que estão distribuídas em diferentes regiões, o que reflete a ampla ocorrência da espécie no país, presente em diferentes biomas e fitofisionomias. Na região Nordeste, ocorre naturalmente em áreas de restinga no litoral dos estados de Sergipe, Alagoas, Pernambuco, Paraíba, Rio Grande do Norte e Ceará (Martins et al. 2012; Santos et al. 2017; Silva et al. 2017a); nas regiões Sudeste, Centro-Oeste, no estado de Tocantins e sul

da Bahia ocorre no bioma Cerrado sob as fitofisionomias de cerradão, cerrado sentido restrito, campo sujo ou campo rupestre (Ganga et al. 2010; Nascimento et al. 2014).

Algumas das universidades que formam a rede de colaboração entre instituições, também figuram como cessionárias de patentes desenvolvidas a partir de *H. speciosa*. Fato que indica a ocorrência de transferência do conhecimento científico gerado nessas instituições em produtos tecnológicos, ainda que com pouca frequência, considerando a discrepância entre o número de artigos desenvolvidos e patentes geradas. A conversão do conhecimento científico em produtos tecnológicos é um aspecto positivo e precisa ser incentivado, levando em conta que a inovação e obtenção de produtos tecnológicos são fatores importantes para o desenvolvimento do país.

Os temas mais desenvolvidos sobre a espécie estão relacionados ao uso como alimento e como planta medicinal, os estudos foram realizados em populações naturais localizadas nas diferentes regiões do país em que a planta é conhecida e utilizada principalmente por comunidades extrativistas no tratamento de doenças e como alimento. Os produtos tecnológicos baseados na espécie também seguem essa tendência e a maior parte apresenta relação direta com o potencial como alimento, a partir do beneficiamento dos frutos e desenvolvimento de produtos como licores, bolos e geleias.

Estudos de diversidade genética estimada por marcadores moleculares também estão entre os temas mais desenvolvidos. Numerosos são os estudos nesse sentido e empregam marcadores como RAPD (Silva et al. 2013), ISSR (Da Costa et al. 2015) e SSR (Silva et al. 2019) para avaliar populações naturais (Silva et al. 2017b) e bancos de germoplasma (Silva et al. 2019). Estudos considerando populações amplamente distribuídas no país e que investigam as diferenças moleculares entre populações e variedades botânicas da espécie são menos frequentes (Chaves et al. 2020; Collevatti et al. 2016, 2018).

Espécies da família Apocynaceae são relatadas como fonte de metabólitos secundários usados para fins terapêuticos com potencial anticarcinogênico, estimulante cardíaco e anti-hipertensivo. Entretanto, *H. speciosa* não está entre as espécies estudadas e o perfil fitoquímico e os constituintes responsáveis por sua bioatividade não foram completamente elucidados (Süntar 2020).

Estudos que visam caracterizar os compostos bioativos presentes no látex, folhas, cascas e frutos da espécie, bem como a atividade funcional desses compostos, são considerados temas especializados e de caráter periférico. Entre os principais compostos bioativos identificados para a espécie estão a rutina, l-(+)-bornesitol, ácido quínico, ácido clorogênico, catequinas, isoquercetinas, β -Caroteno, Lupeol e α / β -amirina.

A rutina é um flavonol abundantemente encontrado em plantas (Ganeshpurkar e Saluja 2017), está presente nas folhas, cascas e frutos de *H. speciosa* (Leite et al. 2020) e possui diversas aplicações farmacológicas como antidiabético (Ghorbani 2017), atividade anticâncer (Perk et al. 2014), anti-inflamatório (Nikfarjam et al. 2017) e mais recente como um inibidor promissor da protease principal e outros alvos proteicos do SARS-CoV2 (Al-Zahrani 2020; Das et al. 2020).

O ácido quínico é um ciclitol que está presente em várias espécies vegetais. É também um dos principais constituintes das folhas de *H. speciosa* e está relacionado à atividade anti-inflamatória pela inibição do fator de transcrição pró-inflamatório (NF- κ B) e da produção do fator de necrose tumoral- α (TNF- α) (Geller et al. 2015).

O l-(+)-bornesitol trata-se do principal constituinte do extrato de folhas de *H. speciosa* (Endringer et al. 2014), é um potente inibidor da enzima de conversão da angiotensina (Moreira et al. 2020) e, portanto, associado ao efeito anti-hipertensivo da espécie (Moreira et al. 2019). O ciclitol também foi citado como agente quimiopreventivo do câncer (Endringer et al. 2009).

O ácido clorogênico, que está associado à atividade antioxidante, está presente nas folhas, frutos e látex (Torres-Rêgo et al. 2016). As catequinas estão presentes nas folhas, frutos, látex e na casca, possuem inúmeras aplicações farmacológicas (Dos Santos et al. 2018;

Isemura 2019; Katalinić et al. 2004; Leite et al. 2020), além de ser um promissor antiviral por impedir a formação do complexo S Proteína-ACE2 do SARS-CoV2 (Jena et al. 2021).

As isoquercetinas, lupeol e α/β -amirina foram relatados para as folhas e estão relacionadas à propriedade antidiabética relatada para a planta (Pereira et al. 2015). O β -caroteno relatado principalmente para frutos, também está presente nas folhas e foi associado às propriedades antioxidante, antidiabético e antiobesidade (Bailão et al. 2015; Dos Santos et al. 2018).

A presença desses compostos bioativos também foi considerada para o desenvolvimento de patentes voltadas para uso farmacológico, produção de reagentes químicos e cosméticos, baseados principalmente nas propriedades do látex e extraídos das folhas.

Vários estudos apontam o potencial farmacológico do látex de *H. speciosa*, entretanto, os constituintes que conferem esses efeitos funcionais ainda não estão completamente elucidados e estão entre os temas importantes para o campo de pesquisa, mas que não foram bem desenvolvidos. Alguns dos constituintes do látex foram identificados, sendo eles o ácido clorogênico, naringenina-7-O-glucosídeo, catequina e procianidina (dos Santos et al. 2016). O mesmo ocorre para os demais compostos bioativos presentes em outras partes da planta para os quais não se tem conhecimento da forma de ação.

O beneficiamento dos frutos, a propagação e conservação da espécie foram considerados temas pouco desenvolvidos. No Brasil, o beneficiamento dos frutos é feito de forma artesanal por comunidades extrativistas que usam a polpa para a produção de doces, geleias, polpas, e em menor proporção pela agroindústria para a produção de polpas e sorvetes nas regiões Centro-Oeste e Nordeste do país. Os frutos são considerados alimentos funcionais e os resultados obtidos por Araújo et al. (2019) sugerem benefícios do consumo por mulheres para prevenção de risco para câncer de mama.

A propagação de *H. speciosa* ocorre principalmente por sementes que são recalcitrantes - sementes que possuem baixa longevidade e perdem rapidamente a capacidade de germinar, por esse motivo é necessário que a semeadura ou uso na produção de mudas ocorra logo após o beneficiamento. A propagação dessa espécie pode ocorrer também *in vitro* a partir de explantes, a exemplo, segmentos internodais (Prudente et al. 2016), contudo, essa forma de propagação é normalmente empregada apenas em bancos ativos de germoplasma.

O fato de *H. speciosa* possuir sementes recalcitrantes tem implicação direta na sua conservação, que ocorre por meio de coleções em campo ou populações naturais em unidades de conservação. Como alternativa à conservação *in situ*, pesquisadores buscam desenvolver protocolos para a criopreservação de explantes (Prudente et al. 2017). Maiores informações são necessárias para a otimização desses protocolos, bem como sobre os substratos adequados para a propagação, espaçamento de plantio, trados culturais e manejo, para que seja possível aproveitar as potencialidades oferecidas pela espécie de forma sustentável.

H. speciosa se encontra em fase de domesticação e os programas de melhoramento ainda estão em desenvolvimento. Embora existam descritores morfológicos estabelecidos para *H. speciosa* (Da Silva Júnior et al. 2018b), não existem cultivares registradas (proteção *sui generis*). Também não foram verificadas tecnologias baseadas em plantas geneticamente modificadas ou genes engenheirados relacionados à espécie e protegidos pelo patenteamento. Em bancos públicos de dados moleculares, informações sobre o genoma nuclear não estão disponíveis até o momento. Nos bancos públicos estão disponíveis somente sequências de nucleotídeos e proteínas que pertencem ao genoma cloroplastidial (cpDNA). Entre essas, sequências de nucleotídeos dos genes *rbcL* e *matk* que funcionam como Códigos de barra de DNA (*DNA barcodes*) são encontradas.

O *DNA barcodes* é um método taxonômico para identificar espécies e tem como objetivo utilizar a diversidade existente dentro do DNA dos seres vivos para discriminar espécies empregando uma sequência de DNA curta (400-800 pb), universal (presente em

todas as espécies), variável (com diferenciação suficiente entre espécies para ser possível distingui-las) e facilmente isolada (Hebert et al. 2003).

Para a maioria das plantas terrestres o cpDNA é conservado e as regiões codificadoras contêm basicamente dois grupos de genes, o primeiro grupo compreende componentes para a maquinaria fotossintética, especificamente para os fotossistemas I e II, o complexo citocromo b6/f e o complexo ATP sintase; o segundo grupo inclui os genes necessários para o sistema genético dos plastídios - subunidades de uma RNA polimerase, rRNAs, tRNAs e proteínas ribossômicas, sendo os genes 16S, 23S, 5S e 27-31 de tRNA suficientes para traduzir todos os aminoácidos, incluindo três genes para a subunidade da RNA (Green 2011; Huang et al. 2013).

O cpDNA é recomendado para estudos filogenéticos por possuir regiões com baixa taxa evolutiva e, conseqüentemente, maior conservação observada entre os genomas em evolução conhecidos, sendo úteis para elucidar questões filogenéticas entre famílias e taxa mais altos. Ao mesmo tempo possui regiões não codificantes e que evoluíram o suficiente para a resolução filogenética em nível de espécie.

Além das características citadas, o cpDNA tem herança uniparental (haploide), ou seja, para a maioria das espécies esse genoma é herdado de um dos pais. Esse fato permite a geração de inferências sobre as contribuições relativas do fluxo de sementes e pólen para a estrutura genética de populações naturais, comparando marcadores nucleares e plastidiais, além de inferências filogenéticas que podem ser feitas em nível intraespecífico (Delplancke et al. 2012; Khadivi-Khub et al. 2014). Cabe destacar que para a maioria das espécies a herança materna é mais difundida do que a herança paterna (Hagemann 2002), com raros casos de herança biparental ou paterna (Zhang e Sodmergen 2010).

Outro aspecto relevante é o tamanho do cpDNA, considerado pequeno quando comparado com o genoma nuclear, dependendo da espécie possui de 120-200 kpb de comprimento e apresenta um comprimento de contorno de cerca de 30 a 60 micrômetros e uma massa de cerca de 80 a 130 milhões de daltons (Shaw et al. 2007). Sua estrutura é simples, conforme destacado anteriormente. Estas características facilitam o desenvolvimento de estudos e o seu sequenciamento, o que pode justificar o maior número de depósitos nos bancos públicos.

Nos dados moleculares que estão disponíveis para *H. speciosa*, foram verificadas estruturas tridimensionais para proteínas do cpDNA obtidas pela modelagem molecular por homologia.

A modelagem por homologia de proteínas é uma ferramenta bem-sucedida de predição de estruturas tridimensionais de proteínas que se baseia na evolução molecular. Partindo do princípio de que a semelhança entre as estruturas primárias desta proteína e de proteínas homólogas de estruturas tridimensionais conhecidas (proteínas-molde) implica em similaridade estrutural entre elas (Muhammed e Aki-Yalcin 2019). Entre suas etapas tem-se o alinhamento da sequência-alvo ou sequência-problema (proteína desconhecida que se quer modelar) com a sequência template ou molde (proteínas homólogas de estruturas tridimensionais conhecidas) (Santos Filho e Alencastro 2003).

A predição da estrutura e função de proteínas é fundamental para determinar a contribuição individual destas para o processo vital de um organismo. Sendo uma ferramenta valiosa, sobretudo para melhorar a compreensão dos mecanismos de reações químicas, estrutura e função, principalmente de proteínas difíceis de serem purificadas em grande escala ou de espécies para as quais existe um volume menor de estudos e informações disponíveis, a exemplo *H. speciosa*.

A disponibilidade de dados moleculares em bancos públicos permite que estes sejam utilizados para o desenvolvimento de estudos *in silico* (Bezerra et al. 2018), para prospecção de genes de resistência a doenças (Debikakas et al. 2014), pragas (Joukhadar et al. 2013), tolerância a diferentes estresses abióticos (Abhayawickrama et al. 2020; Das et al. 2019; Tripathi et al. 2020), análise filogenética (Bauwens et al. 2018; Karakulah e Pavlopoulou

2018) e desenho de primers para estudos de diversidade (Feng et al. 2016; Thatikunta et al. 2016) e expressão gênica (Alonso et al. 2018; Bustin e Huggett 2017).

4.5. Conclusões

H. speciosa figura como uma fonte importante de alimento com alta qualidade nutricional, compostos para aplicação biotecnológica e possui papel bioeconômico relevante. Entretanto, existem lacunas do conhecimento que limitam o melhor aproveitamento e põem em risco a conservação da espécie.

Existe uma vasta literatura sobre a ação anti-inflamatória, angiogênica, osteogênica, efeito antidiabético e propriedades antioxidantes relatadas para diferentes partes da planta. Entretanto, os constituintes que conferem essas atividades funcionais, bem como o modo de ação deles, ainda não estão completamente elucidados. Os genes e rotas metabólicas que estão relacionados à produção dos compostos bioativos em *H. speciosa* também não são conhecidos.

As informações disponíveis nos bancos públicos de dados moleculares são limitadas ao genoma plastidial. Poucos produtos desenvolvidos a partir da espécie foram patenteados e não existem inovações baseadas em biotecnologia e cultivar registrada para *H. speciosa*.

As pesquisas futuras com *H. speciosa* podem direcionar o foco sobretudo para quatro áreas: i) manejo (produção de frutos, colheita e pós-colheita dos frutos); ii) conservação (investigar a diferenciação genética entre populações amplamente distribuídas e monitorar a diversidade genética); iii) domesticação e melhoramento (identificar genótipos de interesse e registro de cultivares); iv) potencial farmacológico (investigar genes, rotas metabólicas e o desenvolvimento de bioprodutos).

4.6. Referências Bibliográficas

- Abhayawickrama B, Gimhani D, Kottearachchi N et al (2020) In silico identification of QTL-based polymorphic genes as salt-responsive potential candidates through mapping with two reference genomes in rice. *Plants*. <https://doi.org/10.3390/plants9020233>
- Almeida LM, Floriano JF, Ribeiro TP et al (2014) *Hancornia speciosa* latex for biomedical applications: physical and chemical properties, biocompatibility assessment and angiogenic activity. *J Mater Sci Mater Med* 25:2153–2162. <https://doi.org/10.1007/s10856-014-5255-8>
- Almeida LM, Magno LN, Pereira AC et al (2019) Toxicity of silver nanoparticles released by *Hancornia speciosa* (Mangabeira) biomembrane. *Spectrochim Acta Part A Mol Biomol Spectrosc* 210:329–334. <https://doi.org/10.1016/j.saa.2018.11.050>
- Alonso GC, Pavarina AC, Sousa TV, Klein MI (2018) A quest to find good primers for gene expression analysis of *Candida albicans* from clinical samples. *J Microbiol Methods* 147:1–13. <https://doi.org/10.1016/j.mimet.2018.02.010>
- Al-Zahrani AA (2020) Rutin as a promising inhibitor of main protease and other protein targets of COVID-19. In *Silico Study Nat Prod Commun*. <https://doi.org/10.1177/1934578X20953951>
- Araújo RL, Savazzi S, Fujimori M et al (2019) Effects of mangaba (*Hancornia speciosa*) Fruit extract adsorbed onto PEG microspheres in MCF-7 breast cancer cells co-cultured with blood cells. *Asian Pac J Cancer Prev* 20:1995–2001. <https://doi.org/10.31557/APJCP.2019.20.7.1995>
- Aria M, Cuccurullo C (2017) bibliometrix: an R-tool for comprehensive science mapping analysis. *J Informetr* 11:959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Aria M, Misuraca M, Spano M (2020) Mapping the evolution of social research and data science on 30 years of social indicators research. *Soc Indic Res* 149:803–831. <https://doi.org/10.1007/s11205-020-02281-3>

- Assumpção CF, Bachiega P, Morzelle MC et al (2014) Characterization, antioxidant potential and cytotoxic study of mangaba fruits. *Ciência Rural* 44:1297–1303. <https://doi.org/10.1590/0103-8478cr20130855>
- Bailão E, Devilla I, da Conceição E, Borges L (2015) Bioactive compounds found in Brazilian Cerrado fruits. *Int J Mol Sci* 16:23760–23783. <https://doi.org/10.3390/ijms161023760>
- Ball R (2017) An introduction to bibliometrics: new development and trends. Elsevier, Cambridge
- Barbosa AM, Dariva C, Padilha FF, et al (2020) Processo de obtenção de um produto contendo compostos bioativos com ação antibacteriana contra bactérias multirresistentes. BR102018076511A2. Deposit: Dez. 19 2018.
- Barroso AJR, Almeida F de AC, Silva LM de M (2018) Liofilização de mangaba para preparação de suco e uso como ingrediente na formulação de produtos alimentícios. BR102016027709. Deposit: Nov. 16 2016.
- Bauwens E, Joosten M, Taganna J et al (2018) In silico proteomic and phylogenetic analysis of the outer membrane protein repertoire of gastric *Helicobacter* species. *Sci Rep* 8:15453. <https://doi.org/10.1038/s41598-018-32476-1>
- Bezerra LCC, Queiroz EWA de, Freire JE da C (2018) Predição físico-química, modelagem e análise do mecanismo de interação da quitinase mo-chi1 [*Moringa oleifera*, Lam.], com polib-(1-4)-N-acetil-D-glucosamine: uma abordagem in silico. *DESAFIOS - Rev Interdiscip da Univ Fed do Tocantins* 5:111–120. <https://doi.org/10.20873/uft.2359-3652.2018vol5n1p111>
- Bitencourt MAO, Pedrosa M de FF, Lima MCJ de S, Langassner SMZ (2014) Extratos, frações, compostos isolados e composição farmacêutica de *Aspidosperma pyrifolium*, *Hancornia speciosa*, *Ipomoea asarifolia* e *Mimosa tenuiflora* aplicados no tratamento de processos de envenenamento por animais peçonhentos. BR102012026958. Deposit: Set. 24 2012.
- Bonete JM, Silva GD, Guidelli ÉJ et al (2020) Tissue reaction and anti-biofilm action of new biomaterial composed of latex from *Hancornia speciosa* Gomes and silver nanoparticles. *An Acad Bras Cienc* 92:1–14. <https://doi.org/10.1590/0001-3765202020191584>
- Braga FC, De França SC, Da Silva GC, Endringer DC (2009) Extrato e fração padronizados de folhas de *Hancornia speciosa* e sua composição farmacêutica. WO2009140749A3. Deposit: May 19 2009.
- Bustin S, Huggett J (2017) qPCR primer design revisited. *Biomol Detect Quantif* 14:19–28. <https://doi.org/10.1016/j.bdq.2017.11.001>
- Campos LZ, Nascimento ALB, Albuquerque UP et al (2018) Use of local ecological knowledge as phenology indicator in native food species in the semiarid region of Northeast Brazil. *Ecol Ind* 95:75–84. <https://doi.org/10.1016/j.ecolind.2018.07.016>
- Cardoso LDM, Reis BDL, Oliveira DDS, Pinheiro-Sant’Ana HM (2014) Mangaba (*Hancornia speciosa* Gomes) from the Brazilian Cerrado: nutritional value, carotenoids and antioxidant vitamins. *Fruits* 69:89–99. <https://doi.org/10.1051/fruits/2013105>
- Cercato LM, White PAS, Nampo FK et al (2015) A systematic review of medicinal plants used for weight loss in Brazil: Is there potential for obesity treatment? *J Ethnopharmacol* 176:286–296. <https://doi.org/10.1016/j.jep.2015.10.038>
- Chaves LJ, Ganga RMD, Guimarães RA, Caldeira AJR (2020) Quantitative and molecular genetic variation among botanical varieties and subpopulations of *Hancornia speciosa* Gomes (Apocynaceae). *Tree Genet Genomes* 16:50. <https://doi.org/10.1007/s11295-020-01444-0>
- Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F (2011) An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the fuzzy sets theory field. *J Informetr* 5:146–166. <https://doi.org/10.1016/j.joi.2010.10.002>
- Collevatti RG, Olivatti AM, Telles MPC, Chaves LJ (2016) Gene flow among *Hancornia speciosa* (Apocynaceae) varieties and hybrid fitness. *Tree Genet Genomes*. <https://doi.org/10.1007/s11295-016-1031-x>
- Collevatti RG, Rodrigues EE, Vitorino LC et al (2018) Unravelling the genetic differentiation among varieties of the Neotropical savanna tree *Hancornia speciosa* Gomes. *Ann Bot* 122:973–984. <https://doi.org/10.1093/aob/mcy060>

- Coradin L, Camilo J, Pareyn FGC (2018) Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial. Ministério do Meio Ambiente, Secretaria de Biodiversidade, Brasília
- Costa DFD, Vieira FDA, Fajardo CG, Chagas KPTD (2015) Diversidade genética e seleção de iniciadores ISSR em uma população natural de mangaba (*Hancornia speciosa* Gomes)(Apocynaceae). Revista Brasileira de Fruticultura 37:970–976. <https://doi.org/10.1590/0100-2945-246/14>
- Da Silva Júnior JF, Léo AS, Muniz AVCS et al (2018a) *Hancornia speciosa*: mangaba. In: Coradin L, Camillo J, Pareyn FC (eds) Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial - Plantas para o Futuro: Região Nordeste, 1a ed. Brasília: Ministério do Meio Ambiente, pp 177–192
- Da Silva Júnior JF, Muniz AVC da S, Léo AS, Maia (2018b) Descriptor for Mangaba (*Hancornia speciosa* Gomes). Bioversity International, Rome, Italy and Brazilian Agriculture Research Corporation, Embrapa Tabuleiros Costeiros, Aracaju, SE, Brazil
- Da Silva MAP, Vieira NF, Martins YAA et al (2015) Sensory profile of fermented milk drinks flavored with fruits from the Brazilian Cerrado. African J Food Sci 9:379–389. <https://doi.org/10.5897/AJFS2015.1293>
- Darrault RO, Schlindwein C (2005) Limited fruit production in *Hancornia speciosa* (Apocynaceae) and pollination by nocturnal and Diurnal insects1. Biotropica 37:381–388. <https://doi.org/10.1111/j.1744-7429.2005.00050.x>
- Das A, Pramanik K, Sharma R et al (2019) In-silico study of biotic and abiotic stress-related transcription factor binding sites in the promoter regions of rice germin-like protein genes. PLoS ONE 14:e0211887. <https://doi.org/10.1371/journal.pone.0211887>
- Das P, Majumder R, Mandal M, Basak P (2020) In-Silico approach for identification of effective and stable inhibitors for COVID-19 main protease (M pro) from flavonoid based phytochemical constituents of *Calendula officinalis*. J Biomol Struct Dyn. <https://doi.org/10.1080/07391102.2020.1796799>
- De Lima JP, Azevedo L, de Souza NJ et al (2015a) First evaluation of the antimutagenic effect of mangaba fruit in vivo and its phenolic profile identification. Food Res Int 75:216–224. <https://doi.org/10.1016/j.foodres.2015.05.045>
- De Lima JP, Fante CA, Freitas Pires CR et al (2015b) The ntioxidative potential and volatile constituents of mangaba fruit over the storage period. Sci Hortic (amsterdam) 194:1–6. <https://doi.org/10.1016/j.scienta.2015.05.035>
- Debibakas S, Rocher S, Garsmeur O et al (2014) Prospecting sugarcane resistance to sugarcane yellow leaf virus by genome-wide association. Theor Appl Genet 127:1719–1732. <https://doi.org/10.1007/s00122-014-2334-7>
- Delplancke M, Alvarez N, Espíndola A et al (2012) Gene flow among wild and domesticated almond species: insights from chloroplast and nuclear markers. Evol Appl 5:317–329. <https://doi.org/10.1111/j.1752-4571.2011.00223.x>
- dos Santos J, Neves MF, Rosalen PL et al (2016) Evaluation of the osteogenic potential of *Hancornia speciosa* latex in rat calvaria and its phytochemical profile. J Ethnopharmacol 183:151–158. <https://doi.org/10.1016/j.jep.2016.02.041>
- Dos Santos UP, Tolentino GS, Morais JS et al (2018) Physicochemical characterization, microbiological quality and safety, and pharmacological potential of *Hancornia speciosa* Gomes. Oxid Med Cell Longev. <https://doi.org/10.1155/2018/2976985>
- Dropa-Almeida D, Gaspar LMAC, Bastos BF, et al (2019) Uso do extrato encontrado de *Hancornia speciosa* Gomes - apocynaceae ou qualquer de seus derivados como agente antimicrobiano. BR102017025846A2. Deposit: Nov. 30 2017.
- EMATER (2019) Fruto do Nordeste surge como fonte de renda para agricultura familiar em Goiás. <https://www.emater.go.gov.br/wp/fruto-do-nordeste-surge-como-fonte-de-rendapara-agricultura-familiar-em-goias/>

- Endringer D, Oliveira O, Braga F (2014) In vitro and in silico inhibition of angiotensin-converting enzyme by carbohydrates and cyclitols. *Chem Pap* 68:37–45. <https://doi.org/10.2478/s11696-013-0407-8>
- Endringer DC, Pezzuto JM, Braga FC (2009) NF- κ B inhibitory activity of cyclitols isolated from *Hancornia speciosa*. *Phytomedicine* 16:1064–1069. <https://doi.org/10.1016/j.phymed.2009.03.022>
- FAO (2021) International network of food data systems (INFOODS). <http://www.fao.org/infoods/infoods/foodbiodiversity/en/>
- Feng S, He R, Lu J et al (2016) Development of SSR markers and assessment of genetic diversity in medicinal *Chrysanthemum morifolium* cultivars. *Front Genet* 7:1–11. <https://doi.org/10.3389/fgene.2016.00113>
- Ferreira HC, Serra CP, Endringer DC et al (2007a) Endothelium-dependent vasodilation induced by *Hancornia speciosa* in rat superior mesenteric artery. *Phytomedicine* 14:473–478. <https://doi.org/10.1016/j.phymed.2006.11.008>
- Ferreira HC, Serra CP, Lemos VS et al (2007b) Nitric oxide-dependent vasodilatation by ethanolic extract of *Hancornia speciosa* via phosphatidyl-inositol 3-kinase. *J Ethnopharmacol* 109:161–164. <https://doi.org/10.1016/j.jep.2006.06.009>
- Floriano JF, Neto FC, da Mota LSL et al (2016) Comparative study of bone tissue accelerated regeneration by latex membranes from *Hevea brasiliensis* and *Hancornia speciosa*. *Biomed Phys Eng Express* 2:045007. <https://doi.org/10.1088/2057-1976/2/4/045007>
- Ganeshpurkar A, Saluja AK (2017) The pharmacological potential of rutin. *Saudi Pharm J* 25:149–164. <https://doi.org/10.1016/j.jsps.2016.04.025>
- Ganga RMD, Ferreira GA, Chaves LJ et al (2010) Caracterização de frutos e árvores de populações naturais de *Hancornia speciosa* Gomes do cerrado. *Rev Bras Frutic* 32:101–113. <https://doi.org/10.1590/s0100-29452010005000019>
- GBIF (2021) Free and open access to biodiversity data. <https://www.gbif.org/>
- Geller FC, Teixeira MR, Pereira ABD et al (2015) Evaluation of the wound healing properties of *Hancornia speciosa* leaves. *Phyther Res* 29:1887–1893. <https://doi.org/10.1002/ptr.5438>
- Ghorbani A (2017) Mechanisms of antidiabetic effects of flavonoid rutin. *Biomed Pharmacother* 96:305–312. <https://doi.org/10.1016/j.biopha.2017.10.001>
- Goyal MR, Ayeleso AO (2018) Bioactive compounds of medicinal plants: properties and potential for human health. CRC Press, Toronto
- Green BR (2011) Chloroplast genomes of photosynthetic eukaryotes. *Plant J* 66:34–44. <https://doi.org/10.1111/j.1365-313X.2011.04541.x>
- Hagemann R (2002) Milestones in plastid genetics of higher plants. In: Esser K, Lüttge U, Beyschlag W, Hellwig F (eds) *Progress in botany (genetics — physiology — systematics — ecology)*. Springer, Berlin Heidelberg, Berlin, pp 5–51
- Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proc R Soc London Ser B Biol Sci* 270:313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Hjørland B (2013) User-based and cognitive approaches to knowledge organization: a Theoretical analysis of the research literature. *Knowl Organ* 40:11–27. <https://doi.org/10.5771/0943-7444-2013-1-11>
- Huang Y-Y, Matzke AJM, Matzke M (2013) Complete sequence and comparative analysis of the chloroplast genome of coconut palm (*Cocos nucifera*). *PLoS ONE* 8:e74736. <https://doi.org/10.1371/journal.pone.0074736>
- Isemura M (2019) Catechin in human health and disease. *Molecules*. <https://doi.org/10.3390/molecules24030528>
- Jena AB, Kanungo N, Nayak V et al (2021) Catechin and curcumin interact with S protein of SARS-CoV2 and ACE2 of human cell membrane: insights from computational studies. *Sci Rep* 11:2043. <https://doi.org/10.1038/s41598-021-81462-7>

- Joukhadar R, El-Bouhssini M, Jighly A, Ogbonnaya FC (2013) Genome-wide association mapping for five major pest resistances in wheat. *Mol Breed* 32:943–960. <https://doi.org/10.1007/s11032-013-9924-y>
- Kang Q, Li H, Jing Z (2020) Mapping and visualization of fuzzy control research: a scientometric analysis during 2010–2019. *ACM Int Conf Proc Ser.* <https://doi.org/10.1145/3421766.3421778>
- Karakülah G, Pavlopoulou A (2018) In silico phylogenetic analysis of hAT transposable elements in plants. *Genes (basel)* 9:284. <https://doi.org/10.3390/genes9060284>
- Katalinić V, Milos M, Modun D et al (2004) Antioxidant effectiveness of selected wines in comparison with (?)-catechin. *Food Chem* 86:593–600. <https://doi.org/10.1016/j.foodchem.2003.10.007>
- Khadivi-Khub A, Zamani Z, Fattahi R, Wunsch A (2014) Genetic variation in wild *Prunus* L. subgen. *Cerasus* germplasm from Iran characterized by nuclear and chloroplast SSR markers. *Trees* 28:471–485. <https://doi.org/10.1007/s00468-013-0964-z>
- Leite NS, da Silva RR, Marques JJ, et al (2018) Uso de um bioadsorvente produzido a partir de sementes de mangaba para remoção de contaminantes de águas e efluentes líquidos. BR102017001445A2. Deposit: Jan. 24 2017.
- Leite SP, Adami TB, Bjerk TR et al (2020) Ultrasonic assisted extraction of bioactive compounds from different parts of *Hancornia Speciosa* Gomes. *J Med Plants Res* 14:300–308. <https://doi.org/10.5897/JMPR2020.6944>
- Marinho DG, Alviano DS, Matheus ME et al (2011) The látex obtained from *Hancornia speciosa* Gomes possesses antiinflammatory activity. *J Ethnopharmacol* 135:530–537. <https://doi.org/10.1016/j.jep.2011.03.059>
- Marques S, Trevisan T, Maia C et al (2018) Improved methods for the rapid formation and prevention of advanced glycation end products (AGEs) in vitro by coupling to the Hypoxanthine/Xanthine oxidase assay system. *Biomedicines* 6:88. <https://doi.org/10.3390/biomedicines6030088>
- Marques SPD, Oliveira IR, Owen RW, Trevisan MTS (2015) Antioxidant capacity, angiotensin I converting enzyme (ACE) and acetylcholinesterase inhibition by extracts of the leaves and bark of *Hancornia speciosa* Gomes. *Hum Journals* 4:171–183
- Martins GV, Martins LSS, Veasey EA et al (2012) Diversity and genetic structure in natural populations of *Hancornia speciosa* var. *speciosa* Gomes in northeastern Brazil. *Rev Bras Frutic* 34:1143–1153. <https://doi.org/10.1590/S0100-29452012000400023>
- McCain KW (1991) Mapping economics through the journal literature: an experiment in journal cocitation analysis. *J Am Soc Inf Sci* 42:290–296. [https://doi.org/10.1002/\(SICI\)1097-4571\(199105\)42:4%3c290::AID-ASI5%3e3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(199105)42:4%3c290::AID-ASI5%3e3.0.CO;2-9)
- Monachino J (1945) A revision of *Hancornia* (Apocynaceae). *Lilloa* 11:19–48
- Moreira LN, Feltrin C, Gonçalves JE et al (2020) Determination of L-(+)-bornesitol, the hypotensive constituent of *Hancornia speciosa*, in rat plasma by LC-MS/MS and its application on a pharmacokinetic study. *Biomed Pharmacother.* <https://doi.org/10.1016/j.biopha.2020.110900>
- Moreira LN, Silva GC, Câmara DV et al (2019) The cyclitol L-(+)-Bornesitol as an active marker for the cardiovascular activity of the Brazilian medicinal plant *Hancornia speciosa*. *Biol Pharm Bull* 42:2076–2082. <https://doi.org/10.1248/bpb.b19-00601>
- Moura LC, Vitorino LC, Megguer CA et al (2016) Influence of refrigeration and cassava starch biofilm use on enzymatic browning in mangaba fruit (*Hancornia speciosa*). *Científica* 44:131. <https://doi.org/10.15361/1984-5529.2016v44n2p131-137>
- Muhammed MT, Aki-Yalcin E (2019) Homology modeling in drug discovery: overview, current applications, and future perspectives. *Chem Biol Drug Des* 93:12–20. <https://doi.org/10.1111/cbdd.13388>

- Mumu JR, Tahmid T, Azad MAK (2020) Job satisfaction and intention to quit: a bibliometric review of work-family conflict and research agenda. Appl Nurs Res. <https://doi.org/10.1016/j.apnr.2020.151334>
- Nascimento RSM, Cardoso JA, Coccozza FDM (2014) Caracterização física e físico-química de frutos de mangabeira (*Hancornia speciosa* Gomes) no oeste da Bahia. Rev Bras Eng Agrícola e Ambient 18:856–860. <https://doi.org/10.1590/1807-1929/agriambi.v18n08p856-860>
- Neves J dos S, Omar NF, Novaes PD (2014) Composição de látex da mangabeira e seu uso na regeneração óssea. BR102012025418A2. Deposit: Oct. 05 2012.
- Nikfarjam BA, Adineh M, Hajiali F, Nassiri-Asl M (2017) Treatment with rutin - A therapeutic strategy for neutrophil-mediated inflammatory and autoimmune diseases: Anti-inflammatory effects of rutin on neutrophils. J Pharmacopuncture 20:52–56. <https://doi.org/10.3831/KPI.2017.20.003>
- Oliveira de KS, Aloufa MAI (2020) Extrativismo e geração de renda da mangaba (*Hancornia speciosa* Gomes) em região litorânea do Rio Grande do Norte. Revista Brasileira de Agroecologia 15:291–304. <https://doi.org/10.33240/rba.v15i5.22898>
- Paula LC, Silva FA, Silva EP et al (2019) Influence of preservation methods on the bioactivity of mangaba (*Hancornia speciosa* Gomes) from the brazilian savannah. Food Sci Technol 39:403–409. <https://doi.org/10.1590/fst.31917>
- Pereira AC, Pereira ABD, Moreira CCL et al (2015) *Hancornia speciosa* Gomes (Apocynaceae) as a potential anti-diabetic drug. J Ethnopharmacol 161:30–35. <https://doi.org/10.1016/j.jep.2014.11.050>
- Pereira AVP, Pereira EBC, Silva Júnior JF, Silva DB (2016) *Hancornia speciosa*: mangaba. In: Vieira, R. F.; Camillo, J.; Coradin, L. (org.). Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial: Plantas para o Futuro - Região Centro-Oeste, 1a ed. Brasília: MMA pp. 237–256
- Pereira DF, Cazarolli LH, Lavado C et al (2011) Effects of flavonoids on α -glucosidase activity: potential targets for glucose homeostasis. Nutrition 27:1161–1167. <https://doi.org/10.1016/j.nut.2011.01.008>
- Perk AA, Shatynska-Mytsyk I, Gerc,ek YC et al (2014) Rutin mediated targeting of signaling machinery in cancer cells. Cancer Cell Int 14:124. <https://doi.org/10.1186/s12935-014-0124-6>
- Prudente DO, Paiva R, Nery FC et al (2016) Indirect in vitro organogenesis of *Hancornia speciosa* Gomes. Biosci J 32:721–729. <https://doi.org/10.14393/BJv32n3a2016-30381>
- Prudente DO, Paiva R, Nery FC et al (2017) Behavior of lateral buds of *Hancornia speciosa* after cryopreservation by encapsulation-vitrification. Acta Sci Biol Sci 39:87–93. <https://doi.org/10.4025/actascibiolsoci.v39i1.31620>
- Ribeiro SS, De Jesus AM, Dos Anjos CS et al (2012) Evaluation of the cytotoxic activity of some Brazilian medicinal plants. Planta Med 78:1601–1606. <https://doi.org/10.1055/s-0032-1315043>
- Ribeiro TP, Sousa TR, Arruda AS et al (2016) Evaluation of cytotoxicity and genotoxicity of *Hancornia speciosa* latex in *Allium cepa* root model. Brazilian J Biol 76:245–249. <https://doi.org/10.1590/1519-6984.20114>
- Rufino M do SM, Alves RE, de Brito ES, et al (2010) Bioactive compounds and antioxidant capacities of 18 non-traditional tropical fruits from Brazil. Food Chem 121:996–1002. <https://doi.org/10.1016/j.foodchem.2010.01.037>
- Santana LCL de A, dos Santos EAL, de Souza FM (2016) Método para a produção de lipase de *Aspergillus niger* utilizando resíduos do processamento de polpas de mangaba como substrato. BR102014013453. Deposit: May 26 2014.
- Santos ACB, Silva MAP, Santos MAF, Leite TR (2013) Levantamento etnobotânico, químico e farmacológico de espécies de Apocynaceae Juss. ocorrentes no Brasil. Rev Bras Plantas Med 15:442–458. <https://doi.org/10.1590/S1516-05722013000300019>
- Santos Filho OA, de Alencastro RB (2003) Modelagem de proteínas por homologia. Quim Nova 26:253–259. <https://doi.org/10.1590/s0100-40422003000200019>

- Santos GG, Silva MR (2012) Mangaba (*Hancornia speciosa* Gomez) ice cream prepared with fat replacers and sugar substitutes. Food Sci Technol 32:621–628. <https://doi.org/10.1590/S0101-20612012005000086>
- Santos PS, dos Santos FL, Silva Santana JG et al (2017) Genetic diversity and the quality of Mangabeira tree fruits (*Hancornia speciosa* Gomes – Apocynaceae), a native species from Brazil. Sci Hortic (amsterdam) 226:372–378. <https://doi.org/10.1016/j.scienta.2017.09.008>
- Santos UP, Campos JF, Torquato HFV et al (2016) Antioxidant, antimicrobial and cytotoxic properties as well as the phenolic content of the extract from *Hancornia speciosa* Gomes. PLoS ONE 11:1–19. <https://doi.org/10.1371/journal.pone.0167531>
- Sasson A, Malpica C (2018) Bioeconomy in Latin America. N Biotechnol 40:40–45. <https://doi.org/10.1016/j.nbt.2017.07.007>
- Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the Tortoise and the hare III. Am J Bot 94:275–288. <https://doi.org/10.3732/ajb.94.3.275>
- Silva AVC, Vitoria MF, Azevedo VRR et al (2019) Research article genetic diversity of the mangaba GeneBank using microsatellites. Genet Mol Res. <https://doi.org/10.4238/gmr18108>
- Silva da AVC, Amorim JAE, da Vitória MF et al (2017b) Characterization of trees, fruits and genetic diversity in natural populations of mangaba. Ciência e Agrotecnologia 41:255–262. <https://doi.org/10.1590/1413-70542017413048416>
- Silva da AVC, Amorim JAE, Melo MF de V, et al (2017a) Genetic Diversity of Remaining Populations of Mangaba (*Hancornia speciosa* Gomes) in Restingas of Brazil. J Agric Sci 9:46. <https://doi.org/10.5539/jas.v9n2p46>
- Silva GC, Braga FC, Lemos VS, Cortes SF (2016) Potent antihypertensive effect of *Hancornia speciosa* leaves extract. Phytomedicine 23:214–219. <https://doi.org/10.1016/j.phymed.2015.12.010>
- Silva GC, Braga FC, Lima MP et al (2011) *Hancornia speciosa* Gomes induces hypotensive effect through inhibition of ACE and increase on NO. J Ethnopharmacol 137:709–713. <https://doi.org/10.1016/j.jep.2011.06.031>
- Silva SA, Cruz EM de O, Reis RV dos, et al (2013) Caracterização morfológica e molecular de genótipos de mangaba. Rev Bras Frutic 35:1093–1100. <https://doi.org/10.1590/S0100-29452013000400021>
- Süntar I (2020) Importance of ethnopharmacological studies in drug discovery: role of medicinal plants. Phytochem Rev 19:1199–1209. <https://doi.org/10.1007/s11101-019-09629-9>
- Thatikunta R, Siva Sankar A, Sreelakshmi J et al (2016) Utilization of in silico EST–SSR markers for diversity studies in castor (*Ricinus communis* L.). Physiol Mol Biol Plants 22:535–545. <https://doi.org/10.1007/s12298-016-0367-x>
- Torres-Rêgo M, Furtado AA, Bitencourt MAO et al (2016) Antiinflammatory activity of aqueous extract and bioactive compounds identified from the fruits of *Hancornia speciosa* Gomes (Apocynaceae). BMC Complement Altern Med 16:275. <https://doi.org/10.1186/s12906-016-1259-x>
- Tripathi S, Srivastava Y, Sangwan RS, Sangwan NS (2020) In silico mining and functional analysis of AP2/ERF gene in *Withania somnifera*. Sci Rep 10:4877. <https://doi.org/10.1038/s41598-020-60090-7>
- Tropicos (2021) Tropicos.org. Missouri botanical garden. <http://www.tropicos.org>
- Vieira Neto RD, Cintra FLD, Ledo A da S, et al (2002) Sistema de produção de mangaba para os tabuleiros costeiros e baixadas litorâneas. Ministério da Agric Pecuária e Abastecimento, EMBRAPA 22.
- WIPO (2020) Guide to the International Patent Classification. World Intellectual Property Organization, WIPO
- Zhang Q, Sodmergen (2010) Why does biparental plastid inheritance revive in angiosperms? J Plant Res 123:201–206. <https://doi.org/10.1007/s10265-009-0291-z>

- Zitha EZM, Machado P da S, Junqueira LA, et al (2020) Impact of processing, packages, and storage on quality of mangaba (*Hancornia speciosa* Gomes) jelly. J Food Process Preserv 44:e14814. <https://doi.org/10.1111/jfpp.14814>
- Zupic I, Čater T (2015) Bibliometric methods in management and organization. Organ Res Methods 18:429–472. <https://doi.org/10.1177/1094428114562629>

5. ARTIGO 2

DIVERSIDADE GENO-FENOTÍPICA EM UMA POPULAÇÃO NATURAL DE *Hancornia speciosa* GOMES: IMPLICAÇÕES PARA A CONSERVAÇÃO E MELHORAMENTO

Periódico (publicado): Genetic Resources and Crop Evolution

RESUMO

Hancornia speciosa Gomes é uma frutífera nativa do Brasil, seus frutos e coprodutos são fonte de renda para famílias extrativistas e agroindústria. A redução de áreas naturais, ausência do plano de manejo sustentável e de programas de conservação e melhoramento, colocam em risco a manutenção de populações naturais da espécie. Portanto, o propósito deste estudo foi estimar a diversidade genética e avaliar a variabilidade fenotípica de frutos, sementes e plântulas em população natural de *H. speciosa* visando fornecer informações que contribuam para a conservação e melhoramento genético da espécie. Foram amostrados indivíduos em uma área de ocorrência natural no Brasil. Variáveis fenotípicas dos frutos, sementes e plântulas foram mensuradas e a diversidade genética estimada por 17 primers ISSR. 53,21% dos frutos apresentaram diâmetro longitudinal entre 27,5 e 35 mm e 63,15% dos frutos apresentaram diâmetro transversal entre 24,5 e 35 mm. As sementes apresentaram valores médios de comprimento, largura e espessura de 9,98 mm, 8,21 mm e 3,97 mm, respectivamente. O comprimento da parte aérea das plântulas variou entre 3,0 a 8,5 cm, comprimento da raiz de 4,0 a 15,8 cm e massa seca de 80 a 103 mg.plântula⁻¹. Verificou-se $H_e = 0,40$, $I = 0,50$, $N_a = 2,0$, $N_e = 1,71$ e 55,29 de similaridade genética. A população apresenta alta diversidade genética e variabilidade fenotípica de frutos, sementes e plântulas, com potencial para fornecer propágulos para a produção de mudas visando o enriquecimento de populações naturais e restauração florestal.

Palavras-chave: árvore frutífera brasileira, conservação, fenótipo, ISSR, mangabeira.

ABSTRACT

Título: Geno-phenotypic diversity in a natural population of *Hancornia speciosa* Gomes: implications for conservation and improvement

Hancornia speciosa Gomes is native tree to South America, and its fruits are a source of income for communities and the food industry. The reduction of natural occurrence areas of *H. speciosa*, the absence of a sustainable management plan conservation and breeding programs threaten its natural populations. Therefore, the purpose of this study was to evaluate the phenotypic variability of fruits, seeds, and seedlings and to estimate the genetic diversity. The main goal is to provide information for the specie conservation and breeding programs. Measurable fruits, seeds, and seedling variables; and genetic diversity was estimated by ISSR markers. Among fruits, 53.21% had a longitudinal diameter from 27.5 to 35 mm and 63.15% had a transversal diameter from 24.5 to 35 mm. The average seed biometry was 9.98 mm, 8.21 mm, and 3.97 mm, for length, width and thickness respectively. The length of seedlings shoots varied from 3.0 to 8.5 cm, roots from 4.0 to 15.8 cm and dry mass from 80 to 103 mg seedling⁻¹. The genetic parameters estimated were $He = 0.40$, $I = 0.50$, $Na = 2.0$, $Ne = 1.71$ and similarity of 55.29%. The population has high phenotypic variability of fruits, seeds and seedlings, and genetic diversity. There is potential for use of this population in the enrichment process for the restoration of natural populations.

Key-words: brazilian fruit tree, conservation, ISSR, phenotype, mangabeira.

5.1. Introdução

A fragmentação florestal e a exploração dos recursos genéticos florestais sem plano de manejo sustentável estão entre as principais ameaças à biodiversidade (Kim et al. 2015). Portanto, a gestão sustentável de remanescentes existentes tem cada vez mais relevância para a proteção do ecossistema florestal e conservação da biodiversidade, especialmente em regiões pouco desenvolvidas (Novello et al. 2018).

Conhecer a diversidade genética de espécies em populações naturais tem um papel importante na gestão de recursos genéticos e pode contribuir para orientar práticas de manejo desses recursos e, conseqüentemente, manter populações geneticamente viáveis de espécies exploradas em remanescentes florestais. A diversidade genética tem grande relevância também para os programas de melhoramento, visto que para estes, o reconhecimento e a medição de tal diversidade e de sua magnitude são essenciais (Heidari et al. 2019).

A diversidade presente nas populações naturais pode ser estudada a partir de análises fenotípicas que são amplamente utilizadas na identificação botânica de espécies e constituem uma etapa necessária para o uso eficiente dos recursos genéticos (Bashiri et al. 2017; Heidari et al. 2019).

A avaliação da informação genética em populações florestais também emprega marcadores moleculares, uma ferramenta amigável para estimar a diversidade (Heidari et al. 2019) que provou ser eficiente para várias espécies florestais e tem sido empregada em estudos com populações naturais (Sousa et al. 2020), em processo de restauração florestal (Cordeiro et al. 2019; Schwarcz et al. 2018; Zucchi et al. 2018), em bancos ativos de germoplasma (Osorio-Guarín et al. 2017; Santos et al. 2017) e melhoramento florestal (Lv et al. 2020).

A flora brasileira é uma excelente fonte de recursos genéticos florestais, com potencial para o extrativismo praticado por comunidades locais e aplicação na agroindústria, especialmente para a produção de alimentos. *Hancornia speciosa* Gomes (Apocynaceae) é uma árvore frutífera, conhecida por mangabeira e consta na lista de recursos genéticos com

potencial alimentício. O seu fruto, a mangaba, além de sabor e aroma agradáveis possui atividade antioxidante, é rico em ácido ascórbico, compostos fenólicos, carotenoides e apresenta altos níveis de potássio, ferro e zinco (Cardoso et al. 2014; Lima et al. 2015a, b).

A mangaba é também um importante produto florestal não madeireiro (PFNM) especialmente para comunidades socioeconomicamente vulneráveis. No Brasil são produzidos anualmente em torno de 1.933 toneladas de mangaba (IBGE 2019), essa produção é oriunda do extrativismo em remanescentes florestais, sendo poucos os plantios comerciais da espécie.

O estado de Sergipe é um dos maiores produtores de mangaba e nesse estado destacam-se as Catadoras de Mangaba, grupo tradicional diferenciado reconhecido pela Lei Estadual nº 288 de 2010 (Sergipe 2010) que usa a mangaba e coprodutos obtidos dessa fruta para geração de renda. O aumento da demanda por frutos e coprodutos de *H. speciosa*, ausência de plano de manejo sustentável e fragmentação das áreas de ocorrência natural colocam em risco a manutenção das populações naturais e a atividade de extrativismo.

A adoção de estratégias de conservação, manejo sustentável e domesticação são imperativas para assegurar a manutenção desse importante recurso genético. Nesse contexto, o conhecimento e monitoramento da diversidade genética pode ser eficaz para avaliar os impactos decorrentes do manejo de PFNM nas espécies exploradas (Dawson et al. 2014), além de contribuir na indicação de indivíduos de interesse para serem inseridos em programas de melhoramento genético.

Estudos de diversidade genética baseados em características fenotípicas (Almeida et al. 2019; Lúcio et al. 2019; Pinheiro et al. 2018) ou marcadores moleculares (Costa et al. 2017; Fajardo et al. 2018; Jimenez et al. 2015; Luz et al. 2020; Yokomizo et al. 2017) já foram realizados com populações naturais de *H. speciosa* e conservadas *in situ*. Normalmente estas análises são realizadas isoladas ou em conjunto apenas para acessos em bancos de germoplasma (Santos et al. 2017; Silva et al. 2019).

Portanto, objetivou-se estimar a diversidade genética e avaliar a variabilidade fenotípica de frutos, sementes e plântulas em população natural de *H. speciosa* visando contribuir para a conservação e melhoramento genético da espécie.

5.2. Material e Métodos

5.2.1 Amostragem da população

O material vegetal foi coletado de 14 árvores em população natural, selecionadas aleatoriamente considerando distâncias entre 30 e 50 m (Seoane et al. 2005), que apresentavam bom aspecto fitossanitário e que durante a coleta dispunham de frutos com sinais de maturação (coloração amarelo claro com pigmentação vermelha e textura tenra).

A área de coleta está localizada no município de Pirambu, litoral do estado de Sergipe, Brasil, e correspondeu a aproximadamente 0,80 km² (Figura 1).

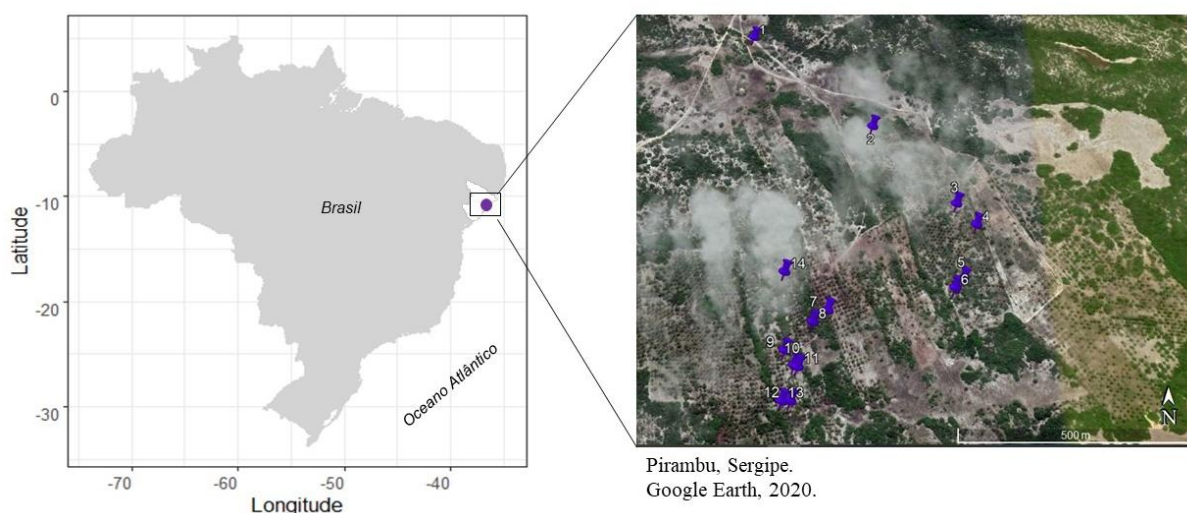


Figura. 1 Distribuição espacial de indivíduos de *Hancornia speciosa* Gomes em área de coleta, município de Pirambu, Sergipe, Brasil.

O clima da região se caracteriza como subúmido, tipo As - tropical chuvoso com verão seco, pela classificação climática de Köppen-Geiger (Alvares et al. 2013). A precipitação anual varia de 1.500 a 1.800 mm, com temperatura média anual de 26 °C e o solo do tipo neossolo quartzarênico (Santos et al. 2018; Sergipe 2016).

A população foi escolhida por apresentar diversidade genética ($He = 0,36$ e $I = 0,53$) e variação fenotípica entre os indivíduos (Freitas 2016) e por estar em uma área de intensa extração pelas Catadoras de Mangaba.

5.2.2 Variabilidade de frutos, sementes e plântulas

Os frutos coletados (171) foram submetidos a análises individualizadas quanto aos aspectos quantitativos: diâmetro longitudinal (DL, $h^2 = 0,77$) e diâmetro transversal (DT, $h^2 = 0,78$), mensurados com o auxílio de paquímetro digital (Messen) e expressos em milímetros; massa fresca da polpa em gramas (MP, $h^2 = 0,56$), número de sementes (NS), massa de sementes (MS, $h^2 = 0,55$) e rendimento de polpa (RP, $h^2 = 0,46$) estimado pela massa da polpa dividida pela massa da polpa mais a massa de sementes ($MP/(MP+MS)$).

As sementes (oito repetições de 25 sementes) foram submetidas a análises individualizadas quantitativas quanto ao comprimento (C, $h^2 = 0,99$), largura (L, $h^2 = 0,98$) e espessura (E, $h^2 = 0,99$), mensurados com o auxílio de paquímetro digital (Messen) e expressos em milímetros.

A principal forma de propagação natural de *H. speciosa* é por sementes, sendo formados bancos de plântulas no solo, portanto, esses propágulos e plântulas foram avaliados. Para avaliar o potencial germinativo das sementes, foi realizado o teste de germinação com oito repetições de 25 sementes, que foram consideradas germinadas quando as plântulas apresentaram todas as estruturas essenciais normais (raiz primária, hipocótilo, cotilédones, epicótilo e protófilos abertos).

O vigor foi avaliado pelo comprimento e massa seca das plântulas. Oito repetições de 10 plântulas foram submetidas a mensurações do comprimento da parte aérea (CPA, $h^2 = 0$) e da raiz (CR, $h^2 = 0,46$), expressos em centímetros. Para determinação da massa seca de plântulas (MSP), oito repetições de 10 plântulas normais obtidas no teste de germinação foram acondicionadas em sacos de papel identificados e levadas à estufa com circulação de ar forçada, a 80 °C por um período de 24h, obtendo-se os dados em $mg.plântula^{-1}$.

As variáveis fenotípicas de frutos, sementes e plântulas foram analisadas por estatística descritiva. Foi calculado o coeficiente de correlação não paramétrico de Spearman (r_s) e o respectivo nível de significância (p). Todas as análises estatísticas foram realizadas no software R v. 4.0.2 (R Core Team, 2020).

5.2.3 Análise molecular

Para o estudo da diversidade genética, realizou-se a extração de DNA genômico de folhas jovens de acordo com o protocolo descrito por Nienhuis et al. (1995) com modificações. O DNA foi solubilizado em TE (1 mM de Tris; 0,1 mM de EDTA) e a qualidade e concentração do DNA de cada amostra determinadas em espectrofotômetro (ASTRAGEN), com base nas absorbâncias A_{260}/A_{280} nm. Para a geração de polimorfismo entre os indivíduos, foram testados 22 *primers* ISSR (Inter Simple Sequence Repeat) (Zietkiewicz et al. 1994) da IDT (*Integrated DNA Technologies*) e selecionados 17 que permitiram a obtenção de fragmentos de alta intensidade e resolução (Tabela 1).

Tabela 1. Primers ISSR usados para o estudo de diversidade de *Hancornia speciosa* Gomes e suas respectivas sequências e temperatura de anelamento (T° A).

Primers ISSR	Sequência	T° A	Primers ISSR	Sequência	T° A
UBC 807	(AG)8-T	43	DAT	(GA)7-RG	43
UBC 808	(AG)8-C	47	GOOFY	(GT) - YG	48
UBC 809	(AG)8-G	48	M1	(GA)5	31
UBC 810	(GA)8-T	43	M2	(GA)8	64
UBC 811	(GA)8-C	45	MANNY	(CAC)4-CR	54
UBC 813	(CT)8-T	44	MAO	(CTC)4-RC	45
UBC 825	(AC)8-T	47	OMAR	(GAG)4-RC	47
UBC 834	(AG)8-YT	47	TERRY	(GTC)A-RC	53
AW3	(GT)6-RG	44			

As reações de amplificação foram conduzidas em termociclador (Uniscience Biometra Tpersonal), em um volume de 12 µL contendo: DNA genômico (5 ng); tampão de PCR com MgCl adicionado; dNTPs (2,5 mM); enzima Taq polimerase (5 U/µL) (Invitrogen); *primer* (2 µM). As reações de PCR foram submetidas a 35 ciclos de amplificação após adesnaturação inicial a 94°C por 3 minutos. Cada ciclo foi constituído de três temperaturas: 94°C por 1 minuto, para a desnaturação da dupla fita de DNA; 31-64°C (considerando a T° A do *primer*) para anelamento do DNA ao *primer*; e por fim, 72°C por 2 minutos para a extensão das novas fitas de DNA. Após os ciclos, o processo foi finalizado a 72°C por 7 minutos. Os produtos da reação foram separados por eletroforese, em cuba horizontal, utilizando-se gel de agarose 1,5% em tampão TBE 1X, a 100 V, corados com Gel Red™ e os géis fotografados sob luz ultravioleta em transiluminador digital (Benchtop). O perfil eletroforético de cada gel foi analisado visualmente e, segundo a ausência (0) ou presença (1) de fragmento amplificado, foi gerada uma matriz binária.

5.2.4 Análises populacionais

A similaridade genética foi calculada por meio do coeficiente de Jaccard, utilizando o programa NTSYS 2.2 (Rohlf 2005). Para a representação dos dados de similaridade foi construído um dendrograma, obtido pelo método de agrupamento *Unweighted Pair-Group Method Arithmetic Average* (UPGMA) (Sneath e Sokal 1973). Com base na matriz de similaridade genética foi realizada a Análise de Coordenadas Principais (*Principal Coordinate Analysis* (*Principal Coordinate Analysis* - PCoA) utilizando o programa GenAlEx 6.5 (Peakall e Smouse 2012).

5.3. Resultados

5.3.1 Variabilidade de frutos, sementes e plântulas

Os frutos de *H. speciosa* são do tipo baga, com formato elipsoidal ou arredondado, exocarpo amarelo com manchas ou estrias avermelhadas, mesocarpo carnoso de coloração amarelada, indeiscente e com número variado de sementes (Figura 2a e b), com maior frequência dos diâmetros longitudinal e transversal correspondentes a 30 cm (Figura 2c).

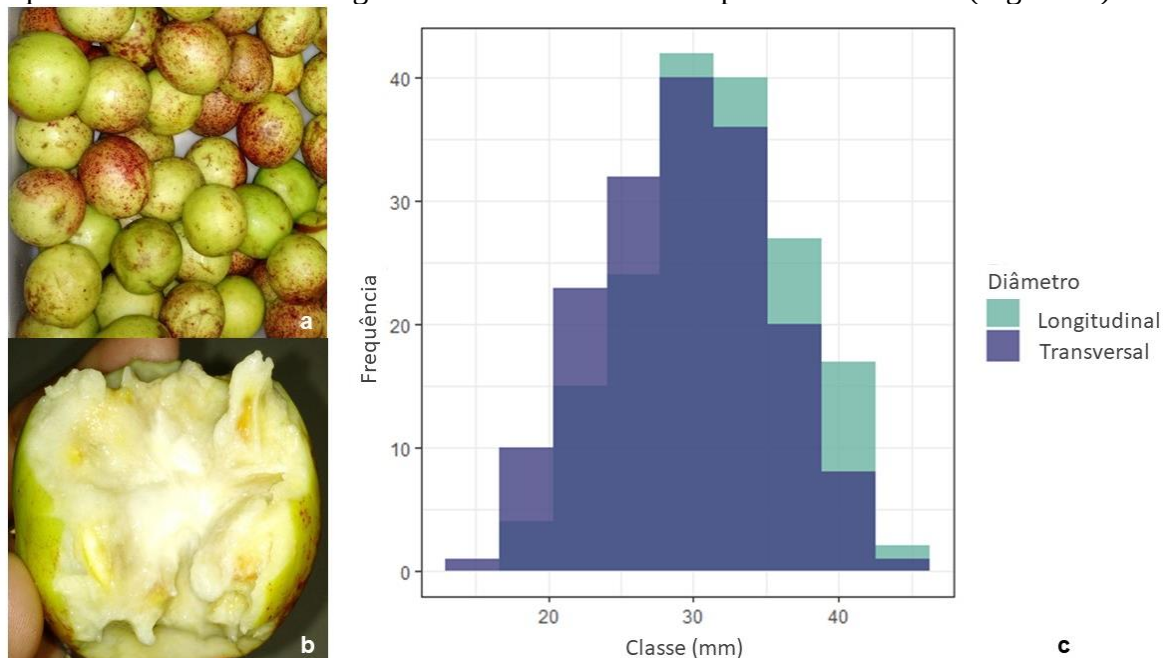


Figura 2. Frutos (a), coloração da polpa (b) e variáveis biométricas (c) para frutos de *Hancornia speciosa* Gomes obtidos em população natural.

O diâmetro longitudinal variou de 17,5 a 45,0 mm e 53,21% dos frutos apresentam entre 27,5 e 35,0 mm. O diâmetro transversal variou de 14,9 a 44,4 mm e 63,15% dos frutos estão entre 24,5 e 35,0 mm (Figura 3).



Figura 3. Variância fenotípica no tamanho de frutos de *Hancornia speciosa* Gomes obtido em população natural.

Para as características massa da polpa, massa de sementes e número de sementes (Figura 4), também se verificou alta variabilidade.

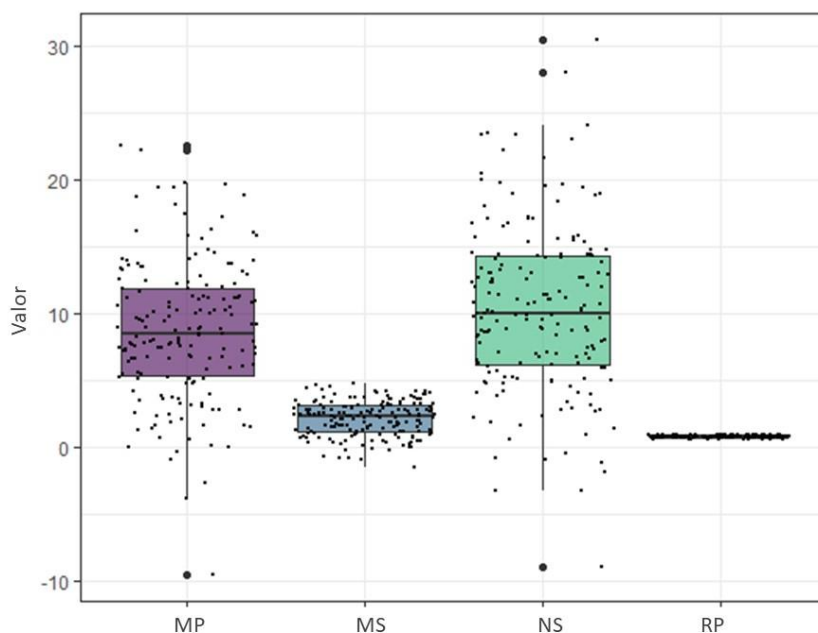


Figura 4. Massa da polpa (MP), massa de sementes (MS), número de sementes (NS) e rendimento da polpa (RP) para frutos de *Hancornia speciosa* Gomes obtidas em população natural.

As estimativas dos coeficientes de correlação de Spearman (r_s) (Figura 5) não foram significativas para variável rendimento da polpa (RP) quando correlacionadas com as demais variáveis. As variáveis massa da polpa (MP), massa de sementes (MS) e número de sementes (NS) apresentaram correlação positiva entre elas e significativas. O diâmetro transversal e diâmetro longitudinal também se correlacionam positivamente ($r = 0.9$).

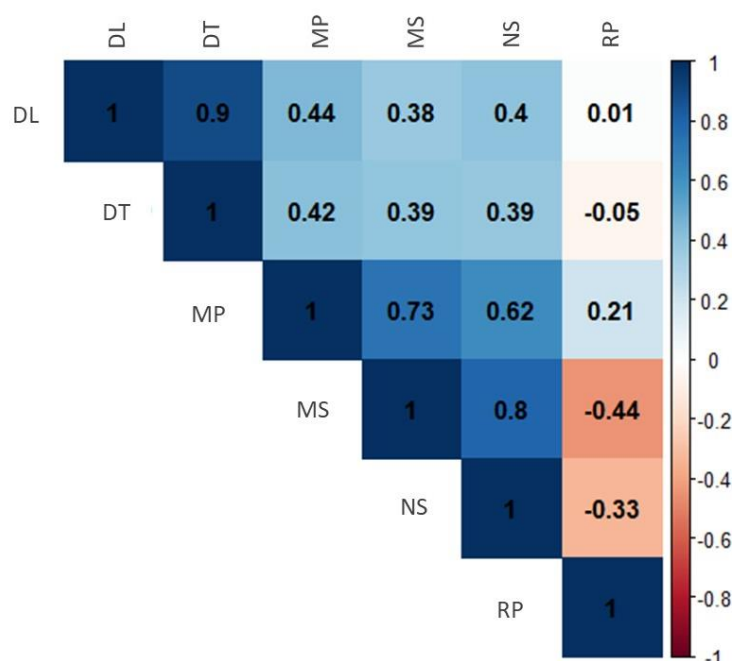


Figura 5. Correlação de Spearman para as variáveis diâmetro longitudinal (DL), diâmetro transversal (DT), massa de polpa (MP), massa de sementes (MS), número de sementes (NS) e rendimento da polpa (RP), para frutos de *Hancornia speciosa* Gomes.

As sementes de *H. speciosa* possuem uma forma ovalada e achatada, com coloração marrom-claro e hilo central (Figura 6a, b). Seus valores médios de comprimento, largura e espessura foram 9,9 mm, 8,2 mm e 3,9 mm, respectivamente (Figura 6c).

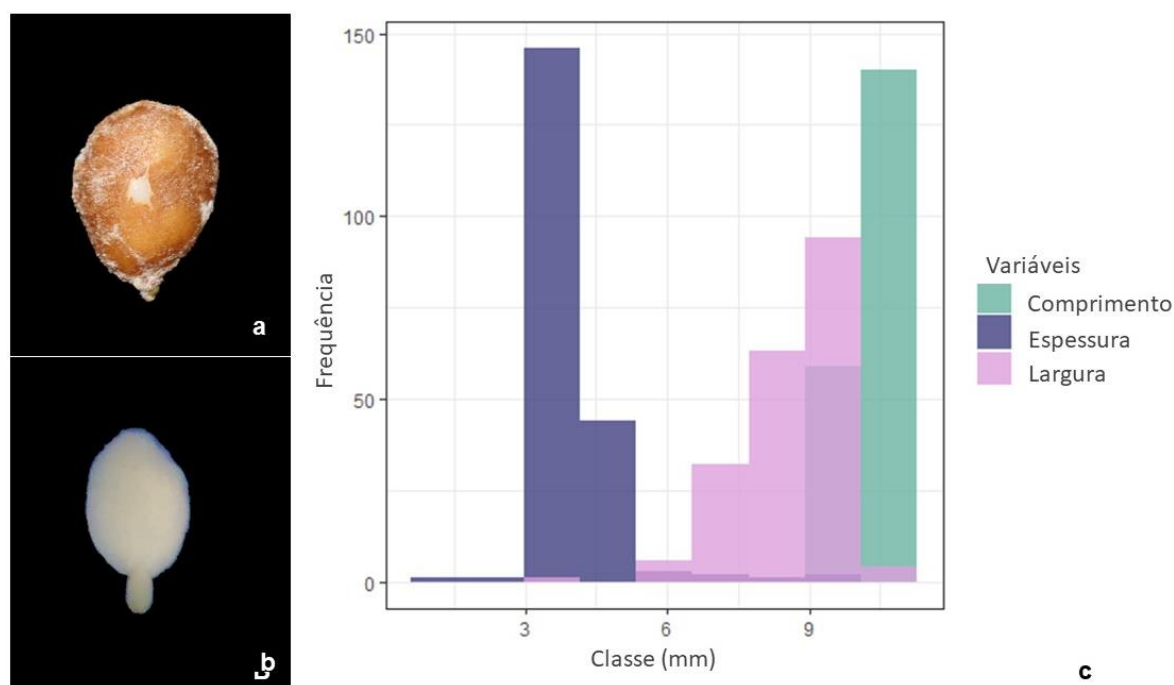


Figura 6. Sementes (a), embrião (b) e variáveis biométricas (c) para sementes de *Hancornia speciosa* Gomes obtidas em população natural.

Para o comprimento, 72,5% das sementes apresentam valores entre 10,0 e 10,5 mm, para a largura, 48,5% das sementes estão na classe de 9,0 a 10,1 mm e 73,5% das sementes apresentam espessura de 3,0 a 4,5 mm.

As sementes de *H. speciosa* apresentaram 94% de germinação, com o comprimento da parte aérea variando de 3,0 cm a 8,5 cm e comprimento da raiz de 4,0 cm a 15,8 cm. Na Figura 7 estão apresentadas plântulas normais e anormais observadas, bem como a frequência dos comprimentos da parte aérea e raiz.

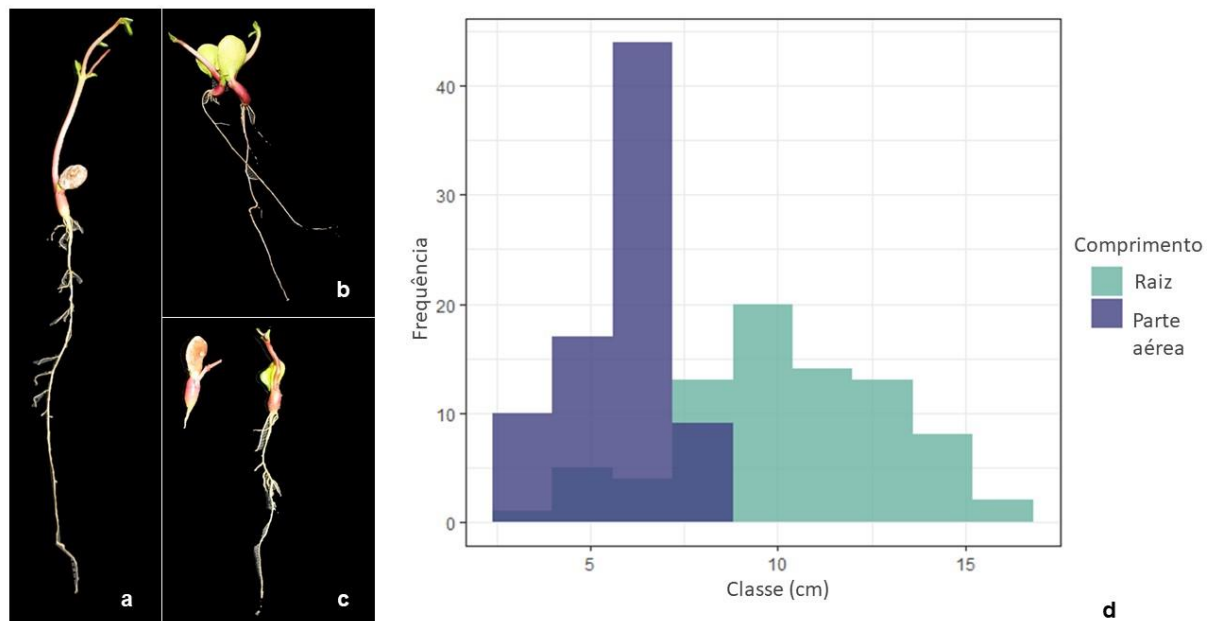


Figura 7. Plântula normal (a), plântulas anormais (b), poliembrionia (c) e variáveis biométricas (d) para plântulas de *Hancornia speciosa* Gomes obtidas em população natural.

As estimativas dos coeficientes de correlação de Spearman (Figura 8) foram positivas para variável comprimento (C) quando correlacionada com a largura (L), espessura de sementes (E) e massa seca de plântulas (MSP), sendo significativa apenas para largura ($r = 0,72$). Para as variáveis referentes às plântulas, a única correlação positiva e significativa foi para a parte aérea e raiz de plântulas ($r = 0,87$).

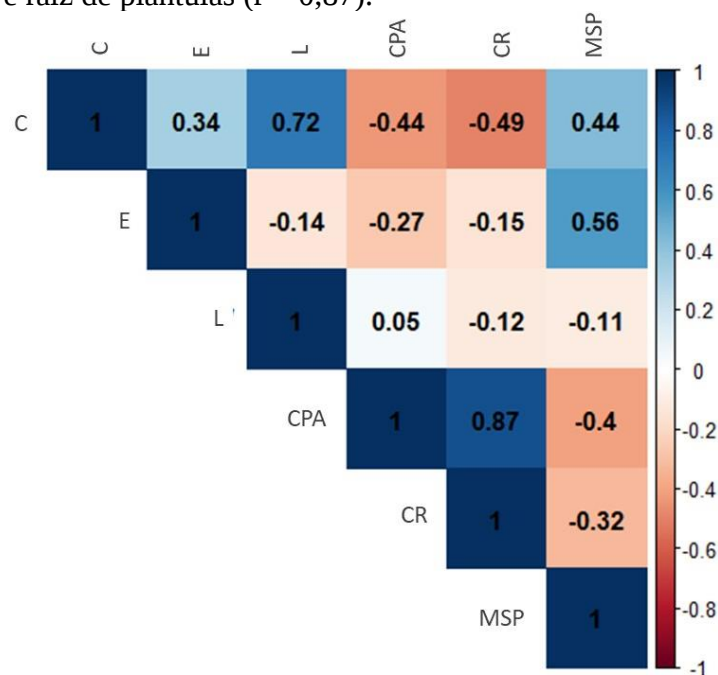


Figura 8. Correlação de Spearman para as variáveis comprimento (C), largura (L) e espessura (E) das sementes e comprimento da parte aérea (CPA), da raiz (CR) e massa seca de plântulas (MSP) de *Hancornia speciosa* Gomes.

5.3.2 Análises populacionais

A população estudada apresentou variação significativa em 81 *locos* polimórficos obtidos de 17 *primers* ISSR que foram amplificados com 100% de polimorfismo. O número médio de fragmentos por *primer* foi de 7,88, o menor valor (2) obtido pela amplificação usando o *primer* M2 e o maior número de fragmentos (12) observado com o *primer* UBC 813.

O número médio de alelos observados (N_a) foi igual a 2,0 e o número médio de alelos efetivos (N_e) correspondeu a 1,71. O valor médio de heterozigosidade estimado para a população com base no Equilíbrio de Hardy-Weinberg (Nei 1978) foi $H_E = 0,40$ e índice de Shannon (I) = 0,58. Os parâmetros genéticos H_E e I variam de 0 a 1,0, sendo que quanto mais próximo de 1, maior a diversidade genética (Souza-Sobreira et al. 2015).

Considerando ainda a caracterização genética desta população, avaliou-se a similaridade genética entre os indivíduos, verificando-se média de 55,29%, erro padrão médio de 1,34% e amplitude das similaridades com variação de 26,19 a 76,19%. Estimada pelo coeficiente de Jaccard, a similaridade genética foi utilizada para construir o dendrograma UPGMA que permitiu a visualização de quatro grupos distintos (Figura 9).

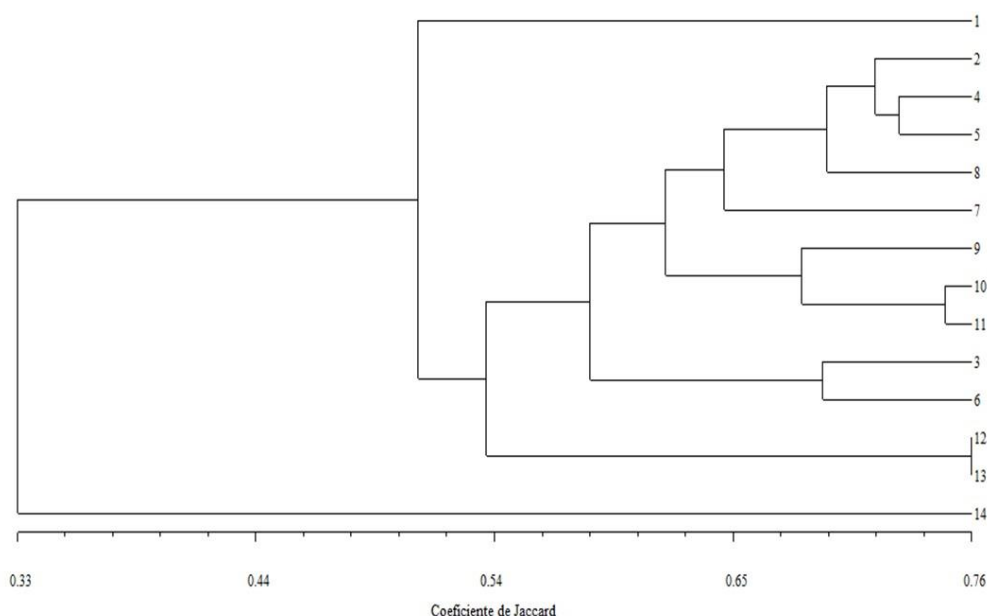


Figura 9. Dendrograma gerado pelo método UPGMA, baseado na similaridade de Jaccard, para população natural de *Hancornia speciosa* Gomes.

Por ser um método hierárquico, o UPGMA agrupa os indivíduos por um processo que se repete em vários níveis até que o dendrograma seja estabelecido (Puiatti et al. 2014). Neste estudo, os indivíduos mais semelhantes foram 12 e 13 ($76\% \pm 0,013$) e os mais divergentes 1 e 14 ($26\% \pm 0,013$). Os quatro grupos são grupo 1: 1, grupo 2 formado por 10 indivíduos (2, 3, 4, 5, 6, 7, 8, 9, 10 e 11), grupo 3 (12 e 13) e o grupo 4: 14.

Os genótipos foram analisados por meio da Análise de Coordenadas Principais (PCoA) com base no coeficiente de similaridade, o primeiro e segundo componentes fornecidos pela análise explicam 31,38% de variação dos indivíduos, contribuindo com 17,29% (eixo 1) e 14,09% (eixo 2).

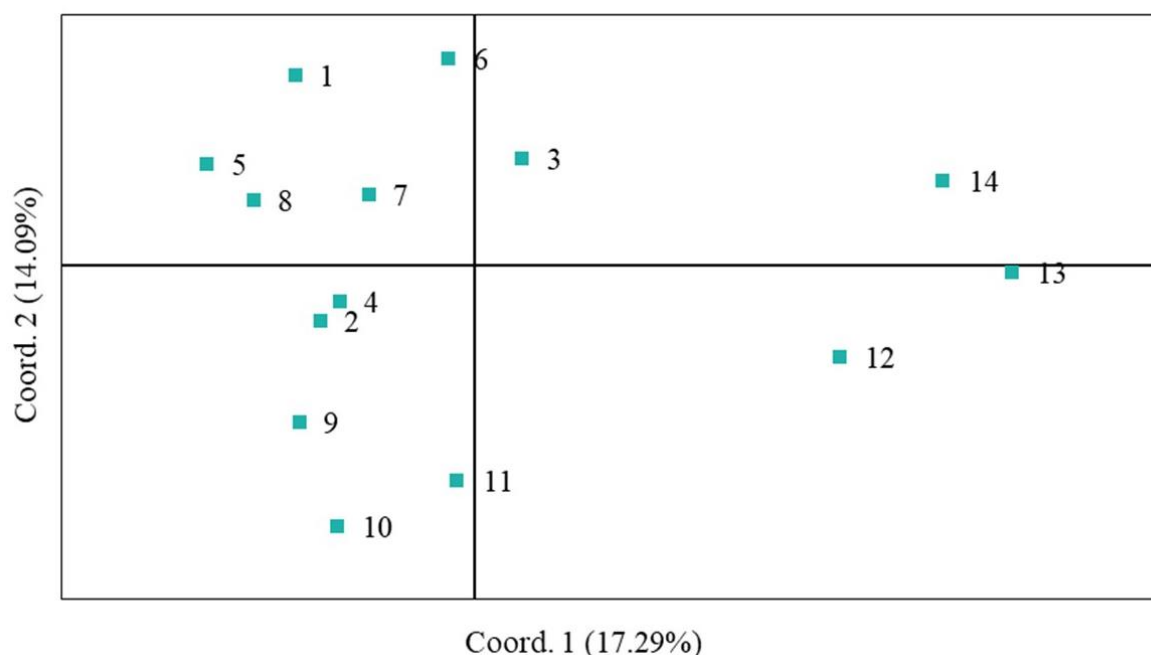


Figura 10. Análise de Coordenadas Principais (PCoA) para população natural de *Hancornia speciosa* Gomes.

O agrupamento PCoA validou o posicionamento dos indivíduos mais próximos agrupados anteriormente pelo método UPGMA, porém adicionou ao grupo o indivíduo 1 que estava em um grupo isolado, o mesmo ocorreu para o indivíduo 14 que passou a compor o segundo grupo juntamente com os indivíduos 12 e 13.

5.4. Discussão

Os resultados deste estudo revelam altos níveis de variabilidade fenotípica para os frutos, sementes e plântulas de *H. speciosa*. Os frutos apresentam diâmetros longitudinal e transversal superiores aos obtidos em estudo com *H. speciosa* no município de Cassilândia, Mato Grosso do Sul, em que 45% dos frutos apresentaram diâmetro longitudinal na faixa entre 23,3 e 28,5 mm e 76% dos frutos apresentando valores em duas classes de diâmetro transversal entre 17,6 a 23,1 mm e de 23,2 a 28,7 mm (Zuffo et al. 2019). A variabilidade de características como essas é comumente observada para espécies que não foram domesticadas e encontram-se em populações naturais.

A massa da polpa (MP) e massa de sementes (MS) dessa população são menores que as obtidas para populações de *H. speciosa* do oeste da Bahia, para as quais verificou-se MP correspondente a 14,7 g e 2,4 g para MS por fruto (Nascimento et al. 2014). Para frutos produzidos em Nova Xavantina, no leste do estado de Mato Grosso, a MP média foi de 40,15 g, MS média de 6,33 g e número de sementes médio de 22,0 (Gonçalves et al. 2013).

As diferenças observadas para o diâmetro dos frutos, massa da polpa e de sementes, podem estar associadas à variabilidade genética ou a variações ambientais devido a localidades geográficas diferentes (Vieira e Gusmão 2008). Especificamente para a população analisada nesta pesquisa, as condições edafoclimáticas correspondem à clima subúmido (As), com precipitação anual variando de 1.500 a 1.800 mm, temperatura média de 25,5 °C e a vegetação considerada Mata Secundária: Mata Atlântica, Manguezal, Restinga (Climate-Data.org 2020; Sergipe 2016). Na área avaliada na Bahia, o clima é tropical (Aw), precipitação média de 1.110 mm e temperatura média de 24,7 °C e inserido no bioma Cerrado (Climate-Data.org 2020; Nascimento; Cardoso et al. 2014). No Mato Grosso, o clima é tropical (Aw) com precipitação média de 1.498 mm, temperatura média anual de 25,1 °C e a vegetação é o Cerrado sensu stricto (Climate-Data.org 2020; Silva et al. 2008).

O rendimento da polpa foi semelhante ao obtido para populações de *H. speciosa* no Cerrado. O conhecimento dessa variável é necessário para o direcionamento da seleção de indivíduos para o cultivo comercial, que depende de árvores que reúnam características superiores às que estão em população natural (Ganga et al. 2010).

H. speciosa é uma espécie alógama e autoincompatível, para que ocorra a produção de frutos é necessária a presença de indivíduos geneticamente diferentes. Adicionalmente, de acordo com Darrault e Schlindwein (2005), a presença de polinizadores também exerce efeitos sobre a produção, tamanho e número de sementes dos frutos, sendo observado que a polinização eficiente contribui para a formação de frutos mais pesados e, consequentemente, com maior quantidade de polpa e maior número de sementes. Ainda de acordo com os autores, os polinizadores mais eficientes para a espécie são mariposas (Sphingidae), abelhas (Euglossini e Centridini) e borboletas (Nymphalidae e Hesperidae).

As correlações positivas para as variáveis mensuradas nos frutos (Figura 5) indicam que as características número de sementes e rendimento de polpa são beneficiadas ou prejudicadas pelas mesmas causas de variação. Desta forma, a escolha de frutos com menor número de sementes implicaria em menor quantidade de polpa, sendo estas variáveis relevantes especialmente no processamento dos frutos, quando se deseja menor massa de sementes e maior de polpa, produto este de maior interesse comercial (Gonçalves et al. 2013; Nascimento et al. 2014).

Para a comercialização dos frutos *in natura*, seria vantajosa a escolha dos maiores frutos, considerando que neste tipo de mercado o tamanho dos frutos juntamente com a coloração são atrativos para os consumidores (Chitarra e Chitarra 2005), no caso de *H. speciosa*, a coloração amarelo claro com pigmentação vermelha está associada a frutos maduros e mais valorizados no mercado (EMBRAPA 2021).

As estimativas de correlação podem ser informativas para o melhoramento da espécie. Isso ocorre porque as características de interesse relacionadas à produtividade e à qualidade de frutos, que são de difícil mensuração, podem apresentar correlação com outra característica de fácil mensuração, o que facilita e acelera a seleção de plantas superiores (Giles et al. 2016).

As variáveis analisadas são altamente influenciadas pelo ambiente. No entanto, os indivíduos neste estudo foram coletados na mesma área e, portanto, com condições ambientais semelhantes. Sendo assim, grande parte da variabilidade observada pode estar associada à variação genética.

A população está localizada em área de extrativismo das Catadores de Mangaba, essas informações podem ser úteis para atividades posteriores visando contribuir com essa comunidade que obtém parte de sua renda pela comercialização de frutos e co-produtos obtidos a partir da espécie.

As sementes de *H. speciosa* avaliadas no presente trabalho possuem comprimento, largura e espessura semelhantes aos obtidos em estudos com *H. speciosa* provenientes do BAG Mangaba mantido pela Embrapa Tabuleiros Costeiros em Sergipe (Soares et al. 2019) e com população natural em Cassilândia, no Mato Grosso do Sul (Zuffo et al. 2019).

As sementes possibilitam que as plantas se estabeleçam no tempo e no espaço e garantem a próxima geração de indivíduos (Mitchell et al. 2017). Elas também são fontes de variabilidade, que pode se manifestar em diferentes escalas, entre populações, dentro da população e no próprio indivíduo. A variabilidade existente em nível intrapopulacional aumenta as chances de sobrevivência das gerações da população em um ambiente incerto, o que reduz o risco de extinção em condições variáveis (Mitchell et al. 2017).

Após a germinação, os valores observados neste estudo para o comprimento da parte aérea e raiz também são semelhantes aos valores reportados na literatura para a espécie, que corresponderam a 4,96 cm e 8,5 cm (Santos et al. 2010). A massa seca de plântulas variou de 80 a 103 mg.plântula⁻¹, valores superiores aos obtidos para esta espécie em estudo

desenvolvido por Pinto et al. (2014). A massa seca de plântulas foi semelhante aos resultados obtidos por Soares et al. (2015).

A altura e massa seca das plântulas estão relacionadas ao vigor do lote de sementes. O vigor está relacionado à emergência rápida e uniforme das plantas sobre uma ampla gama de condições em campo (Rajjou et al. 2012). Existem vários fatores que exercem influência sobre o vigor das sementes, entre eles a genética, condições ambientais durante e após a formação das sementes, danos mecânicos, ação de micro-organismos, insetos e condições de armazenamento (Carvalho e Nakagawa 2012).

Para espécies recalcitrantes, como *H. speciosa*, o vigor tem implicações para a regeneração. Sementes recalcitrantes possuem baixa longevidade e para o sucesso reprodutivo as espécies investiram em outras estratégias como a rápida germinação e estabelecimento de plântulas, formando um banco de plântulas no solo (Pammenter e Berjak 2000).

O estudo do crescimento das plantas é fundamental para aumentar o campo de conhecimento sobre o comportamento dos indivíduos de uma população, além disso, parâmetros morfológicos são amplamente utilizados para determinar o padrão de qualidade das mudas. Desse modo, o conhecimento da variabilidade fenotípica permite selecionar materiais para conservação e também com melhores possibilidades de domesticação.

Em se tratando da diversidade genética das árvores adultas avaliadas por marcadores moleculares, os resultados encontrados neste estudo para número de *locos* polimórficos (81) foi superior ao relatado para populações de *H. speciosa* estudadas por meio de marcadores ISSR no Rio Grande do Norte com 30 *locos* (Costa et al. 2015) e para acessos do Banco Ativo de Germoplasma de Mangaba (BAG Mangaba) mantido pela Embrapa Tabuleiros Costeiros em Sergipe com 49 *locos* polimórficos (Santos et al. 2017). Para populações de Pernambuco verificou-se 93 *locos* polimórficos (Jimenez et al. 2015). O número médio de *locos* por *primer* no presente estudo foi de 7,88, enquanto que para o BAG Mangaba foi de 4,08 e para a população de Pernambuco foi de 11,5.

O percentual de *locos* polimórficos é uma informação relevante, sobretudo para estudos que usam marcadores dominantes, como o ISSR, pois é considerado como uma medida de diversidade genética (Lorenzoni et al. 2014). Portanto, o alto grau de polimorfismo verificado para a população estudada sugere que ela apresenta variabilidade genética.

O número médio de alelos observados (N_a) e o número médio de alelos efetivos (N_e) informam sobre a diversidade alélica e a diferença na distribuição entre os genótipos, indicando a existência de diversidade genética. O N_e é conhecido também por tamanho efetivo da população e refere-se ao número de indivíduos que se reproduzem e deixam descendentes, desta maneira, transmitindo os genes para a próxima geração. Esse conceito é fundamental para fins evolutivos e conservação genética (Charlesworth e Willis 2009; Whitlock et al. 2000).

Na genética de populações, o N_e tem grande importância porque mede a taxa de deriva genética e endogamia e permite a estimativa e comparação com o tamanho mínimo viável ou mínimo adequado, o qual assegura que a população mantenha a viabilidade ao longo do tempo e que exista sustentabilidade genética na sua descendência (Vencovsky e Crossa 2003).

O Índice de Nei é considerado um índice próprio para avaliar diversidade genética através da heterozigosidade genotípica (Nei 1973), sendo um dos índices mais utilizados para estudos de diversidade. Ao usar marcadores dominantes, como o ISSR, o índice de Shannon (I) é útil para análise populacional porque permite que os efeitos de distorção, causados pela incapacidade de detectar *locos* heterozigotos, sejam minimizados (Silva et al. 2017).

Índices de diversidade genética mais elevados são esperados em estudos envolvendo amostras amplas (Cavallari et al. 2008). Entretanto, estudos com populações de *H. speciosa* com maior número de indivíduos do que o avaliado neste trabalho apresentaram valores menores para estes índices. A exemplo, populações do Rio Grande do Norte com $H_E = 0,18$ e $I = 0,26$ (Costa et al. 2015), populações de Sergipe com $I = 0,29$ a $0,45$ e $H_E = 0,20$ a $0,31$ (Silva et al. 2017) e para acessos do BAG Mangaba $H_E = 0,22$ e $I = 0,32$ (Santos et al. 2017),

enquanto que na população estudada os valores foram $H_E = 0,40$ e $I = 0,58$. Estes dados confirmam a variabilidade genética presente na população natural estudada e o seu potencial como fonte de material genético.

A diversidade genética dentro e entre as populações é fundamental para a conservação da biodiversidade. Melhor potencial de adaptação é esperado em populações com altos níveis de diversidade genética (Grenier et al. 2016). O sistema reprodutivo também contribui para a variabilidade, *H. speciosa* é uma espécie alógama, com flores hermafroditas e autoincompatíveis que são polinizadas por diferentes espécies de abelhas, mariposas e borboletas (Darrault e Schlindwein 2005). A autoincompatibilidade foi relatada também por Collevatti et al. (2016) em estudo de polinização cruzada entre variações botânicas da espécie. Maior diversidade genética é esperada nas populações naturais de espécies com essas características.

Os valores médios de N_a , N_e , H_E e I obtidos para a população estudada no presente trabalho, comparados com outros estudos para a espécie empregando marcadores ISSR, reportados na literatura, sugerem que a população apresenta alta diversidade genética. Portanto, é recomendável que os órgãos públicos atuem para minimizar a antropização da área, visando assegurar a manutenção desta população. Manejo florestal sustentável (MFS) é um tema altamente relevante na área florestal e na política ambiental (Wolfslehner et al. 2005). A elaboração de um plano de manejo sustentável para a espécie pode ser uma das estratégias adotadas, visto que a população está inserida em área de intenso extrativismo.

Um fator importante que contribui para a manutenção da diversidade genética em populações naturais é o fluxo gênico. Na literatura é relatada evidências de fluxo gênico entre áreas remanescentes naturais e restauradas da Mata Atlântica, além de níveis semelhantes de diversidade genética nessas áreas (Cordeiro et al. 2019; Schwarcz et al. 2018; Sujii et al. 2017).

A população estudada está próxima a uma área natural protegida, a Reserva Biológica Santa Isabel (Decreto nº 96.999, de 20 de outubro de 1988), na qual foi registrada a presença de *H. speciosa* (Oliveira e Landim 2014), o que pode ter contribuído para a manutenção do fluxo gênico e diversidade genética populacional. A possibilidade de ocorrência do fluxo gênico entre essas populações evidencia a importância das unidades de conservação que podem ser fonte de alelos para fragmentos previamente isolados, contribuindo para o aumento do seu tamanho efetivo.

A formação de grupos distintos pelos métodos de UPGMA e PCoA indica a existência de diversidade genética na população estudada. Estudos recentes apontam que o conhecimento da variabilidade genética das populações naturais permite inferir sobre a sustentabilidade da população (Cordeiro et al. 2019; Sujii et al. 2017), além de fornecer informações sobre as espécies e permitir a seleção de indivíduos mais divergentes ou mais semelhantes geneticamente conforme objetivo de estudo da pesquisa.

A diversidade genética das populações de plantas envolve interações de vários processos e características intrínsecas das espécies. Entre os processos tem-se a fragmentação de habitat, mutação, isolamento reprodutivo, deriva genética, fluxo gênico, seleção, entre outros (Hepsibha et al. 2010). Quanto às características intrínsecas, tem-se por exemplo, o mecanismo de dispersão de sementes, o sistema reprodutivo, que para *H. speciosa* ocorre pela polinização cruzada favorecendo, portanto, o surgimento de novas combinações gênicas e, consequentemente, maior diversidade.

Estudos de diversidade genética juntamente com a caracterização fenotípica são de grande relevância para compreender a genética e o desenvolvimento das plantas, sendo informações essenciais para a gestão de recursos genéticos. Para *H. speciosa*, frutífera com alto potencial, essas informações podem ser úteis no desenvolvimento do plano de manejo sustentável, obtenção de propágulos para a produção de mudas e seleção de indivíduos para programas de melhoramento florestal, ações estas indispensáveis para assegurar a manutenção de populações naturais da espécie.

5.5. Conclusões

Existe alta diversidade genética estimada pelos marcadores ISSR entre as árvores adultas analisadas e elevada variabilidade fenotípica para os frutos, sementes e plântulas.

Os frutos e as plântulas apresentaram características superiores quando comparados a outras populações da espécie, e as sementes, alto percentual de germinação.

Os achados obtidos neste estudo revelam o potencial da população para fornecer propágulos a serem utilizados em estratégias de conservação e melhoramento genético de *H. speciosa*.

5.6. Referências Bibliográficas

- Almeida GQ, Chaves LJ, Vieira MC, Ganga RMD (2019) Agronomic evaluation of a *Hancornia speciosa* Gomes germplasm collection from the Brazilian Cerrado. *Crop Breed Appl Biotechnol* 19:8–14. <https://doi.org/10.1590/1984-70332019v19n1a02>
- Alvares CA, Stape JL, Sentelhas PC, Gonçalves JLM, Sparovek G (2013) Köppen's climate classification map for Brazil. *Meteorol Zeitschrift* 22(6):711–728. <https://doi.org/10.1127/0941-2948/2013/0507>
- Bashiri H, Cheghamirza K, Arji I, Mahmodi N (2017) Assessing genetic diversity of *Pyrus* spp. in the central Zagros mountains based on morphological characters. *Genet. Resour. Crop Evol.* 64(2):391–404. <https://doi.org/10.1007/s10722-016-0367-2>
- Cardoso MC, Reis BL, Oliveira DS, Pinheiro-Sant'ana HM (2014) Mangaba (*Hancornia speciosa* Gomes) from the Brazilian Cerrado: nutritional value, carotenoids and antioxidants vitamins. *Fruits* 69:89–99. <https://doi.org/10.1051/fruits/2013105>
- Carvalho NM, Nakagawa J (2012) Sementes: ciência, tecnologia e produção. FUNEP, Jaboticabal
- Cavallari MM, Billot C, Bouvet JM, Favreau B, Zucchi MI, Palmier DA, Gimenes MA (2008) Isolation and characterization of microsatellite markers for *Casearia sylvestris* Sw (Salicaceae), a neotropical medicinal tree. *Mol Ecol Resour* 8(4):802–804. <https://doi.org/10.1111/j.1755-0998.2007.02069.x>
- Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. *Nat Rev Genet* 10(11):783–796. <https://doi.org/10.1038/nrg2664>
- Chitarra AB, Chitarra MIF (2005) Pós-colheita de frutas e hortaliças: Fisiologia e manuseio. Lavras: UFLA, 2005. 785 p.
- Cilma-Data.org. Climate data for world cities. Available in <https://pt.climate-data.org/info/sources/>. Acesso 8 set. 2020.
- Collevatti RG, Olivatti AM, Telles MP, Chaves LJ (2016) Gene flow among *Hancornia speciosa* (Apocynaceae) varieties and hybrid fitness. *Tree Genet Genomes* 12(4):74. <https://doi.org/10.1007/s11295-016-1031-x>
- Cordeiro EM, Macrini CM, Sujii PS, Schwarcz KD, Pinheiro JB, Rodrigues RR et al (2019) Diversity, genetic structure, and population genomics of the tropical tree *Centrolobium tomentosum* in remnant and restored Atlantic forests. *Conserv Genet* 20(5):1073–1085. <https://doi.org/10.1007/s10592-019-01195-z>
- Costa CF, Collevatti RG, Chaves LJ, Lima JS, Soares TN, Telles MPC (2017) Genetic diversity and fine-scale genetic structure in *Hancornia speciosa* Gomes (Apocynaceae). *Biochem Syst Ecol* 72:63–67. <https://doi.org/10.1016/j.bse.2017.03.001>
- Costa DFD, Vieira FDA, Fajardo CG, Chagas KPTD (2015) Genetic diversity and ISSR initiators selection in a natural population of mangaba (*Hancornia speciosa* Gomes) (Apocynaceae). *Rev Bras Frutic* 37(4):970–976. <https://doi.org/10.1590/0100-2945-246/14>
- Darrault RO, Schlindwein C (2005) Limited fruit production in *Hancornia speciosa* (Apocynaceae) and pollination by nocturnal and diurnal insects. *Biotropica* 37(3):381–388. <https://doi.org/10.1111/j.1744-7429.2005.00050.x>

- Dawson IK, Leakey R, Clement CR, Weber JC, Cornelius JP, Roshetko JM et al (2014) The management of tree genetic resources and the livelihoods of rural communities in the tropics: Non-timber forest products, smallholder agroforestry practices and tree commodity crops. *For Ecol Manag* 333:9–21. <https://doi.org/10.1016/j.foreco.2014.01.021>
- EMBRAPA. Árvore do conhecimento Mangaba. Disponível em: <https://www.agencia.cnptia.embrapa.br/gestor/mangaba/Abertura.html> Acesso em: 28 out. 2021
- Farjado CG, Costa DF, Chagas KPT, Vieira FA (2018) Genetic diversity in natural populations of *Hancornia speciosa* Gomes: implications for conservation of genetic resources. *Ciênc Agrotec* 42(6):623–630. <https://doi.org/10.1590/1413-70542018426019018>
- Freitas BAL (2016) Conhecimento local, diversidade morfogênica como subsídios para conservação da mangaba. Dissertação, Universidade Federal de Sergipe
- Ganga RMD, Ferreira GA, Chaves LJ, Naves RV, Nascimento JLD (2010) Caracterização de frutos e árvores de populações naturais de *Hancornia speciosa* Gomes do cerrado. *Rev Bras Frutic* 32(1):101–113. <https://doi.org/10.1590/S0100-29452010005000019>
- Giles JAD, Oliari LSO, Rocha ACB, Schmildt ER, Silva W, França JM (2016) Correlações entre características físicas, químicas e físico-químicas de frutos de cirigueleira. *Revista Agro@mbiente On-line* 10(1):30–35. <https://doi.org/10.18227/1982-8470ragro.v10i1.2763>
- Gonçalves LGV, Andrade FR, Marimon Junior BH, Schossler TR, Lenza E, Marimon BS (2013) Biometria de frutos e sementes de mangaba (*Hancornia speciosa* Gomes) em vegetação natural na região o leste de Mato Grosso. *Brasil Rev Ciênc Agrár* 36(1):31–40
- Grenier S, Barre P, Litrico I (2016) Phenotypic plasticity and selection: nonexclusive mechanisms of adaptation. *Scientifica*. <https://doi.org/10.1155/2016/7021701>
- Heidari P, Rezaei M, Sahebi M, Khadivi A (2019) Phenotypic variability of *Pyrus boissieriana* Buhse: Implications for conservation and breeding. *Sci Hortic* 247:1–8. <https://doi.org/10.1016/j.scienta.2018.11.075>
- Hepsibha BT, Premalakshmi V, Sekar T (2010) Genetic diversity in *Azima tetraacantha* (Lam) assessed through RAPD analysis. *Indian J Sci* 3(2):170–173
- IBGE. Produção da Extração Vegetal e da Silvicultura. Disponível em: <https://sidra.ibge.gov.br/pesquisa/pevs/quadros/brasil/2018>. Acesso em 30 jul. 2020.
- Jimenez HJ, Martins LSS, Montarroyos AVV, Silva Junior JF, Alzate-Marin AL, Moraes Filho RM (2015) Genetic diversity of the Neotropical tree *Hancornia speciosa* Gomes in natural populations in Northeastern Brazil. *Genet Mol Res* 14(4):17749–17757. <https://doi.org/10.4238/2015.December.21.48>
- Kim DH, Sexton JO, Townshend JR (2015) Accelerated deforestation in the humid tropics from the 1990s to the 2000s. *Geophys Res Lett* 42(9):3495–3501. <https://doi.org/10.1002/2014GL062777>
- Lima JP, Azevedo L, Souza NJ, Nunes EE, Boas EVBV (2015a) First evaluation of the antimutagenic effect of mangaba fruit in vivo and its phenolic profile identification. *Food Res Int* 75:216–224. <https://doi.org/10.1016/j.foodres.2015.05.045>
- Lima JP, Fante CA, Pires CRF, Nunes EE, Alves RR, Elias HHS et al (2015b) The antioxidative potential and volatile constituents of mangaba fruit over the storage period. *Sci Hortic* 194:1–6. <https://doi.org/10.1016/j.scienta.2015.05.035>
- Lorenzoni RM, Soares TCB, Santiago VF, Silva JA, Coelho RI (2014) Utilização de marcadores ISSR na avaliação da divergência genética entre acessos de biribazeiro. *Rev Bras Frutic* 36:251–257
- Lúcio DM, Silva RAR, Chagas KPT, Vieira FA (2019) Biometry of fruits and seeds of *Hancornia speciosa* Gomes (Apocynaceae) in natural populations. *Agricult Forest* 65(1):99–110. <https://doi.org/10.17707/AgricultForest.65.1.11>
- Luz GA, Santos JA, Oliveira KP, Martins PP, Valente SES, Maia MCC, Lima PSC (2020) Genetic diversity and population structure of mangabeira (*Hancornia speciosa*) estimated using ISSR markers. *Acta Sci Biol Sci* 42(1):e53606. <https://doi.org/10.4025/actascibiolsci.v42i1.53606>

- Lv J, Li C, Zhou C, Chen J, Li F, Weng Q, Li M, Wang Y, Chen S, Chen J, Gan S (2020) Genetic diversity analysis of a breeding population of *Eucalyptus cloeziana* F. Muell. (Myrtaceae) and extraction of a core germplasm collection using microsatellite markers. *Ind Crop Prod* 145:112157. <https://doi.org/10.1016/j.indcrop.2020.112157>
- Mitchell J, Johnston IG, Bassel GW (2017) Variability in seeds: biological, ecological, and agricultural implications. *J Exp Bot* 68(4):809–817. <https://doi.org/10.1093/jxb/erw397>
- Nascimento RS, Cardoso JA, Coccoza FD (2014) Caracterização física e físico-química de frutos de mangabeira (*Hancornia speciosa* Gomes) no oeste da Bahia. *Rev Bras Eng Agríc Ambient* 18(8):856–860. <https://doi.org/10.1590/1807-1929/agriambi.v18n08p856-860>
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70(12):3321–3323
- Nienhuis J, Tivang J, Sckroch P, Santos JB (1995) Genetic relationships among cultivars and lines of lima bean (*Phaseolus lunatus* L.) as measured by RAPD markers. *J Am Soc Hort Sci* 120(2):300–306
- Novello M, Viana JPG, Alves-Pereira A, Silvestre EA, Nunes HF, Pinheiro JB et al (2018) Genetic conservation of a threatened Neotropical palm through community-management of fruits in agroforests and second-growth forests. *For Ecol Manag* 407:200–209. <https://doi.org/10.1016/j.foreco.2017.06.059>
- Oliveira EVS, Landim MF (2014) Caracterização fitofisionômica das restingas da Reserva Biológica de Santa Isabel, litoral norte de Sergipe. *Sci Plena* 10(10):109902-1
- Osorio-Guarín JA, Berdugo-Cely J, Coronado RA, Zapata YP, Quintero C, Gallego-Sánchez G, Yockteng R (2017) Colombia a source of cacao genetic diversity as revealed by the population structure analysis of germplasm bank of *Theobroma cacao* L. *Front Plant Sci* 8:1994. <https://doi.org/10.3389/fpls.2017.01994>
- Pammenter NW, Berjak P (2000) Aspects of recalcitrant seed physiology. *Rev. Bras. Fisiol. Veg.* 12(Edicao Especial):56–69
- Peakall ROD, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28(19):2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Pinheiro EA, Coimbra RR, Silva KLF, Ferreira WM (2018) Characterization and phenotypic variability in natural populations of mangabeira in the state of Tocantins. *Braz Rev Caatinga* 31(3):560–571. <https://doi.org/10.1590/1983-21252018v31n304rc>
- Pinto RJ, Mapeli NC, Cremon C, Silva EF (2014) Germinação e crescimento inicial de mangaba (*Hancornia speciosa* Gomes) em função de preparados homeopáticos *Carbo vegetabilis* e dias após o despolpamento para semeadura. *Agrarian* 7(24):244–250
- Puiatti GA, Cecon PR, Nascimento M, Nascimento ACC, Finger FL, Puiatti M et al (2014) Comparação dos métodos de agrupamento de Tocher e UPGMA no estudo de divergência genética em acessos de alho. *Revista da Estatística da Universidade Federal de Ouro Preto* 3(3):275–279
- R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Rajjou L, Duval M, Gallardo K, Catusse J, Bally J, Job C, Job D (2012) Seed germination and vigor. *Annu Rev Plant Biol* 63:507–533. <https://doi.org/10.1146/annurev-arplant-042811-105550>
- Rohlf FJ (2005) NTSYS-pc: numerical taxonomy and multivariate analysis system, version 2.20. Applied Biostatistics, New York
- Santos HG, Jacomine PKT, Anjos LHC, Oliveira VA, Lumbreras JF, Coelho MR et al (2018) Sistema Brasileiro de Classificação de Solos. Embrapa, Brasília
- Santos PCG, Alves EU, Guedes RS, Silva KB, Cardoso EA, Lima CR (2010) Qualidade de sementes de *Hancornia speciosa* Gomes em função do tempo de secagem. *Semina Ciênc Agrár* 31(2):343–351

- Santos PS, Freitas LS, Santana JGS, Muniz EN, Rabbani ARC, Silva AVC (2017) Genetic diversity and the quality of Mangabeira tree fruits (*Hancornia speciosa* Gomes - Apocynaceae), a native species from Brazil. *Sci Hortic* 226:372–378. <https://doi.org/10.1016/j.scienta.2017.09.008>
- Schwarcz KD, Silvestre EA, Campos JB, Sujii PS, Grando C, Macrini CMT et al (2018) Shelter from the storm: Restored populations of the neotropical tree *Myroxylon peruiferum* are as genetically diverse as those from conserved remnants. *For Ecol Manag* 410:95–103. <https://doi.org/10.1016/j.foreco.2017.12.037>
- Seoane CES, Sebbenn AM, Kageyama PY (2005) Sistema de reprodução em duas populações naturais de *Euterpe edulis* M. sob diferentes condições de fragmentação florestal. *Scientia Forestalis* 69:13–24
- Sergipe (2016). Digital Atlas on Water Resources of Sergipe, Aracaju, SE: state secretariat for the environment and water resources. CD-ROM
- Sergipe. Law nº. 7.082, December 16, 2010. Acknowledge of the Catadoras de Mangaba as a differentiated cultural group and establishes self-recognition as a criterion of law and provides other measures. Disponível em: <http://www.catadorasdemangaba.com.br/includes/cdmangaba/13/Lei288-2010.pdf>. Acesso: 22 jun. 2020.
- Silva AVC, Amorim JAE, Melo MFV, Ledo AS, Rabbani ARC (2017) Genetic diversity of remaining populations of mangaba (*Hancornia speciosa* Gomes) in restingas of Brazil. *J Agric Sci.* <https://doi.org/10.5539/jas.v9n2p46>
- Silva AVC, Nascimento ALS, Soares ANR, Rabbani ARC, Silva Júnior JF, Ledo AS (2019) Identification and preliminary characterization of early fruiting mangabeira (*Hancornia speciosa* - Apocynaceae). *Revista Agro@mbiente On-line* 13:115–128. <https://doi.org/10.18227/1982-8470ragro.v13i0.5462>
- Silva FAM, Assad ED, Evangelista BA (2008) Caracterização Climática do Bioma Cerrado. In Sano, SM, Almeida SP, Ribeiro JF. (Orgs.). *Cerrado: ecologia e flora*. Brasília: Embrapa, pp 70–88.
- Sneath PH, Sokal RR (1973) Numerical taxonomy. The principles and practice of numerical classification.
- Soares ANR, Melo MIFV, Vitória MF, Silva AVC (2015) Physiological quality of mangaba seeds submitted to drying. *Afr J Agric Res* 10(52):4806–4813. <https://doi.org/10.5897/AJAR2015.10424>
- Soares ANR, Muniz ADS, Muniz EN, Melo MDV, Santos OS, Ledo ADS (2019) Biometry, emergence and initial growth of accessions and mangaba progenies. *J Agric Sci.* <https://doi.org/10.5539/jas.v11n4p436>
- Sousa VA, Reeves PA, Reilley A, Aguiar AV, Stefenon VM, Richards CM (2020) Genetic diversity and biogeographic determinants of population structure in *Araucaria angustifolia* (Bert.) O. Ktze *Conserv Genet* 21:217–229. <https://doi.org/10.1007/s10592-019-01242-9>
- Souza-Sobreira FB, Souza GB, Rosado CCG, Miranda FD et al (2015) Genetic diversity in three natural populations of *Pitcairnia flammea* (L.) John (Bromeliaceae) estimated by ISSR markers. *Genet Mol Res* 14:15892–15901. <https://doi.org/10.4238/2015.December.3.1>
- Sujii PS, Schwarcz KD, Grando C, Silvestre EA, Mori GM, Brancalion PH, Zucchi MI (2017) Recovery of genetic diversity levels of a Neotropical tree in Atlantic Forest restoration plantations. *Biol Conserv* 211:110–116. <https://doi.org/10.1016/j.biocon.2017.05.006>
- Vencovsky R, Crossa J (2003) Measurements of representativeness used in genetic resources conservation and plant breeding. *Crop Sci* 43(6):1912–1921. <https://doi.org/10.2135/cropsci2003.1912>
- Vieira FDA, Gusmão E (2008) Biometria, armazenamento de sementes e emergência de plântulas de *Talisia esculenta* Radlk (Sapindaceae). *Ciênc Agrotec* 32(4):1073–1079. <https://doi.org/10.1590/S1413-70542008000400006>

- Whitlock MC, Ingvarsson PK, Hatfield T (2000) Local drift load and the heterosis of interconnected populations. *Heredity* 84(4):452–457
- Wolfslehner B, Vacik H, Lexer MJ (2005) Application of the analytic network process in multi-criteria analysis of sustainable forest management. *For Ecol Manag* 207(1–2):157–170. <https://doi.org/10.1016/j.foreco.2004.10.025>
- Yokomizo GKI, Maia MCC, Trindade CF (2017) Morphological dissimilarity among mangabeira tree populations from Amapá and Paraíba. *Brazil Rev Caatinga* 30(2):521–529. <https://doi.org/10.1590/1983-21252017v30n228rc>
- Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20(2):176–183. <https://doi.org/10.1006/geno.1994.1151>
- Zucchi MI, Sujii PS, Mori GM, Viana JP, Grando C, Silvestre EDA et al (2018) Genetic diversity of reintroduced tree populations in restoration plantations of the Brazilian Atlantic Forest. *Restor Ecol* 26(4):694–701. <https://doi.org/10.1111/rec.12620>
- Zuffo AM, Busch A, Steiner F, Alves CZ, Alcântara Neto F, Santos MA et al (2019) Biometric characteristics of fruits, seeds and plants of *Hancornia speciosa* Gomes (Apocynaceae). *Aust J Crop Sci* 13(4):622. <https://doi.org/10.21475/ajcs.19.13.04.p1651>

6. ARTIGO 3

GENÔMICA E ESTRUTURA GENÉTICA DE POPULAÇÕES DE *Hancornia speciosa*: FONTE DE ALIMENTOS FUNCIONAIS E FÁRMACOS

Periódico (a ser submetido): Biological Conservation

RESUMO

Para identificar áreas com adequabilidade ambiental e estimar a diversidade genômica e estrutura genética de populações naturais da árvore *H. speciosa* foi combinada uma abordagem integrada de modelos de distribuição de espécies (SDM) e genômica populacional baseada em polimorfismos de nucleotídeo único (SNPs). Com base em registros de ocorrências e variáveis bioclimáticas usando SDM, foram identificadas áreas de maior adequabilidade e amostradas seis populações naturais, que foram genotipadas por sequenciamento para análise da diversidade genômica e estrutura populacional. As áreas com maior adequabilidade (0,75 a 1,00) estão na Mata Atlântica (litoral nordestino) e Cerrado (Brasil central). Foram identificados 5.046 SNPs (4.691 neutros e 355 *outliers*), os quais permitiram inferir que existe alta diversidade e excesso de heterozigotos ($f = -0,123$ a $-0,066$) em todas as populações e 315 alelos privados. Maior variação genética ocorre dentro das populações. Foi verificada diferenciação genética entre as populações do Nordeste e Centro-Oeste ($F_{ST} = 0,442$ a $0,538$). E ao avaliar a estruturação genética, verificou-se que a população de Coxim apresentou diferenciação acentuada em relação às demais populações, sobretudo considerando a análise de discriminantes principais. O conjunto de SNPs *outliers* revela uma estruturação genética formada por três grupos que estão correlacionados com os diferentes biomas em que as populações estão inseridas. Para dois SNPs *outlier* foi possível identificar anotação funcional. Esses resultados contribuem para avançar no conhecimento sobre a espécie e no desenvolvimento de estratégias para seu uso sustentável e conservação.

Palavras-chave: genômica populacional, genotipagem por sequenciamento, mangaba, modelo de distribuição de espécie, SNP.

ABSTRACT

Título: Genomics and genetic structure of *Hanconia speciosa* populations: a source of functional foods and pharmaceuticals

To identify areas with environmental suitability and to estimate the genomic diversity and genetic structure of the *H. speciosa* tree natural populations, an integrated approach of species distribution models (SDM) and population genomics based on single nucleotide polymorphisms (SNPs) were combined. Based on occurrence records and bioclimatic variables using SDM, areas of greatest suitability were identified, and six natural populations were sampled, which were genotyped by sequencing in genomic diversity and population structure analyze. The areas with the highest suitability (0.75 to 1.00) are the Atlantic Forest (northeastern coast) and Cerrado (central Brazil). We identified 5,046 SNPs (4,691 neutrals and 355 outliers), which allowed us to infer that there is high diversity and excess of heterozygotes ($f = -0.123$ to -0.066) in all populations, and 315 private alleles. Greater genetic variation occurs within populations. Genetic differentiation was verified between the Northeast and Midwest populations ($F_{ST} = 0.442$ to 0.538). When evaluating the genetic structure, it was found that the Coxim population presented differentiation in comparison to other populations, especially considering the analysis of main discriminators. The set of outlier SNPs reveals a genetic structure formed by three groups that are correlated with the different biomes in which the populations are inserted. For two outlier SNPs it was possible to identify functional annotation. These results contribute to advancing knowledge about the species and developing of strategies for its sustainable use and conservation.

Key-words: population genomics, sequencing genotyping, mangaba, species distribution model, SNP.

6.1. Introdução

As florestas são ecossistemas dinâmicos com uma variedade de funções, cobrem atualmente 31% da área terrestre global (United Nations, 2021) e fornecem inúmeros serviços ecossistêmicos, sejam eles de provisão, apoio, reguladores e culturais (Brockerhoff et al., 2017). Especialmente em regiões tropicais, representam o tipo de habitat mais rico em espécies e abrigam a maior parte da biodiversidade mundial (Lewis et al., 2015).

Entre as muitas espécies tropicais tem-se *Hanconia speciosa* (Apocynaceae), a qual é uma árvore frutífera e medicinal com importância econômica e ecológica. Suas populações possuem ampla distribuição geográfica na América do Sul (Tropicos, 2021). O gênero *Hanconia* é monotípico e existe alta variabilidade fenotípica nas populações naturais com diferenças morfológicas contrastantes em folhas e frutos para populações localizadas em regiões geográficas diferentes (Monachino, 1945).

Considerando características morfológicas, foram identificadas seis variedades botânicas para *H. speciosa*, sendo elas: *Hanconia speciosa* var. *speciosa* (Nordeste ao Centro-Oeste e Norte); *H. speciosa* var. *maximilliani* e *H. speciosa* var. *lundii* (Sudoeste); *H. speciosa* var. *cuyabensis* (Mato Grosso e Centro-Oeste); *H. speciosa* var. *gardneri* e *H. speciosa* var. *pubescens* (Brasil central) (Monachino, 1945). Entretanto, pela revisão botânica atual, apenas duas variedades são reconhecidas: *H. speciosa* var. *pubescens* e *H. speciosa* var. *speciosa* (Flora do Brasil, 2021).

Estudos que investigam a diversidade genética e a diferenciação entre populações de *H. speciosa* amplamente distribuídas são menos frequentes (Chaves et al., 2020; Collevatti et al.,

2018) e seus resultados não são congruentes com as classificações botânicas existentes (Collevatti et al., 2018), portanto o número de variedades botânicas permanece controverso.

Devido à ampla distribuição geográfica, algumas das diferenciações morfológicas podem ser devido a respostas à adaptação local (Le Corre e Kremer, 2012), nesse sentido, a genômica de populações pode fornecer maior poder de resolução para identificar a diversidade e divergência dentro e entre populações, isto porque a diferenciação genética pode ser influenciada por forças que atuam em diferentes regiões do DNA (Díaz et al., 2021).

Os estudos genômicos permitem identificar e separar os efeitos decorrentes de deriva, fluxo gênico e endogamia que informam de maneira confiável sobre a demografia da população e a história filogenética e permitem identificar também efeitos específicos como seleção, mutação, acasalamento e recombinação que ajudam a identificar genes importantes para a aptidão e adaptação (Black et al., 2001; Luikart et al., 2003)

As plataformas de sequenciamento de próxima geração (NGS) possibilitaram avançar nos estudos genômicos florestais (Mora-Sala et al., 2019; Novaes et al., 2008; Tosso et al., 2018), sobretudo pela genotipagem por sequenciamento (GBS) (Elshire et al., 2011; Poland et al., 2012) que permite a identificação e sequenciamento de milhares de SNPs no genoma, inclusive para espécies não-modelo, como *H. speciosa*, fornecendo um grande número de marcadores genéticos para estudos populacionais a custos relativamente baixos (Ulaszewski et al., 2021).

Adicionalmente, outras abordagens como os modelos de distribuição de espécies (SDMs) que utilizam associações entre ocorrências conhecidas de espécies e condições ambientais para estimar as distribuições geográficas potenciais (Feng et al., 2019), podem contribuir para guiar levantamentos de campo (Fois et al., 2018) e predizer áreas com adequabilidade ambiental para espécies de interesse. Combinados com análises genômicas, oferecem contribuições promissoras e podem nortear as estratégias de manejo, conservação e melhoramento.

Estudos utilizando SDMs foram realizados para *H. speciosa* visando investigar os efeitos das mudanças climáticas na distribuição e produção de frutos no Brasil (Nabout et al., 2016) e para conhecer a paleodistribuição da espécie (Collevatti et al., 2018). Entretanto, até o momento não foram relatados estudos empregando GBS para a identificação e genotipagem de SNPs.

Considerando as controvérsias a respeito das variedades botânicas de *H. speciosa* e o reduzido número de estudos com populações amplamente distribuídas, uma abordagem integrada de SDM e genômica populacional baseada em SNPs foi empregada para identificar áreas com adequabilidade ambiental, a diversidade genômica e estrutura genética das populações naturais de *H. speciosa*.

6.2. Material e Métodos

6.2.1 Modelos de distribuição de espécies

Modelos de distribuição de espécies foram utilizados para identificar as áreas com maior adequabilidade ambiental para *H. speciosa* e gerar o mapa de distribuição potencial, o que posteriormente norteou a amostragem de populações em diferentes regiões no Brasil para estudos genômicos. Os registros de ocorrência da espécie foram obtidos junto a *Global Biodiversity Information Facility* - GBIF (<https://www.gbif.org/>) acessados em abril de 2021. GBIF é uma rede internacional de dados mantida por governos mundiais que foi escolhida por compilar em seu conjunto de dados informações depositadas em diferentes bases (GBIF, 2021).

Foram obtidos 1.033 registros de ocorrência, que passaram por limpeza e o viés de amostragem foi removido utilizando o pacote *CoordinateCleaner* R (Zizka et al., 2019), o que possibilitou a remoção de coordenadas atribuídas às capitais, coordenadas com latitude igual à

longitude, coordenadas iguais à exatamente zero, coordenadas baseadas em centroides de províncias e referências corrigidas de países.

Os registros foram filtrados para as ocorrências entre 1970 e 2021 e usando o pacote R *ecospat* (Di Cola et al., 2017) realizou-se a desagregação espacial com 2,5 arco-minutos. Após a limpeza restaram 196 registros de ocorrência incluídos na análise que foram mapeados em uma grade de células de $0,5^\circ \times 0,5^\circ$ (longitude $\times$ latitude), abrangendo a América do Sul (Figura 1).



Figura 1. Registro da ocorrência de *Hancornia speciosa* na América do Sul após limpeza das coordenadas.

Para construir o modelo de distribuição foram utilizadas 19 variáveis bioclimáticas obtidas no *WorldClim* (*WorldClim* versão 2 - <https://worldclim.org/data/worldclim21.html>, acessadas em abril de 2021), um banco de dados climáticos e meteorológicos global (Fick; Hijmans, 2017). Essas variáveis têm sido amplamente utilizadas em SDM (Porfirio et al., 2014) e foram convertidas para uma resolução espacial de $0,5^\circ$ (Figura 2). Para remover a colinearidade entre as variáveis bioclimáticas, estas foram submetidas à correlação de Spearman, e com o auxílio da função *findCorrelation* do pacote R *caret* (Kuhn, 2008) foram selecionadas somente aquelas com correlação abaixo de 0,7. Estas foram variáveis de temperatura: bio02 - Faixa Diurna Média (Média do mês (temp. Máx. - temp. Mín.)), bio03 - Isotermalidade e bio08 - Temperatura Média do Trimestre Mais Úmido; e de precipitação bio15 - Sazonalidade de Precipitação (coeficiente de variação) e bio18 - Precipitação do Trimestre Mais Quente (Fick e Hijmans, 2017).

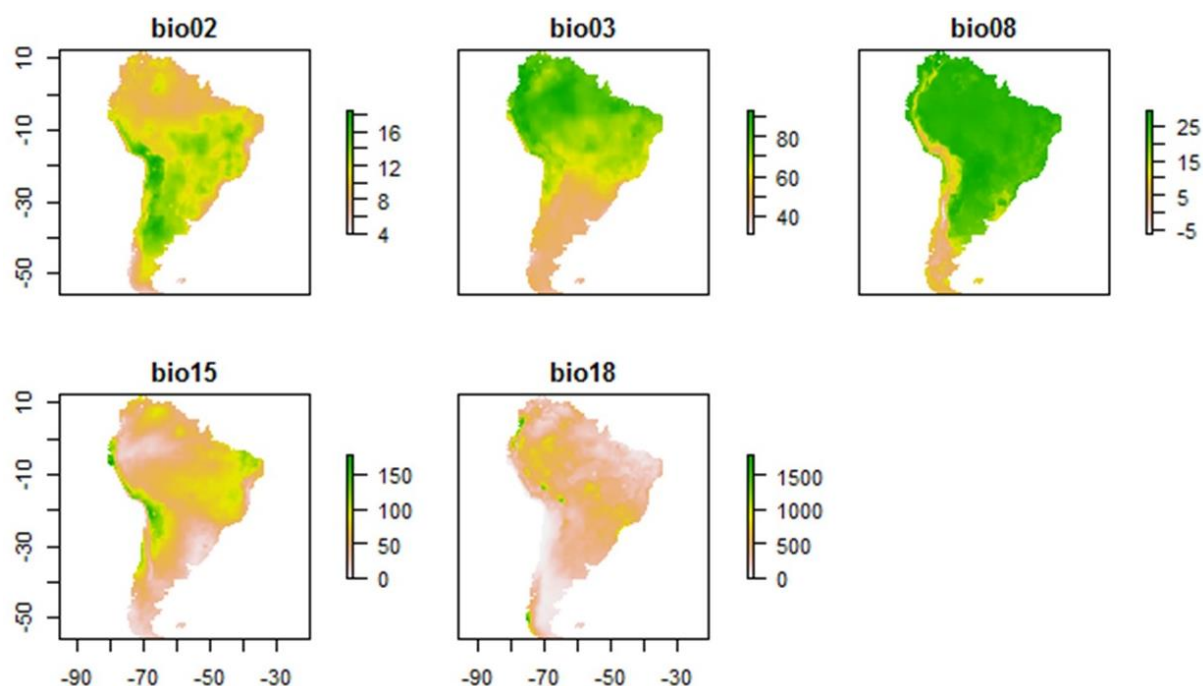


Figura 2. Variáveis bioclimáticas que foram incluídas no estudo ($r \leq 0,70$).

Para gerar o modelo de distribuição foram combinados cinco algoritmos matemáticos, considerando suas especificidades. A combinação deles possibilita a predição de um modelo com maior confiabilidade (Araújo e New, 2007; Diniz-Filho et al., 2009).

Os algoritmos utilizados foram: Bioclim (Nix, 1986), Domain (Distância/Domínio de Gower) (Carpenter et al., 1993) e Distância de Mahalanobis (Farber e Kadmon, 2003), que são baseados apenas em dados de presença da espécie; e os algoritmos de aprendizado de máquina, Entropia máxima (MaxEnt) (Phillips e Dudík, 2008) baseado em dados de presença-fundo, e Máquina de Vetores de Suporte (Support Vector Machine - SVM) (Tax e Duin, 2004) baseado em dados de presença-ausência.

Para avaliar os modelos gerados, foi adotada a abordagem de partição dupla com 70% e 30% de treino e teste, respectivamente. Esta abordagem particionada de ocorrência foi utilizada para calcular os valores da verdadeira estatística de habilidade (*True Skill Statistics* - TSS) de cada modelo. O TSS leva em consideração os erros de omissão e comissão, varia de -1 a +1 (+1 indica concordância perfeita; e valores de zero ou menores indicam um desempenho não melhor do que aleatório) (Allouche et al., 2006).

Essa etapa foi realizada de forma randomizada, 10 vezes para cada algoritmo e foram obtidos 50 mapas (5 algoritmos $\times$ 10 vezes). Para transformar mapas contínuos em mapas binários foram utilizadas a máxima sensibilidade e especificidade. Por fim, mapas gerados para o mesmo algoritmo (réplicas) e os mapas entre os algoritmos com um TSS maior que 0,5 foram sobrepostos para prever o mapa final (conjunto) de distribuição potencial.

6.2.2 Populações amostradas para análises moleculares

O material vegetal foi prospectado em seis populações naturais de *H. speciosa*, localizadas em áreas com adequabilidade ambiental para a espécie de acordo com o mapa predito, na região Nordeste, em áreas do bioma Mata Atlântica (Estância-SE, Pirambu-SE, Marechal Deodoro-AL e Conde-BA) e na região Centro-Oeste, no bioma Cerrado (Barra do

Garças-MT) e em área de transição Cerrado/Pantanal (Coxim-MS) (Figura 3). *H. speciosa* é uma das espécies consideradas prioritárias para o desenvolvimento de pesquisas nestas regiões (Coradin et al., 2018; Vieira et al., 2016).

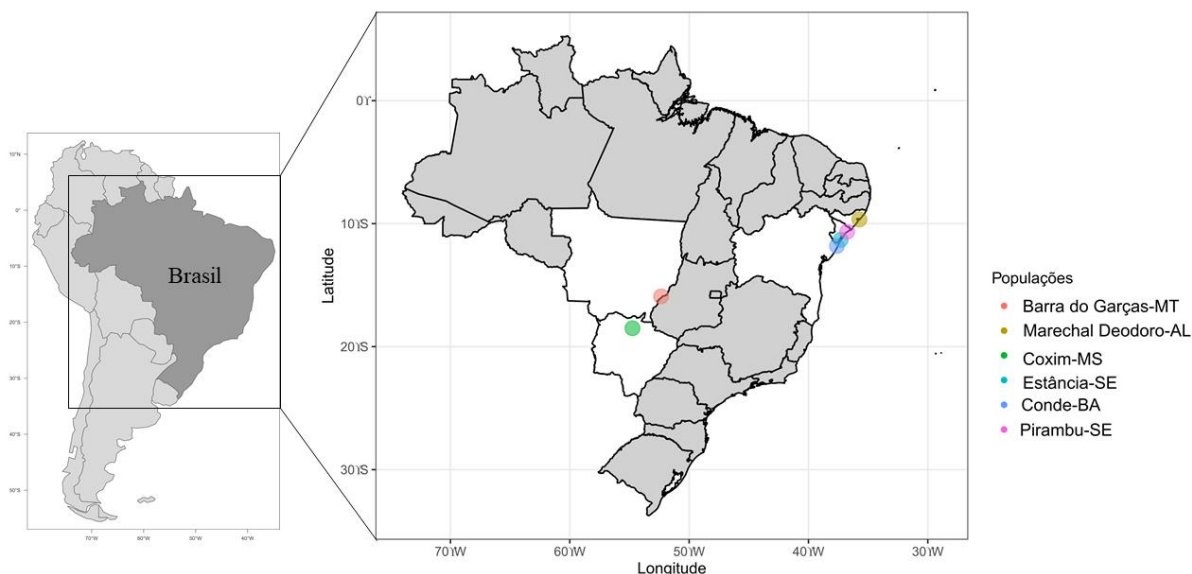


Figura 3. Mapa ilustrando as populações naturais de *Hancornia speciosa* que foram amostradas.

O presente estudo foi registrado no Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado (SISGEN) sob o número A21EE65, de acordo com a Lei 13.123/2015 (Brasil, 2015).

6.2.3 Extração do DNA, preparação e sequenciamento da biblioteca genômica

Folhas jovens de 86 indivíduos (Estância - 18, Pirambu - 15, Conde - 20, Marechal Deodoro - 10, Coxim - 09 e Barra do Garças - 14) foram coletadas e armazenadas a -80 °C até o momento da extração do DNA. O DNA genômico total foi extraído utilizando o protocolo de Doyle e Doyle (1990). A qualidade e a quantidade de DNA foram avaliadas em gel de agarose a 1% (p/v), corado com *SYBR safe DNA Gel Stain* (Invitrogen), por comparação visual com amostras de DNA do fago λ (Invitrogen) de concentrações conhecidas. Após a quantificação, as amostras foram normalizadas para concentração de 30 ng/ μ L.

A biblioteca genômica foi desenvolvida utilizando a técnica de Genotipagem por Sequenciamento (*Genotyping-by-Sequencing, GBS*) de acordo com o protocolo de dupla digestão descrito por Poland et al. (2012). O DNA genômico de cada indivíduo foi digerido por duas enzimas de restrição, *Nsi I* e *Mse I* (definidas previamente por meio de testes considerando o tamanho do genoma e a quantidade de fragmentos). Posteriormente, os fragmentos gerados na digestão para cada indivíduo foram ligados a adaptadores específicos para a tecnologia Illumina (contendo as sequências de código de barras e complementares aos *primers* Illumina™ para sequenciamento). Após a ligação dos adaptadores, as amostras foram então combinadas em um único pool (multiplex), que foi purificado usando o kit de purificação de PCR QIAquick (QIAGEN) e amplificado por PCR, visando enriquecer os fragmentos em ambos os locais de restrição. Após esta etapa, a biblioteca foi novamente purificada, validada e quantificada em equipamento Invitrogen™ Fluorômetro Qubit™ V3.0 (Thermo Fisher Scientific, Carlsbad, CA, EUA) e então foram sequenciadas na plataforma Illumina (Illumina, Inc., San Diego, CA).

6.2.4 Processamento das sequências, filtragem e identificação dos SNPs

O processamento das leituras de sequências brutas, a filtragem e identificação dos SNPs foram realizadas com o programa Stacks v. 1.42 (Catchen et al., 2011). Inicialmente foi utilizado o módulo *process_radtags* para demultiplexar as amostras e remover as leituras de baixa qualidade. Como não existe genoma de referência para *H. speciosa*, foi utilizado o *pipeline denovo*.

O módulo *ustacks* foi utilizado para o alinhamento das sequências e formação das pilhas de sequências homólogas e curtas (*stacks*) para identificar locos com nucleotídeos polimórficos. Esta etapa foi realizada para cada amostra separadamente com os seguintes parâmetros: profundidade mínima de sequenciamento ($-m$) = 3 e distância máxima entre os *stacks* ($-M$) = 2. Posteriormente, com o módulo *cstacks*, os locos de cada indivíduo foram agrupados em um catálogo, permitindo uma distância máxima de dois nucleotídeos ($-n$ 2) entre os locos de cada indivíduo. Os grupos de *stacks* foram comparados com o catálogo usando *sstacks* para determinar o conjunto de locos de cada indivíduo.

Os locos com valores de probabilidade logarítmica média menor que -10 foram removidos como o módulo de correção *rxstacks* e os SNPs foram filtrados pelo módulo *populations*, usando os seguintes parâmetros: frequência do alelo menos frequente ($LFA \geq 0,01$); profundidade mínima dos *stacks* ($m \geq 3$) e retenção de um marcador único por tag, ocorrendo em pelo menos 75% dos indivíduos em cada população.

O conjunto total de SNPs foi avaliado para identificação e separação de locos neutros e *outliers* (locos supostamente sob seleção). Essa análise ocorreu como base nas relações F_{ST} (índice de fixação que informa sobre a diferenciação genética entre populações) e H_E (heterozigosidade esperada) esperadas em um modelo de migração de ilha (Wright, 1965) que identifica os locos com F_{ST} excessivamente alto ou baixo (locos *outlier*), utilizando o pacote R *fsthet*. Este método usa quantis suavizados com base na distribuição empírica de $F_{ST} - H_E$ para identificar *outliers*. Os intervalos de confiança suavizados foram gerados com 1.000 réplicas de *bootstrap*, e os *outliers* foram identificados como os mais extremos nos dois lados da distribuição, considerando intervalo de confiança de 0,95 (Flanagan e Jones, 2017).

6.2.5 Diversidade e estrutura genômica

A diversidade genômica populacional foi estimada com base em marcadores SNPs neutros, considerando os seguintes parâmetros populacionais: número médio de alelos (A), número de alelos privados (A_p), heterozigosidade observada (H_O) e esperada (H_E), Índice de Shannon (I) e o coeficiente de endogamia de Wright (1965) (f) que foram estimados usando software GenAlEx v. 6.5 (Peakall e Smouse, 2012) e plotados com o pacote R *ggplot* (Wickham, 2016).

A distribuição da variação genômica entre e dentro das populações naturais de *H. speciosa* foi avaliada por meio de uma análise de variância molecular (AMOVA) e a diferenciação genética foi estimada por meio dos valores de F_{ST} de pares (Weir e Cockerham, 1984) e suas significâncias testadas com 9.999 permutações utilizando o software GenAlEx v. 6.5 (Peakall e Smouse, 2012). O resultado do F_{ST} de pares (Weir e Cockerham, 1984) foi plotado com o pacote R *heatmap3* (Zhao et al., 2014).

Para inferência da estrutura populacional duas abordagens foram utilizadas: a abordagem baseada em Modelo Bayesiano e a Análise Discriminante de Componentes Principais realizadas para os dois conjuntos de dados, marcadores SNP neutros e marcadores SNP *outliers*.

A inferência da estrutura populacional pela análise bayesiana foi realizada com o software *Structure* 2.3.2 (Pritchard et al., 2000) que permite determinar o número aproximado de agrupamentos genéticos (K) e atribuir indivíduos ao agrupamento mais apropriado.

Os indivíduos são inicialmente atribuídos aleatoriamente a um número predeterminado de grupos. Posteriormente, as frequências variantes são estimadas em cada grupo e os indivíduos reatribuídos com base nessas estimativas de frequência. Esse processo se repete

por muitas vezes e resulta em uma convergência progressiva em direção a estimativas de frequência de alelos confiáveis em cada população e probabilidades de adesão de indivíduos a uma população. Essa abordagem tem como suposição que os locos não estão ligados e as populações estão em Equilíbrio de Hardy-Weinberg (HWE) (Porrás-Hurtado et al., 2013; Pritchard et al., 2000).

A análise foi realizada com base no modelo de mistura com frequências alélicas correlacionadas. Os modelos foram testados para valores de K variando de 1 a 10, com 10 execuções independentes por valor de K. Para cada simulação, o período inicial (*burn-in*) foi de 250.000 etapas seguidas por 250.000 interações *Markov Chain Monte Carlo* (MCMC). Para determinar o valor mais provável de K, o método ΔK , que expressa a taxa de mudança de segunda ordem do logaritmo da probabilidade marginal dos dados (Evanno et al., 2005), foi utilizado conforme implementado na ferramenta on-line *Harvester Structure* (Earl e VonHoldt, 2012).

A análise de discriminantes principais (*Discriminant Analysis of Principal Components* - DAPC) foi realizada com o pacote R *Adegenet* (Jombart e Ahmed, 2011). Essa abordagem não depende da suposição de HWE, os dados são transformados utilizando uma análise de componentes principais (PCA) e em seguida é realizada uma análise discriminante (DA) para os componentes principais retidos. A identificação do número de grupos (K) foi realizada empregando o algoritmo *K-means* e o número estimado usando o Critério de Informação Bayesiano (*Bayesian Information Criterion*, BIC), que atinge um valor mínimo quando ocorre a designação melhor suportada de indivíduos para o número apropriado de grupos (Jombart e Ahmed, 2011; Jombart et al., 2010).

6.2.6 Detecção e caracterização de SNPs outliers

Os SNPs supostamente sob seleção (*locos outliers*) identificados com o *fsthet* (Flanagan e Jones, 2017) foram submetidos a pesquisas BLAST (*Basic Local Alignment Search Tool*) usando Blast2GO com configurações padrão (Götz et al., 2008) para encontrar sequências semelhantes depositadas no banco de dados não redundante do *National Center for Biotechnology Information* (NCBI). O algoritmo de escolha foi o *blastx* que é usado para encontrar produtos de tradução potenciais de uma sequência de nucleotídeos desconhecida.

Sequências dos SNPs supostamente sob seleção com ocorrências BLAST significativas foram utilizados para recuperar as anotações dos termos *Gene Ontology* (GO) que resumem as informações sobre funções moleculares, processos biológicos e componentes celulares (Blake, 2013), que posteriormente foram atribuídas às sequências de consulta. Estas etapas foram realizadas adotando a configuração padrão do Blast2GO.

6.3 Resultados

6.3.1 Modelos de distribuição de espécies

O mapa de distribuição potencial para identificar áreas com adequabilidade ambiental para *H. speciosa* (Figura 4), predito a partir de registros de ocorrência e variáveis bioclimáticas, foi confiável considerando o valor de TSS igual a 0,72.

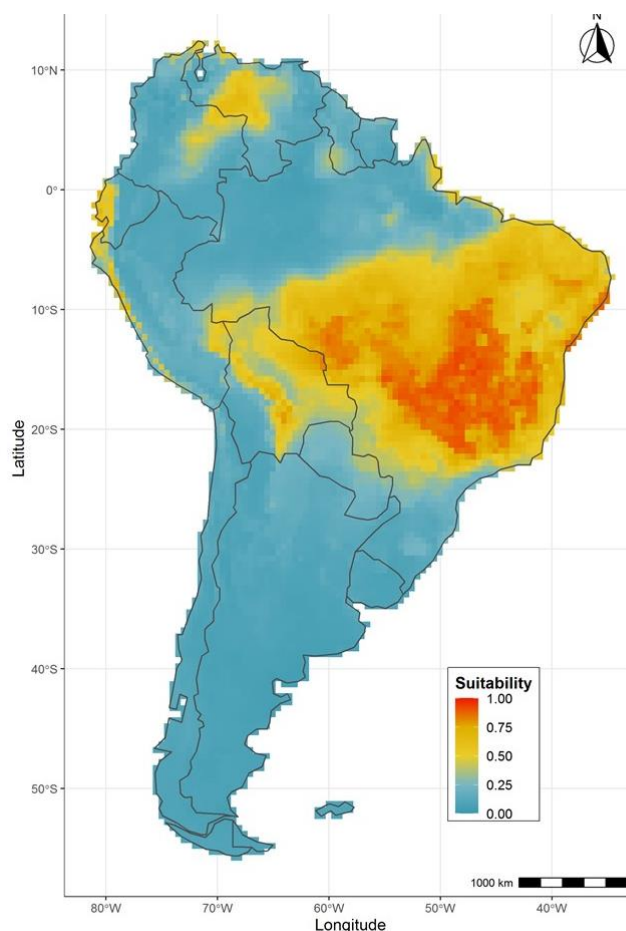


Figura 4. Distribuição geográfica potencial de *Hancornia speciosa*, considerando o cenário climático atual. A distribuição potencial é indicada por um gradiente de cor variando de vermelho (alta adequabilidade) a azul (baixa adequabilidade).

A área total modelada correspondeu a 17,8 milhões de km² e destes apenas 29,78% possuem adequabilidade ambiental para a ocorrência de *H. speciosa*. As áreas com maior adequabilidade no Brasil estão distribuídas nos biomas Mata Atlântica (restinga) e Cerrado.

A aptidão ambiental para a espécie foi reduzida nas áreas mais extremas do Norte e Sul do Brasil. Ainda de acordo com o modelo predito, para os demais países da América do Sul nos quais há registros de ocorrência da espécie, foram observadas apenas áreas de média adequabilidade (0,50): na Bolívia - biomas Floresta Amazônica, Cerrado/Savana, Floresta Tropical Sazonalmente Seca; e em parte da região Madre de Dios no Peru onde o bioma predominante é a Floresta Amazônica (Miranda et al., 2018).

A partir do modelo de distribuição nota-se também áreas com adequabilidade de 0,25 a 0,50 na Venezuela, Colômbia e Equador, porém nestes países não há registros de ocorrência da espécie.

6.3.2 Diversidade genômica e estrutura genética

O sequenciamento resultou em 156.209.732 reads e, após a limpeza, a demultiplexagem e filtragem dos dados faltantes foram identificadas 5.046 SNPs de alta qualidade, a profundidade média de sequenciamento por indivíduo variou de 2,08 a 84,62;

enquanto a profundidade média por loco variou de 8,46 a 123,5. E foram retidos 74 indivíduos para as análises subsequentes.

A análise com o *fsthet* detectou 355 locos *outlier*, ou seja, SNPs supostamente sob seleção, e 4.691 SNPs neutros que foram utilizados para inferir a diversidade genômica populacional e estrutura genética. A população de Barra do Garças-MT apresentou os maiores valores para todas as variáveis e juntamente com as populações de Coxim-MS e Conde-BA apresentou alelos privados (Tabela 1).

Tabela 1. Parâmetros de diversidade genômica usando 4.691 SNPs neutros de seis populações naturais de *Hancornia speciosa* Gomes. Número de indivíduos (*N*), número médio de alelos (*A*), número de alelos efetivos (*Ne*) e alelos privados (*Ap*).

População	N	A ± SE	Ne ± SE	Ap
Estância-SE	10	9,22 (±0,010)	1,281 (±0,005)	0
Pirambu-SE	15	14,64 (±0,010)	1,268 (±0,005)	0
Conde-BA	20	19,64 (±0,015)	1,305 (±0,005)	9
Marechal Deodoro-AL	9	8,865 (±0,006)	1,251 (±0,005)	0
Coxim-MS	8	7,123 (±0,010)	1,305 (±0,005)	85
Barra do Garças-MT	12	11,479 (±0,012)	1,405 (±0,005)	221

A população de Barra do Garças-MT apresentou maior diversidade genética ($H_E = 0,259$) e para todas as populações a heterozigosidade observada foi maior que a esperada e o coeficiente de endogamia (*f*) variou de -0,123 (Estância-SE) a -0,066 (Barra do Garça-MT); portanto, as populações apresentam excesso de heterozigotos (Figura 5).

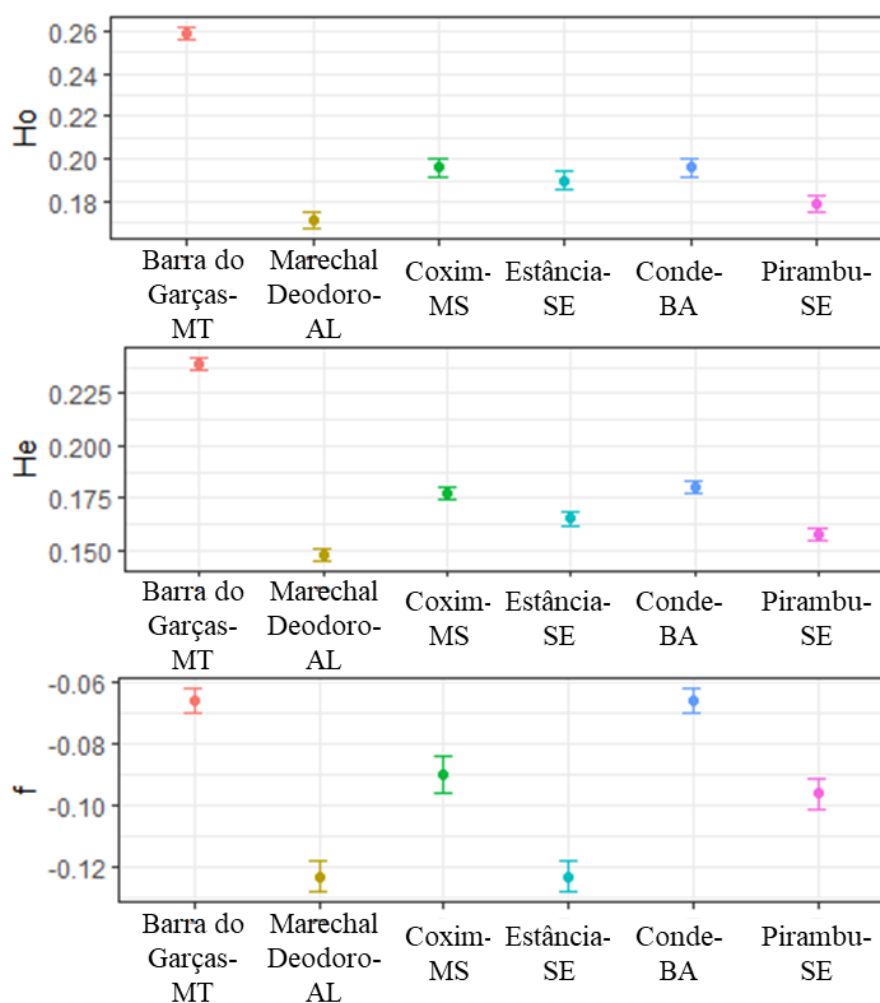


Figura 5. Estimativas de diversidade genômica e endogamia baseada em 4.691 SNPs neutros de seis populações naturais de *Hancornia speciosa* Gomes.

Nas variações da distribuição genética dentro e entre as populações, a análise de variância molecular (AMOVA), considerando todas as populações, mostrou que a maior variação estava presente nos indivíduos da mesma população (65%), o mesmo foi observado quando foram avaliadas somente as populações da região Nordeste (93%). Além disso, ao observar as duas AMOVA percebe-se que as populações no Centro-Oeste são responsáveis por aumentar a variação entre populações (Tabela 2).

Tabela 2. Análise de variância molecular (AMOVA) baseada em 4.691 marcadores SNP neutros para seis populações naturais de *Hancornia speciosa* e apenas para as populações da região Nordeste.

Populações das regiões Nordeste e Centro-Oeste					
Fonte de Variação	Graus de liberdade	Soma dos Quadrados	Quadrado médio	Variância	Porcentagem
Entre populações	5	41.410,176	8.282,035	319,279	35%
Dentro de populações	142	82.555,986	581,380	581,380	65%
Total	147	123.966,162		900,659	100%
Populações da região Nordeste					

Entre populações	3	4.745,773	1.581,924	40,723	7%
Dentro de populações	104	54.195,986	1.581,924	521,115	93%
Total	107	58.941,759		561,838	100%

Considerando a amplitude do grau de diferenciação genética analisada pelas estimativas do F_{ST} par-a-par, verificou-se pequena de diferenciação genética de ($F_{ST} = 0,046$ a $0,099$) entre as populações localizadas na região Nordeste (bioma Mata Atlântica) e diferenciação genética muito grande ($F_{ST} = 0,442$ a $0,538$) ao comparar essas populações e as duas populações localizadas no Centro-Oeste (Cerrado e transição Cerrado/Pantanal) (Figura 6).

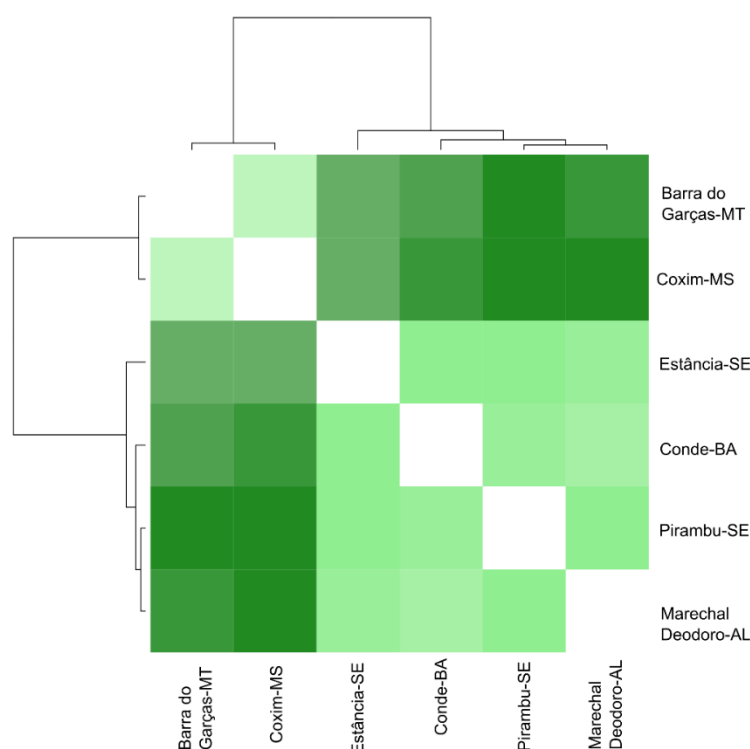


Figura 6. Heatmap para as estimativas de F_{ST} par-a-par baseadas em 4.691 marcadores SNP neutros para seis populações naturais de *Hancornia speciosa* Gomes. Regiões de coloração mais intensa indicam maior F_{ST} .

As análises bayesiana para acessar a estrutura genômica dos 74 indivíduos de *H. speciosa* com base em 4.691 SNPs neutros (7A) e com base em 355 SNPs *outlier* (Figura 7B) sugerem a existência de dois grupos genéticos (Figura 6) com base no método ΔK (Evanno et al., 2005). Os grupos formados são bem definidos e fortemente associados à origem geográfica dos indivíduos. O grupo 1 (em vermelho, na Figura 7A e B) foi composto por 54 indivíduos de *H. speciosa* com origem em áreas de Mata Atlântica (Estância-SE, Pirambu-SE, Conte-BA e Marechal Deodoro-AL) no Nordeste do Brasil. O grupo 2 (em verde, na Figura 7A e B) foi composto por 20 indivíduos com origem geográfica no Centro-Oeste (Coxim-MS e Barra do Garças-MT) no bioma Cerrado e em área de transição Cerrado/Pantanal.

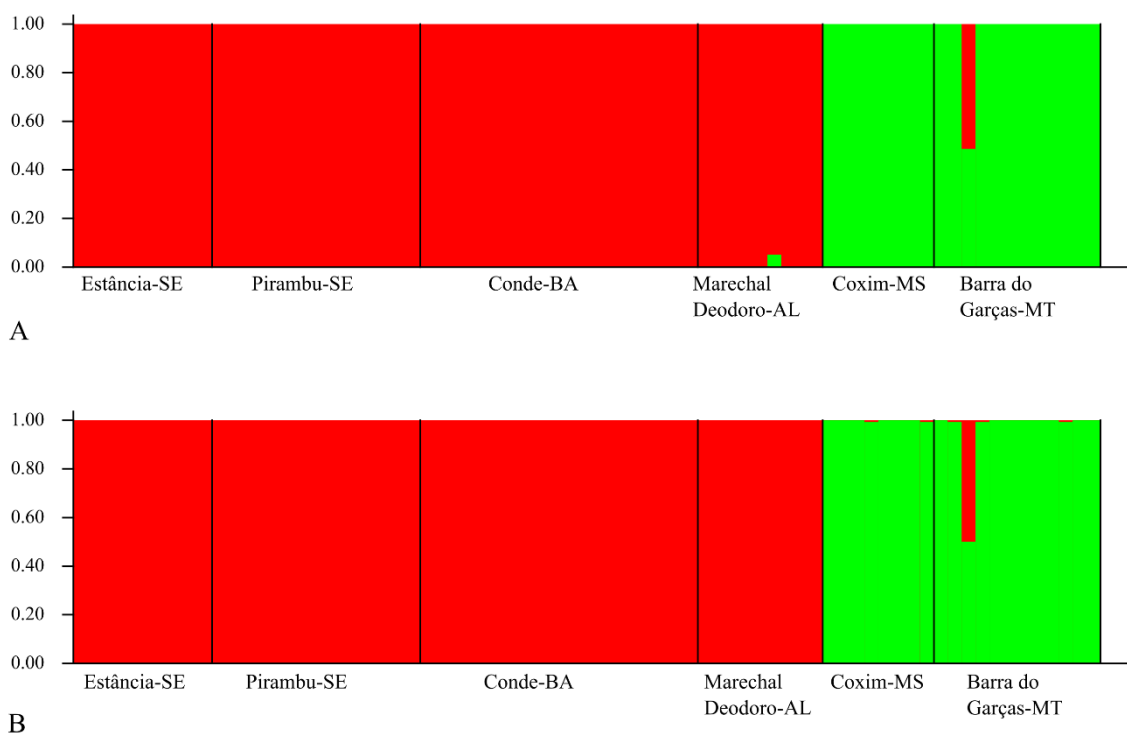


Figura 6. Análise Bayesiana com marcadores SNPs para 74 indivíduos de *Hancornia speciosa*. Probabilidades de atribuição dos indivíduos aos agrupamentos com base em 4.691 marcadores SNP neutros (A) e com base em 355 marcadores SNP outliers (B).

Complementarmente, as análises realizadas por meio de DAPC revelaram padrões um pouco diferentes, na análise inicial sem informações prévias sobre os agrupamentos, por meio do método de *K-means*, o número de agrupamentos que melhor explicou os dados de variação genômica entre os indivíduos foi $K = 3$ (considerando os valores de BIC), para as análises realizadas com marcadores SNPs neutros e também para *outliers*. Quando os grupos de populações são representados nas DAPC pode-se observar uma marcante separação entre a população de Coxim-MS e demais populações para os SNPs neutros (Figura 8A e C) e SNPs *outliers* (Figura 8B e D).

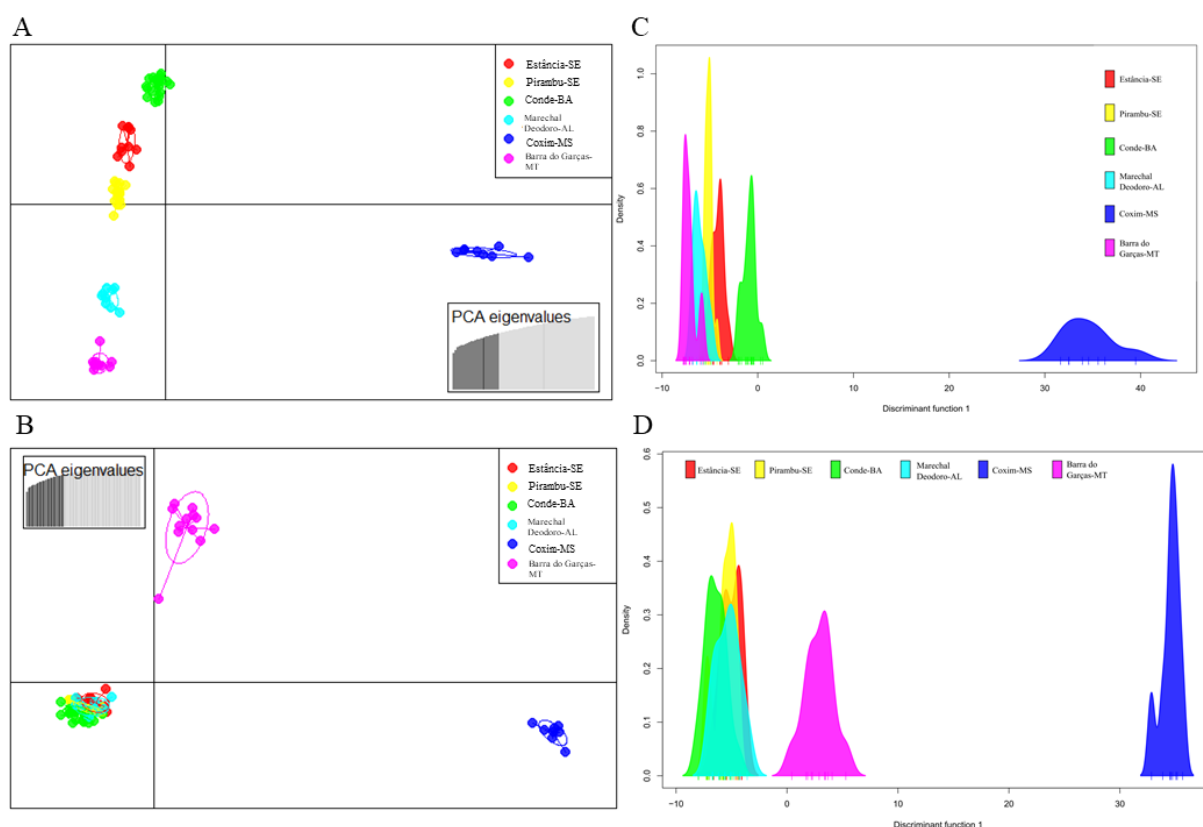


Figura 8. Análises Discriminantes de Componentes Principais (DAPC) com marcadores SNPs para 74 indivíduos de *Hancornia speciosa*. Gráfico de dispersão (A) e gráfico de densidade (C) com base em 4.691 marcadores SNPs com comportamento neutro; Gráfico de dispersão (B) e gráfico de densidade (D) com base em 355 marcadores SNPs com comportamento *outlier*.

A DAPC baseada nos marcadores SNPs neutros mostrou a formação de três grupos principais distintos, sendo as populações de Estância-SE, Pirambu-SE e Conde-BA formando um deles; as populações de Marechal Deodoro-AL e Barra do Garças-MT mais proximamente relacionados e Coxim-MS com maior diferenciação genética (Figura 8A e C). Na DAPC para os SNPs *outliers* as populações da região Nordeste (Estância-SE, Pirambu-SE, Conde-BA e de Marechal Deodoro -AL) apresentaram baixa diferenciação genética e as populações da região Centro-Oeste apresentaram maior diferenciação genética (Figura 8B e D).

6.3.3 Caracterização de SNPs supostamente sob seleção

Para as 355 sequências de SNPs com comportamento *outlier* que foram analisadas com o Blast2GO, apenas para 27 delas ocorreu BLAST com êxito, destas apenas 22 foram mapeadas e para duas foi possível chegar na anotação funcional baseado nos termos GO (Figura 9).

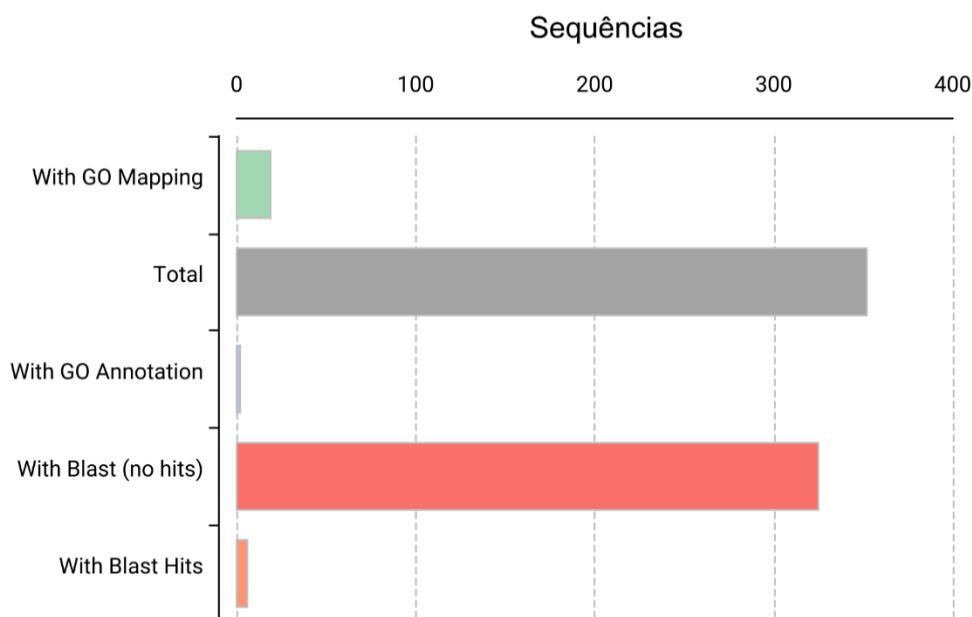


Figura 9. Resumo da análise de 355 SNPs supostamente sob seleção no Blast2GO para populações naturais de *Hancornia speciosa*.

Para a anotação funcional da ontologia genética classificada de acordo com o termo GO, as duas sequências para as quais foi possível chegar a anotação funcional foram atribuídas a diferentes grupos funcionais: componente celular - as partes de uma célula ou seu ambiente extracelular; função molecular - atividades elementares de um produto gênico no nível molecular, como ligação ou catálise; processo biológico - operações ou conjuntos de eventos moleculares com início e fim definidos, pertinentes ao funcionamento de unidades vivas integradas (células, tecidos, órgãos e organismos) (Ding et al., 2018) (Tabela 3).

Tabela 3. Identificação e anotação funcional dos marcadores SNPs supostamente sob seleção para populações naturais de *Hancornia speciosa*.

ID	SNP	Termos GO		
		Componente celular	Função molecular	Processo biológico
SNP_20625	G/A		GO:0003676 ligação de ácido nucleico GO:0008270 ligação de íon zinco	GO:0015074 Integração de DNA
SNP_14078	A/G	GO:0000139 Membrana de Golgi GO:0005789 membrana de retículo endoplasmático GO:0016021 componente integral da membrana	GO:0004190 atividade de endopeptidase de tipo aspártico	GO:0007219; Via de sinalização Notch GO:0016485 processamento de proteína

6.4 Discussão

Estudos genômicos populacionais em espécies florestais podem fornecer informações relevantes para diferentes áreas, como a ecologia, conservação e melhoramento. Nesse

sentido, este é o primeiro estudo utilizando GBS para a identificação de SNPs em *H. speciosa* e sua aplicação para inferir a diversidade genômica e estrutura populacional.

As populações estudadas estão inseridas em áreas com alta adequabilidade ambiental de acordo com o mapa predito pelo SDM (TSS= 0,72), utilizando variáveis bioclimáticas relacionadas à temperatura e precipitação. A maior parte das áreas com alta adequabilidade estão situadas na região central do Brasil (Figura 4) (Centro-Oeste; parte do Sudeste - Minas Gerais; Nordeste - Extremo Oeste Baiano), sendo que nessas regiões a espécie está inserida no bioma Cerrado. A outra parte está concentrada no litoral nordestino no bioma Mata Atlântica.

Com exceção da população Conde, as populações do Nordeste apresentaram menor diversidade genética (Figura 5), o que pode ser resultado de interferências antrópicas nessas áreas do litoral nordestino em que o bioma Mata Atlântica foi praticamente substituído por plantações de cana-de-açúcar, cacau e por cidades (a maioria das capitais nordestinas estão no domínio da Mata Atlântica) (MMA, 2015).

Nas populações de Estância e Pirambu ocorre a exploração da espécie para a coleta de frutos com maior intensidade e sem plano de manejo, o que afeta a capacidade reprodutiva da espécie, sua regeneração natural, podendo contribuir para redução da diversidade genética. Considerando que a espécie têm importância socioeconômica para as comunidades extrativistas da região, que obtêm parte de sua renda pela comercialização dos frutos (Lima et al., 2013), uma estratégia propositiva visando a conservação e uso sustentável seria o desenvolvimento de um Plano de Manejo Florestal Comunitário e Familiar (MFCF) (Brasil, 2006; Brasil, 2009), que consiste na união do planejamento de ações e de técnicas adequadas para o uso econômico da floresta (Reis et al., 2013).

A população com maior diversidade genética foi Barra do Garças ($H_e = 0,239$) e está localizada no Parque Estadual Serra Azul-MT, aspecto que pode ter contribuído para maior diversidade. Este fato reforça a importância das Unidades de Conservação como uma estratégia de conservação *in situ*, sobretudo para espécies com sementes recalcitrantes que não toleram a perda de água e, portanto, não podem ser conservadas *ex situ* em condições convencionais de banco de sementes (Pammenter e Berjak, 2014), como é o caso de *H. speciosa*.

Nos dois biomas (Cerrado e Mata Atlântica) em que estão as áreas de maior adequabilidade preditas para *H. speciosa*, existem ameaças que colocam em risco a manutenção das populações naturais. No Cerrado, as ameaças consistem no corte para a produção de carvão, que prejudica a persistência de longo prazo da população e o fogo que é um distúrbio comum nesse bioma (Miranda et al., 2002) e compromete a sobrevivência de adultos e mudas da espécie (Lima et al., 2013).

Mesmo suscetível a diferentes ameaças é possível considerar que as populações avaliadas detêm alta diversidade genética, as estimativas de N_e e H_e foram semelhantes às obtidas em populações naturais no Rio Grande do Norte em domínio da Mata Atlântica empregando marcadores ISSR (Costa et al., 2015) que apresentaram $N_e = 1,8 - 1,49$ e $H_e = 0,16 - 0,29$; e A semelhante ao obtido para 28 populações naturais localizadas na região central do Brasil (Cerrado) e avaliadas por microsatélites, $A = 4,6 - 13,1$ (Collevatti et al., 2018). Também foram identificados alelos privados em três das populações estudadas, Barra do Garças ($A_p = 221$), Coxim ($A_p = 85$) e Conde ($A_p = 9$) que possuem diversidade não encontrada nas demais. Essa informação pode ser utilizada para ampliar a representatividade genética em bancos ativos de germoplasma e também para montar uma coleção de trabalho visando o melhoramento genético da espécie (Santos et al., 2015).

Para as populações estudadas, além da alta diversidade genética estimada, com a análise hierárquica de variância, maior variação foi observada dentro das populações do que entre populações (Tabela 2), esses resultados podem ser decorrentes do sistema reprodutivo da espécie que consiste na fecundação cruzada (Darrault e Schlindwein, 2005) para o qual é esperado altos níveis de diversidade genética dentro das populações (Cruz Neto et al., 2014). Experimentos de polinização realizados com *H. speciosa* demonstram que a espécie é

autoincompatível, apresenta altas taxas de cruzamento (próximo de 1,0) (Collevatti et al., 2016) e que seu mecanismo de polinização é altamente eficiente com características que dificultam a autopolinização e favorecem a polinização cruzada (Darrault e Schlindwein, 2005).

As estimativas do F_{ST} par-a-par demonstram que entre as populações de mesma região, como esperado, existe menor diferenciação genética (Figura 6). F_{ST} fornece informações importantes sobre os processos evolutivos que influenciam a estrutura da variação genética dentro e entre as populações (Holsinger e Weir, 2009). Por ser uma árvore tropical com ampla distribuição geográfica, *H. speciosa* pode experimentar alta heterogeneidade ambiental e condições ecológicas variadas, fatores que afetam a distribuição da diversidade e diferenciação genética (Le Corre e Kremer, 2012). A diferenciação entre populações correlacionada com a distribuição geográfica pode ser explicada considerando a limitação do fluxo gênico por meio da dispersão de pólen ou sementes pela distância geográfica entre as regiões.

As populações foram separadas em grupos distintos que correspondem às regiões geográficas de origem quando avaliada a estrutura genética pelo método bayseano considerando marcadores neutros (Figura 7A) e *outliers* (Figura 7B). Baseado em características fenotípicas, populações de *H. speciosa* que ocorrem na região Nordeste em direção ao Centro-Oeste e Norte do Brasil são consideradas variedade botânica *H. speciosa* var. *speciosa*, enquanto que na região central do Brasil são considerados como *H. speciosa* var. *pubescens* (Flora do Brasil, 2021). Entretanto, nas análises de DAPC (Figuras 8A-D), as populações do Centro-Oeste são agrupadas separadamente e não suportam a classificação em apenas duas variedades botânicas. De acordo Chaves et al. (2020), as variedades botânicas *H. speciosa* var. *pubescens*, *H. speciosa* var. *gardneri* e *H. speciosa* var. *speciosa* tendem a ser mais semelhantes, enquanto a variedade *H. speciosa* var. *cuyabensis* mostrou-se mais diferenciada das demais.

Sobre as diferentes variedades botânicas, um consenso é que no litoral nordestino ocorre apenas *H. speciosa* var. *speciosa*, que parece ser a variedade típica, e apresenta a maior distribuição geográfica no país, diferente das demais variedades que estão atribuídas a regiões específicas do país. Não existem barreiras reprodutivas entre variedades botânicas (Collevatti et al., 2016), o que pode favorecer o cruzamento entre as variedades em zonas de transição entre uma região e outra.

Outro fator que pode favorecer o cruzamento entre variedades e populações é que *H. speciosa* conta com visitantes florais de várias espécies, como *Eulaema nigrata* (Euglossini), *Heliconius nanna* (Nymphalidae), *Perichares filetes adela* (Hesperiidae) e *Erinnys ello* (Sphingidae) e embora não tenha sido detectado relação espécie-específica entre a planta e seus visitantes florais, espécies da família Lepidoptera são mais frequentes, sobretudo mariposas que possuem aparelho bucal longo o suficiente para chegar ao fundo do tubo de flores (Darrault e Schlindwein, 2005). Esses polinizadores podem percorrer grandes áreas durante suas rotas de forrageamento e contribuem para a dispersão de pólen a longa distância (Skogen et al., 2016; Aguilar et al., 2019).

Para as populações estudadas, ao avaliar a estruturação genética, verificou-se que a população de Coxim apresentou diferenciação acentuada em relação às demais populações, sobretudo considerando a DAPC em que independente do conjunto de marcadores SNPs (neutros ou supostamente sob seleção) analisados, os indivíduos foram agrupados isolados dos demais. Durante a coleta, verificou-se que diferente das outras populações, os seus indivíduos estavam em área com vegetação fechada e sombreados por outras espécies arbóreas. Essa população está localizada em uma área de transição entre os biomas Cerrado e Pantanal e aparentemente é mais isolada em relação às demais populações estudadas, motivo que talvez explique seu maior grau de diferenciação quando comparada às demais populações.

Os resultados obtidos para o conjunto de SNPs supostamente sob seleção, revelam uma estruturação genética formada por três grupos que estão correlacionados com os

diferentes biomas e fitofisionomias em que cada população está inserida, Mata Atlântica (Estância, Pirambu Conde e de Marechal Deodoro), Cerrado (Barra do Garças) e área de transição entre Cerrado e Pantanal (Coxim), sendo esses resultados evidências de adaptações locais. Ainda que o mapa de distribuição predito indique alta adequabilidade ambiental em todas as populações, a mudança de biomas exigiu a adaptação da espécie, indicando a necessidade de incluir camadas com informações sobre a vegetação para gerar mapas preditivos mais resolutivos, especialmente quando a espécie estudada apresentar ampla distribuição geográfica, como é o caso de *H. speciosa*.

A adaptação local ocorre quando a seleção natural faz com que as populações locais evoluam e se adaptem melhor à heterogeneidade das condições bióticas e abióticas, podendo resultar em características fenotípicas divergentes e adaptações genéticas que alteram a composição genética das populações ao longo das gerações (Le Corre e Kremer, 2012). A capacidade de adaptação pode, no entanto, ser comprometida em pequenas populações devido à diversidade genética reduzida e deriva genética (Leimu et al., 2006; Stockwell et al., 2003). O fluxo gênico também exerce efeitos sobre a adaptação local, o nível de adaptação aumenta com o aumento da distância entre as populações para as quais o fluxo gênico é reduzido e ocorre o aumento da diferenciação genética (Becker et al., 2006).

O baixo número de SNPs supostamente sob seleção para os quais existe anotação funcional é esperado em espécies não-modelo, como é o caso de *H. speciosa*. Além disso, até o momento não existem informações sobre o genoma nuclear de *H. speciosa* em bancos públicos de dados moleculares (Nunes et al., 2022), sendo este trabalho o primeiro a investigar sequências genômicas e possíveis anotações funcionais. As funções da maioria dos genes de espécies vegetais permanecem desconhecidas, sobretudo ao considerar espécies não-modelo (Zwaenepoel et al., 2018), nesse contexto, o uso de ferramentas baseadas em similaridade de sequência é importante e frequentemente aplicado para fornecer uma noção inicial sobre a função de um gene em um genoma recém-sequenciado e anotado (Bains et al., 2019; Mehravi et al., 2021).

Em *H. speciosa* foi identificada anotação funcional com função molecular envolvida na ligação ao ácido nucleico (GO:0003676), essa anotação também foi observada para *Triticum aestivum* L. (trigo) em análise como MicroRNAs (miRNAs) (Achakzai et al., 2018) e em fatores de transcrição dedo-homeodomínio de zinco (ZF-HDs) (Liu et al., 2021). Também foi identificada em fatores de transcrição de folhas de videira (*Vitis vinifera* L.) sob estresse salino (Upadhyay et al., 2018) e em miRNAs em amostras de folhas de *Solanum lycopersicum* L. (tomate) tratadas com as diferentes qualidades de luz (Dong et al., 2020).

Para espécies florestais a anotação funcional GO:0003676 foi relacionada a ESTs (*Expressed Sequence Tags*) de *Quercus mongólica* (Ueno et al., 2008) e de *Pinus sylvestris* var. *mongólica* (Fang et al., 2014), além de atribuída a genes de *Picea glauca* (Moench) Voss) (abeto branco) e *Picea mariana* (Mill.) (abeto preto) (Pavy et al., 2012) e a miRNAs *Malus domestica* Borkh. (macieira) potencialmente envolvidos na regulação do crescimento de gemas e na indução e desenvolvimento de flores (Xing et al., 2016).

Outra função molecular identificada para os SNPs supostamente sob seleção de *H. speciosa* foi GO: 0008270 (ligação de íons de zinco) também identificada para *Pinus sylvestris* var. *mongólica* (Fang et al., 2014) e em genes plantas de *Citrus reticulata* Blanco (tangerina) afetadas por Huanglongbing (HLB) e tratadas por termoterapia (Pan et al., 2021).

Ambas anotações são classificadas como atividade de um produto gênico em nível molecular, nas espécies *S. lycopersicum*, *T. aestivum*, *Vitis vinifera* e *C. reticulata* foram anotadas para sequências de genes ou produtos gênicos oriundos de material vegetal submetido a tratamentos como luz, termoterapia e estresses abióticos como alta salinidade, altas temperaturas e frio (Upadhyay et al., 2018; Dong et al., 2020; Liu et al., 2021; Pan et al., 2021). Para *H. speciosa* não foi possível associar essa informação a nenhuma característica adaptativa específica. Entretanto, as sequências dos SNPs identificados para populações

amplamente distribuídas, nesse estudo, podem servir de referência para novos trabalhos com a espécie.

6.5. Conclusão

Foi possível identificar um grande número de SNPs neutros e supostamente sob seleção.

O modelo da distribuição potencial de *H. speciosa* foi predito com sucesso e revela a ocorrência da espécie em regiões que possuem clima mais quente. As áreas de maior adequabilidade estão situadas na região central do Brasil (parte do Centro-Oeste, Sudeste e Nordeste) e no litoral nordestino nos biomas Cerrado e Mata Atlântica.

A estrutura genômica revela a formação de grupos genéticos bem definidos e confirmou a distinção entre populações da região Nordeste e Centro-Oeste. SNPs supostamente sob seleção apresentam uma estruturação genética formada por três grupos que estão correlacionados com os diferentes biomas que as populações estão inseridas, indicando adaptação local. Apenas para dois SNPs, supostamente sob seleção, foi possível identificar anotação funcional.

Esses resultados contribuem para avançar no conhecimento sobre a espécie e para o desenvolvimento de estratégias visando o manejo, conservação e melhoramento genético.

6.6 Referências Bibliográficas

- Achakzai, H.K., Barozai, M.Y.K., Din, M., Baloch, I.A., Achakzai, A.K.K., 2018. Identification and annotation of newly conserved micrnas and their targets in wheat (*Triticum aestivum* L.). PLoS ONE 13, e0200033. <https://doi.org/10.1371/journal.pone.0200033>
- Aguilar, A., Twardowski, T., Wohlgemuth, R., 2019. Bioeconomy for Sustainable Development. Biotechnology Journal 14, 1800638.
- Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology 43, 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. Trends in Ecology & Evolution 22, 42–47. <https://doi.org/10.1016/J.TREE.2006.09.010>
- Bains, S., Thakur, V., Kaur, J., Singh, K., Kaur, R., 2019. Elucidating genes involved in sesquiterpenoid and flavonoid biosynthetic pathways in *Saussurea lappa* by de novo leaf transcriptome analysis. Genomics 111, 1474–1482. <https://doi.org/10.1016/j.ygeno.2018.09.022>
- Becker, U., Colling, G., Dostal, P., Jakobsson, A., Matthies, D., 2006. Local adaptation in the monocarpic perennial *Carlina vulgaris* at different spatial scales across Europe. Oecologia 150, 506–518.
- Black, 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, 441–469. <https://doi.org/10.1146/ANNUREV.ENTO.46.1.441>
- Blake, J.A., 2013. Ten Quick Tips for Using the Gene Ontology. PLOS Computational Biology 9, e1003343. <https://doi.org/10.1371/JOURNAL.PCBI.1003343>
- Brasil, 2006. Lei nº 11.284 de 2 de março de 2006. Presidência da República, Brasília.
- Brasil, 2009. Decreto nº 6.874, de 5 de junho de 2009. Presidência da República, Brasília.
- Brasil, 2015. Lei nº 13.123 de 20 de maio de 2015. Presidência da República, Brasília.
- Brockhoff, E.G., Barbaro, L., Castagnyrol, B., Forrester, D.I., Gardiner, B., González-Olabarria, J.R., Lyver, P.O., Meurisse, N., Oxbrugh, A., Taki, H., Thompson, I.D., van der Plas, F., Jactel, H., 2017. Forest biodiversity, ecosystem functioning and the provision of ecosystem services. Biodiversity and Conservation 26, 3005–3035. <https://doi.org/10.1007/S10531-017-1453-2>

- Carpenter, G., Gillison, A.N., Winter, J., 1993. DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation* 2, 667–680. <https://doi.org/10.1007/BF00051966>
- 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, 171–182. <https://doi.org/10.1534/g3.111.000240>
- Chaves, L.J., Ganga, R.M.D., Guimarães, R.A., Caldeira, A.J.R., 2020. Quantitative and molecular genetic variation among botanical varieties and subpopulations of *Hancornia speciosa* Gomes (Apocynaceae). *Tree Genetics and Genomes* 16. <https://doi.org/10.1007/s11295-020-01444-0>
- Collevatti, R.G., Olivatti, A.M., Telles, M.P.C., Chaves, L.J., 2016. Gene flow among *Hancornia speciosa* (Apocynaceae) varieties and hybrid fitness. *Tree Genetics and Genomes* 12. <https://doi.org/10.1007/s11295-016-1031-x>
- Collevatti, R.G., Rodrigues, E.E., Vitorino, L.C., Lima-Ribeiro, M.S., Chaves, L.J., Telles, M.P.C.C., 2018. Unravelling the genetic differentiation among varieties of the Neotropical savanna tree *Hancornia speciosa* Gomes. *Annals of botany* 122, 973–984. <https://doi.org/10.1093/aob/mcy060>
- Coradin, L., Camilo, J., Pareyn, F.G.C., 2018. Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial. Ministério do Meio Ambiente, Secretaria de Biodiversidade, Brasília.
- Costa, D.F. da, Vieira, F. de A., Fajardo, C.G., Chagas, K.P.T. das, 2015. Diversidade genética e seleção de iniciadores issr em uma população natural de mangaba (*Hancornia speciosa* Gomes) (Apocynaceae). *Revista Brasileira de Fruticultura* 37, 970–976. <https://doi.org/10.1590/0100-2945-246/14>
- Cruz Neto, O., Aguiar, A. v., Twyford, A.D., Neaves, L.E., Toby Pennington, R., Lopes, A. v., 2014. Genetic and Ecological Outcomes of *Inga vera* Subsp. *affinis* (Leguminosae) Tree Plantations in a Fragmented Tropical Landscape. *PLOS ONE* 9, e99903. <https://doi.org/10.1371/JOURNAL.PONE.0099903>
- Darrault, R.O., Schlindwein, C., 2005. Limited fruit production in *Hancornia speciosa* (Apocynaceae) and pollination by nocturnal and diurnal insects. *Biotropica* 37, 381–388. <https://doi.org/10.1111/j.1744-7429.2005.00050.x>
- Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F.T., D'Amen, M., Randin, C., Engler, R., Pottier, J., Pio, D., Dubuis, A., Pellissier, L., Mateo, R.G., Hordijk, W., Salamin, N., Guisan, A., 2017. ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography* 40, 774–787. <https://doi.org/10.1111/ecog.02671>
- Díaz, B.G., Zucchi, M.I., Alves-Pereira, A., Almeida, C.P. de, Moraes, A.C.L., Vianna, S.A., Azevedo-Filho, J., Colombo, C.A., 2021. Genome-wide SNP analysis to assess the genetic population structure and diversity of *Acrocomia* species. *PLOS ONE* 16, e0241025. <https://doi.org/10.1371/JOURNAL.PONE.0241025>
- Ding, R., Qu, Y., Wu, C.H., Vijay-Shanker, K., 2018. Automatic gene annotation using GO terms from cellular component domain. *BMC Medical Informatics and Decision Making* 18, 97–106. <https://doi.org/10.1186/S12911-018-0694-7/TABLES/2>
- Diniz-Filho, J.A.F., Mauricio Bini, L., Fernando Rangel, T., Loyola, R.D., Hof, C., Nogués-Bravo, D., Araújo, M.B., 2009. Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography* 32, 897–906. <https://doi.org/10.1111/j.1600-0587.2009.06196.x>
- Dong, F., Wang, C., Dong, Y., Hao, S., Wang, L., Sun, X., Liu, S., 2020. Differential expression of microRNAs in tomato leaves treated with different light qualities. *BMC Genomics* 21. <https://doi.org/10.1186/s12864-019-6440-4>
- Doyle, JJ, Doyle, JL, 1990. Isolamento de DNA de planta de tecido fresco. *Focus* 12, 13–15.
- Earl, D.A.; Vonholdt, B.M., 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4, 359–361.

- 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, e19379. <https://doi.org/10.1371/JOURNAL.PONE.0019379>
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fang, P., Niu, S., Yuan, H., Li, Z., Zhang, Y., Yuan, L., Li, W., 2014. Development and Characterization of 25 EST-SSR markers in *Pinus sylvestris* var. *mongolica* (Pinaceae) . *Applications in Plant Sciences* 2, 1300057. <https://doi.org/10.3732/apps.1300057>
- Farber, O., Kadmon, R., 2003. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling* 160, 115–130. [https://doi.org/10.1016/S0304-3800\(02\)00327-7](https://doi.org/10.1016/S0304-3800(02)00327-7)
- Feng, X., Park, D.S., Walker, C., Peterson, A.T., Merow, C., Papeş, M., 2019. A checklist for maximizing reproducibility of ecological niche models. *Nature Ecology and Evolution* 3, 1382–1395. <https://doi.org/10.1038/s41559-019-0972-5>
- Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37, 4302–4315. <https://doi.org/10.1002/joc.5086>
- Flanagan, S.P., Jones, A.G., 2017. Constraints on the F_{ST} -Heterozygosity Outlier Approach. *Journal of Heredity* 108, 561–573. <https://doi.org/10.1093/jhered/esx048>
- Flora do Brasil, 2021. *Hancornia speciosa* Gomes. <http://floradobrasil.jbrj.gov.br/reflora/listaBrasil/FichaPublicaTaxonUC/FichaPublicaTaxonUC.do?id=FB15558> (acesso 20.07.21).
- Fois, M., Cuena-Lombrana, A., Fenu, G., Bacchetta, G., 2018. Using species distribution models at local scale to guide the search of poorly known species: Review, methodological issues and future directions. *Ecological Modelling* 385, 124–132. <https://doi.org/10.1016/J.ECOLMODEL.2018.07.018>
- GBIF, 2021. Free and open access to biodiversity data. <https://www.gbif.org/> (acesso 04.08.21).
- Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J., Conesa, A., 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research* 36, 3420–3435. <https://doi.org/10.1093/NAR/GKN176>
- Holsinger, K.E., Weir, B.S., 2009. Genetics in geographically structured populations: defining, estimating and interpreting F_{ST} . *Nature Reviews Genetics* 10, 639–650.
- Jombart, T., Ahmed, I., 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27, 3070–3071. <https://doi.org/10.1093/BIOINFORMATICS/BTR521>
- Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* 11, 1–15. <https://doi.org/10.1186/1471-2156-11-94>
- Kuhn, M., 2008. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software* 28, 1–26. <https://doi.org/10.18637/JSS.V028.I05>
- Le Corre, V., Kremer, A., 2012. The genetic differentiation at quantitative trait loci under local adaptation. *Molecular ecology* 21, 1548–1566. <https://doi.org/10.1111/J.1365-294X.2012.05479.X>
- Leimu, R., Mutikainen, P., Koricheva, J., Fischer, M., 2006. How general are positive relationships between plant population size, fitness and genetic variation? *Journal of Ecology* 94, 942–952. <https://doi.org/10.1111/J.1365-2745.2006.01150.X>
- Lewis, S.L., Edwards, D.P., Galbraith, D., 2015. Increasing human dominance of tropical forests. *Science* 349, 827–832. <https://doi.org/10.1126/SCIENCE.AAA9932>
- Lima, I.L.P., Scariot, A., Giroldo, A.B., 2013. Sustainable Harvest of Mangaba (*Hancornia speciosa*) Fruits in Northern Minas Gerais, Brazil. *Economic Botany* 67, 234–243. <https://doi.org/10.1007/s12231-013-9244-5>

- Liu, H., Yang, Y., Zhang, L., 2021. Zinc finger-homeodomain transcriptional factors (ZF-HDS) in wheat (*Triticum aestivum* L.): Identification, evolution, expression analysis and response to abiotic stresses. *Plants* 10, 1–15. <https://doi.org/10.3390/plants10030593>
- 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, 981–994. <https://doi.org/10.1038/nrg1226>
- Mehravi, S., Ranjbar, G.A., Mirzaghaderi, G., Severn-Ellis, A.A., Scheben, A., Edwards, D., Batley, J., 2021. De novo snp discovery and genotyping of iranian pimpinella species using ddRAD sequencing. *Agronomy* 11, 1342. <https://doi.org/10.3390/agronomy11071342>
- Miranda, H.S., Bustamante, M.M.C., Miranda, A.C., 2002. O fator de fogo, In: *Os Cerrados Do Brasil: Ecologia e História Natural de Uma Savana Neotropical*. Columbia University Press, Nova York, pp. 51–68.
- Miranda, P.L.S. de, Oliveira-Filho, A.T., Pennington, R.T., Neves, D.M., Baker, T.R., Dexter, K.G., 2018. Using tree species inventories to map biomes and assess their climatic overlaps in lowland tropical South America. *Global Ecology and Biogeography* 27, 899–912. <https://doi.org/10.1111/geb.12749>
- MMA, 2015. Mapa de Vegetação Nativa na Área de Aplicação da Lei nº 11.428/2006 – Lei da Mata Atlântica (ano base 2009). Ministério de Meio Ambiente.
- Monachino, J., 1945. A Revisión de *Hancornia* (Apocynaceae). *Lilloa*, San Miguel de Tucumán, v. 11, p. 19–48.
- Mora-Sala, B., Gramaje, D., Abad-Campos, P., Berbegal, M., 2019. Diversity of *Phytophthora* Species Associated with *Quercus ilex* L. in Three Spanish Regions Evaluated by NGS. *Forests* 10, 979. <https://doi.org/10.3390/F10110979>
- Nabout, J.C., Magalhães, M.R., de Amorim Gomes, M.A., da Cunha, H.F., 2016. The Impact of Global Climate Change on the Geographic Distribution and Sustainable Harvest of *Hancornia speciosa* Gomes (Apocynaceae) in Brazil. *Environmental Management* 57, 814–821. <https://doi.org/10.1007/s00267-016-0659-5>
- Nix, H.A., 1986. A biogeographic analysis of Australian elapid snakes. In: *Atlas of elapid snakes of Australia*, in: *Atlas of Elapid Snakes of Australia*. Australian Government Publishing Service, Canberra, pp. 4–15.
- Novaes, E., Drost, D.R., Farmerie, W.G., Pappas, G.J., Grattapaglia, D., Sederoff, R.R., Kirst, M., 2008. High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics* 9, 1–14. <https://doi.org/10.1186/1471-2164-9-312>
- Nunes, V.V., Silva-Mann, R., Souza, J.L., Calazans, C.C., 2022. Pharmaceutical, food potential, and molecular data of *Hancornia speciosa* Gomes: a systematic review. *Genetic Resources and Crop Evolution* 69, 525–543. <https://doi.org/10.1007/s10722-021-01319-w>
- Pammenter, N.W., Berjak, P., 2014. Physiology of Desiccation-Sensitive (Recalcitrant) Seeds and the Implications for Cryopreservation. *International Journal of Plant Sciences* 175, 21–28. <https://doi.org/10.1086/673302>
- Pan, R., Chen, Y., Deng, X., Xu, M., 2021. Digital Gene Expression Analysis of Huanglongbing Affected Mandarins (*Citrus reticulata* Blanco) In Response to Thermotherapy. *Horticultural Plant Journal* 7, 1–12. <https://doi.org/10.1016/j.hpj.2020.12.004>
- Pavy, N., Pelgas, B., Laroche, J., Rigault, P., Isabel, N., Bousquet, J., 2012. A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. *BMC Biology* 10. <https://doi.org/10.1186/1741-7007-10-84>
- Peakall R.O.D., Smouse P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28, 2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Phillips, S.J.; Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31, 161–175.

- Poland, J.A., Brown, P.J., Sorrells, M.E., Jannink, J.L., 2012a. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7. <https://doi.org/10.1371/journal.pone.0032253>
- Poland, J.A., Brown, P.J., Sorrells, M.E., Jannink, J.L., 2012b. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7, e32253. <https://doi.org/10.1371/journal.pone.0032253>
- Porfirio, L.L., Harris, R.M.B., Lefroy, E.C., Hugh, S., Gould, S.F., Lee, G., Bindoff, N.L., Mackey, B., 2014. Improving the Use of Species Distribution Models in Conservation Planning and Management under Climate Change. PLoS ONE 9, e113749. <https://doi.org/10.1371/journal.pone.0113749>
- Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., Lareu, M., 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. Frontiers in Genetics 4, 1-13. <https://doi.org/10.3389/FGENE.2013.00098>
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
- Reis, S.L., Couto, C.s., Pinheiro, C.S.; Espada, A.L.V., Lima, J.A., Lentini, M.W., 2013. Técnicas Pré-Exploratórias para o Planejamento da Exploração de Impacto Reduzido no Manejo Florestal Comunitário e Familiar. Belém: Instituto Floresta Tropical, 150 p.
- Santos, E.A., Souza, R.M. e, 2015. Territorialidade das catadoras de mangaba no litoral sul de Sergipe. Geosaberes: Revista de Estudos Geoeducacionais 6, 629–642.
- Skogen, K.A., Jogesh, T., Hilpman, E.T., Todd, S.L., Rhodes, M.K., Still, S.M., Fant, J.B., 2016. Land use change has no detectable effect on reproduction of a disturbance-adapted, hawkmoth-pollinated plant species. American Journal of Botany 103, 1950–1963. <https://doi.org/10.3732/ajb.1600302>
- Stockwell, C.A., Hendry, A.P., Kinnison, M.T., 2003. Contemporary evolution meets conservation biology. Trends in Ecology & Evolution 18, 94–101. [https://doi.org/10.1016/S0169-5347\(02\)00044-7](https://doi.org/10.1016/S0169-5347(02)00044-7)
- Stockwell, C.A., Hendry, A.P., Kinnison, M.T., 2003. Contemporary evolution meets conservation biology. Trends in Ecology & Evolution 18, 94–101. [https://doi.org/10.1016/S0169-5347\(02\)00044-7](https://doi.org/10.1016/S0169-5347(02)00044-7)
- Tax, D.M.J., Duin, R.P.W., 2004. Support Vector Data Description. Machine Learning 54, 45–66.
- Tosso, F., Hardy, O.J., Doucet, J.L., Daïnou, K., Kaymak, E., Migliore, J., 2018. Evolution in the Amphi-Atlantic tropical genus *Guibourtia* (Fabaceae, Detarioideae), combining NGS phylogeny and morphology. Molecular Phylogenetics and Evolution 120, 83–93. <https://doi.org/10.1016/J.YMPEV.2017.11.026>
- Tropicos, 2018. Tropicos.com.org. Disponível em: <https://www.tropicos.org/home>. (acesso 15.09.2021).
- Ueno, S., Taguchi, Y., Tsumura, Y., 2008. Microsatellite markers derived from *Quercus mongolica* var. *crispula* (Fagaceae) inner bark expressed sequence tags, Genes Genet. Systems 83, 179–187. <https://doi.org/10.1266/ggs.83.179>
- Ulaszewski, B., Meger, J., Burczyk, J., 2021. Comparative Analysis of SNP Discovery and Genotyping in *Fagus sylvatica* L. and *Quercus robur* L. Using RADseq, GBS, and ddRAD Methods. Forests 12, 222. <https://doi.org/10.3390/F12020222>
- United Nations, 2021. The Global Forest Goals Report 2021. New York.
- Upadhyay, A., Gaonkar, T., Upadhyay, A.K., Jogaiah, S., Shinde, M.P., Kadoo, N.Y., Gupta, V.S., 2018. Global transcriptome analysis of grapevine (*Vitis vinifera* L.) leaves under salt stress reveals differential response at early and late stages of stress in table grape cv. Thompson Seedless. Plant Physiology and Biochemistry 129, 168–179. <https://doi.org/10.1016/j.plaphy.2018.05.032>
- Vieira, R.F., Camillo, J., Coradin, L., 2016. Espécies Nativas da Flora Brasileira de Valor Econômico Atual ou Potencial: Plantas para o Futuro - Região Centro-Oeste. 1. ed. Brasília: MMA, v. 1

- Weir, B.S., Cockerham, C.C., 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* 38, 1358–1370.
- Wickham, H., 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wright, S., 1965. The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. *Evolution* 19, 395–420.
- Xing, L., Zhang, D., Zhao, C., Li, Y., Ma, J., An, N., Han, M., 2016. Shoot bending promotes flower bud formation by miRNA-mediated regulation in apple (*Malus domestica* Borkh.). *Plant Biotechnology Journal* 14, 749–770. <https://doi.org/10.1111/pbi.12425>
- Zhao, S., Guo, Y., Sheng, Q., Shyr, Y., 2014. Heatmap3: an improved heatmap package with more powerful and convenient features. *BMC Bioinformatics* 15, 1–16. <https://doi.org/10.1186/1471-2105-15-S10-P16>
- Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., Farooq, H., Herdean, A., Ariza, M., Scharn, R., Svantesson, S., Wengström, N., Zizka, V., Antonelli, A., 2019. CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* 10, 744–751. <https://doi.org/10.1111/2041-210X.13152>
- Zwaenepoel, A., Diels, T., Amar, D., van Parys, T., Shamir, R., van de Peer, Y., Tzfadia, O., 2018. MorphDB: Prioritizing Genes for Specialized Metabolism Pathways and Gene Ontology Categories in Plants. *Frontiers in Plant Science* 9, 1–13. <https://doi.org/10.3389/fpls.2018.00352>

7. CONSIDERAÇÕES FINAIS

Para desenvolver novos estudos e pesquisas de forma propositiva é importante conhecer o contexto no qual está inserida a área estudada. Uma boa ferramenta são as revisões sistemáticas de literatura, que empregam um processo metodológico e ajudam a reduzir o risco de viés de estudos incluídos. Nesse contexto, para o desenvolvimento desta tese, no primeiro capítulo (Artigo 1: Potencial farmacológico, alimentício e dados moleculares de *Hancornia speciosa*: uma revisão sistemática) foi realizado um mapeamento científico, tecnológico e molecular sobre a espécie estudada com o propósito de aprofundar o conhecimento sobre o estado atual da pesquisa e contribuir para elucidar questões como: Quais são os compostos bioativos naturais da espécie *H. speciosa* e suas funções? Existem produtos tecnológicos desenvolvidos a partir dessa espécie? Quais instituições/pesquisadores estão trabalhando com a espécie? Quais aspectos foram investigados? Existem dados moleculares disponíveis em bancos públicos?

Ao investigar tais questões, verificou-se que a espécie é ativo com potencial a ser explorado na estratégia nacional de bioeconomia, pois é uma fonte de alimentos com qualidade nutricional e funcional, além de possuir compostos bioativos para uso farmacológico e para a conversão em bioprodutos. Foram identificadas também as principais lacunas a serem consideradas em estudos com a espécie e estas direcionaram os capítulos seguintes.

Especificamente no Capítulo 2 (Artigo 2: Diversidade geno-fenotípica em uma população natural de *Hancornia speciosa* Gomes: implicações para a conservação e melhoramento) foram avaliadas a diversidade genética de árvores e a variabilidade fenotípica de frutos, sementes e plântulas em população natural. Essas informações são relevantes no direcionamento das estratégias de manejo, conservação, domesticação e melhoramento, as quais fazem parte das principais lacunas sobre a espécie.

As informações sobre as sementes e plântulas são importantes considerando que um dos primeiros passos para a domesticação se dá pela propagação das espécies. O conhecimento da diversidade genética pode ser aproveitado na etapa do diagnóstico para o desenvolvimento de um plano de manejo, além de útil para indicar matrizes mais divergentes geneticamente, as quais se recomenda para coleta de sementes usadas no enriquecimento de áreas naturais ou para o melhoramento.

No Capítulo 3 (Artigo 3: Genômica e estrutura genética de populações de *Hancornia speciosa*: fonte de alimentos funcionais e fármacos) buscou-se contribuir para ampliar o conhecimento sobre as áreas com maior adequabilidade ambiental para o desenvolvimento da espécie e sobre a diversidade genômica e estrutura genética de populações amplamente distribuídas no país. Nesse estudo foi aplicado pela primeira vez o protocolo GBS, que foi eficiente na identificação e genotipagem de marcadores SNPs. Com esses marcadores foi possível avaliar a diversidade e divergência dentro e entre populações com maior poder de resolução, gerando novas informações.

As sequências do genoma nuclear obtidas neste estudo serão as primeiras depositadas em bancos públicos de dados moleculares para a espécie, e podem servir como referências para o desenho de *primers* específicos e usadas em estudos *in silico*, esses estudos por sua vez, podem contribuir para esclarecer os constituintes que conferem atividade bioativa aos frutos, extratos foliares e látex.

Novos estudos podem concentrar esforços na investigação de genes e rotas metabólicas que estão relacionadas ao potencial farmacológico da espécie, bem como no desenvolvimento de bioprodutos.

ANEXOS



RESEARCH ARTICLE

Geno-phenotypic diversity in a natural population of *Hancornia speciosa* Gomes: implications for conservation and improvement

Valdinete Vieira Nunes · Renata Silva-Mann · Juliana Lopes Souza ·
 Crislaine Costa Calazans

Received: 28 September 2020 / Accepted: 1 March 2021
 © The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract *Hancornia speciosa* Gomes is native tree to South America, and its fruits are a source of income for communities and the food industry. The reduction of natural occurrence areas of *H. speciosa*, the absence of a sustainable management plan conservation and breeding programs, threaten its natural populations. Therefore, the purpose of this study was to evaluate the phenotypic variability of fruits, seeds, and seedlings and to estimate the genetic diversity. The main goal is to provide information for the conservation and breeding programs of the species. Measurable variables of fruits, seeds, and seedling; and genetic diversity was estimated by ISSR markers. Among fruits, 53.21% had a longitudinal diameter from 27.5 to 35 mm and 63.15% had a transversal diameter from 24.5 to 35 mm. The average seed biometry was 9.98 mm, 8.21 mm, and 3.97 mm, for length, width and thickness respectively. The length of seedlings shoots varied from 3.0 to 8.5 cm, roots from 4.0 to 15.8 cm and dry mass from 80 to 103 mg seedling⁻¹.

The genetic parameters estimated were $H_e = 0.40$, $I = 0.50$, $N_a = 2.0$, $N_e = 1.71$ and similarity of 55.29%. The population has high phenotypic variability of fruits, seeds and seedlings, and genetic diversity. There is potential for use of this population in the enrichment process for the restoration of natural populations.

Keywords Conservation · Phenotype · Brazilian fruit tree · ISSR · Mangabeira

Introduction

Among the main threats to biodiversity are forest fragmentation and the over-exploitation of forest genetic resources without a sustainable management plan (Kim et al. 2015). Therefore, the sustainable management of remnant vegetation is relevant to protect forest ecosystems and plays an important role in conserving biodiversity, especially in poorly developed regions (Novello et al. 2018) which involve detailed knowledge of the genetic diversity of species.

At the same time identifying genetically divergent individuals is relevant to breeding programs. Since the measurement of genetic diversity begins by estimating such diversity and its magnitude, these aspects are essential (Heidari et al. 2019) to population maintenance.

Part of Doctoral Thesis of the first author.

V. V. Nunes · J. L. Souza · C. C. Calazans
 Post-Graduation Program in Agriculture and Biodiversity,
 Universidade Federal de Sergipe, São Cristóvão,
 Sergipe, Brazil

R. Silva-Mann (✉)
 Department of Engineer Agronomy, Universidade Federal
 de Sergipe, São Cristóvão, Sergipe, Brazil
 e-mail: renatamann@academico.ufs.br

Published online: 14 March 2021



REVIEW

Pharmaceutical, food potential, and molecular data of *Hancornia speciosa* Gomes: a systematic review

Valdinete Vieira Nunes · Renata Silva-Mann · Juliana Lopes Souza ·
 Crislaine Costa Calazans

Received: 30 March 2021 / Accepted: 8 December 2021
 © The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract *Hancornia speciosa* Gomes is a fruit and medicinal tree species native to South America, which in Brazil is considered of potential economic value and priority for research and development. We present a map of the state-of-art, including articles, patents, and molecular data of the species to identify perspectives for future research. The annual scientific production, intellectual, social, and conceptual structure were evaluated, along with the number of patent deposits, components of the plant used, countries of deposit, international classification and assignees, and the accessibility of available molecular data. Brazil has the most significant publications (306) between 1992 and 2020. Technological products (29) have been developed from different tissues of the plant. Most of the articles and patents were developed by researchers from public universities from different regions of Brazil. The molecular data are sequences of nucleotides (164) and proteins (236) of the chloroplast

genome and are described to identify the species as DNA barcodes and proteins involved in photosynthesis. The compilation and report of scientific, technological, and molecular information in the present review allowed the identification of new perspectives of research to be developed based on the gaps in knowledge regarding the species and perspectives for the definition of future research.

Keywords Mangaba · Biodiversity · Bioeconomy · Ethnopharmacology · Scientific mapping · Patents

Introduction

Fundamental concepts within Latin America have led to the adoption of bio-economy principles, which consists of the evolution to a new economic system based on the sustainable use of resources (Lewandowski 2018), with different levels of socio-economic impact. They highlight the valorization of biodiversity resources in medicine as one of the critical areas addressed (Sasson and Malpica 2018). Given economic limitations, there is a demand for low-cost, accessible, and effective medicines that can assist as an alternative or complementary medication for healing diseases, especially for middle-high, medium-low, and low-income citizens (Goyal and Ayeleso 2018).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10722-021-01319-w>.

V. V. Nunes · J. L. Souza · C. C. Calazans
 Post-Graduation Program in Agriculture and Biodiversity,
 Universidade Federal de Sergipe, São Cristóvão,
 Sergipe, Brazil

R. Silva-Mann (✉)
 Department of Agronomic Engineering, Universidade
 Federal de Sergipe, São Cristóvão, Sergipe, Brazil
 e-mail: renatamann@academico.ufs.br

Published online: 18 January 2022

Springer

Sequências de SNPs *outliers* identificados para *Hancornia speciosa* Gomes

>CLocus_547GCATATTGTTAGATACAGTGATGAGTCAAAGGCATACAAATCGTATAATCCATTGTTAGAAACGATGA
>CLocus_702 TGCATAGCTTGCACACGTCCTTGTAGCCATAATGTATTAGAGACACTGCCATGCACGTTACGTGTTTT
>CLocus_839 TGCATATTACGAGTAGACAAAGGTGGTAATTTCGGCTTCATGGCATGTTTATTGATTTCATTGGATTGAT
>CLocus_1353TGCATGAAACCGAGTCTTCGGAAAAAGTAGGCACGCGACCACGTAAGGCTATTGATCAGTGAGAAACGGAA
>CLocus_1501TGCATAAACTATCCCAACAAAGGCCAAAAATAGTTTGTAAACGCCACACTCCCTTTTACCATTGGAACAA
>CLocus_1599TGCATGATACAACTGCAATTACGCCAATATGCTATGAAAGACTTTGTCCTTACACCAATTTTTTACAG
>CLocus_1625TGCATTTATTGGTTACTAGAATATGAAATAGATATCAAAAGAATGAATTCAATTAGGTCAACAAGACAAT
>CLocus_1775TGCATATGATCATTGATGAGTCGCAAAAGGCCCTTCTACGCAATAAACTTATCACTGATAATGCTATCAT
>CLocus_1929TGCATTGACATTTTCGATGCGTACAAAGAGAAGTAAACAGCTATGCAATGGAATCAGGATTTCACTTCCT
>CLocus_3160TGCATGTAACAAATTTTGGTGTATTAGTGGTAACCTCAAATGTACCATAAGTGTAATAAAATCTGTTTAC
>CLocus_4176TGCATCATTAGATCTGAAATGCTGTGTGACGTGACGCTGATAGGGCAAAATCTTAAGTATGTAGACACTTT
>CLocus_4190TGCATAAATTTAGGAATATACATTTCTTTGCCCTCACTCAAACTTACCATTGTTTATATAGATATAAGTA
>CLocus_5678TGCATAATTTCAATGACCGCTCCATAAGCTGCTATAGTTTCATAGACCCAGGTTGCTTGGAAATTATG
>CLocus_5718TGCATACAAAAATTTCCGCATCAGATCGAACCTATAAAATATACTCCCTCCGGTCACAATTTATATGTCAC
>CLocus_5917TGCATATTTGATTCACTTCCCTGGAGTCTTGGTCTGTAGTCTCATCGTTCATCAACTTATATTTGACC
>CLocus_5966TGCATACATTTTACACACACAATGCACAGATTTTACACACACAATGCACAAATGACCATTCGACCTTT
>CLocus_6238TGCATTTTCAGATTGCTGCTAGCTGTTCCAGAAATGGAGACCTCAGACTGCAAAACAGGACAGGCTTAA
>CLocus_7015TGCATATAAAGGCGTGAAGTGAAGTGTAAAGTCGCTTGTATTTTCGCTAAGAGTAGATCAAAAA
>CLocus_7664TGCATGTTGTGCAACTCAAAGTGAATGCCGCGGGCTAGTGATCATCATATCCATCGCTCCATTACGA
>CLocus_7694TGCATGAATGGAATGACGAGGATCGTCTAAAGCCTTGGAAAAAGATGAGTGACATTTTGGGGACCATTTG
>CLocus_7829TGCATATCATAGAAACCTTGTACGTCTTGGTACATGGAAGCATTTGCTCGAGAATTCGTTAGTTTTCCTG
>CLocus_7922TGCATGAAAAATGATGATTCAAATCTAGAAAAATGTCACACCATCATCTTATGGATCCAAATGGAGTTTCAAC
>CLocus_8456TGCATTGTGTCTTAGAACTAAGTAGGACATTTGGTTTGGTAAAGTATATTAGCCCTAATTGGCCTCTGA
>CLocus_9061TGCATTTTCTAAGTAAGCAATGCCAATGCATCCATGTATATAACAGTGTCTCTTCTGTAGAACCTCG
>CLocus_9270TGCATCATCTAGGTTAGCACTGATTACAGACACTGCATTATGCTGAGTCTCTGCTTACAATTACAGA
>CLocus_9348TGCATAATGGGGAAGAAATCGTTGCTAGAGTAGGTTTGACATGGTTGGATATATCTCAGAGATTTAGTG
>CLocus_9544TGCATGGTCTTGGTGGTAGGCTTTACGGATCTTTGAGAGGATGCAGAAAAACAACTAGGCCGAATCA
>CLocus_10280TGCATGCAAAAGAGGATGTAAGTTTCCAAAACAGCAGTAGGCTCCCATGAAACCTTTTGTTTTATTCTTT
>CLocus_10312TGCATCTACCGCGCTCTTTTCTTCACTTTTCTTTTACTCAGGAATGTGAGGAAAGTAGTTTTC
>CLocus_10743TGCATCAGGTCTCAAAGATATGCTTACTCTGAGCAGACACTTGCACGCGTATGTATCATGGTGATCAA
>CLocus_10868TGCATATGTTGAATGGGAATGTACATCACGTGATGACTATTACTACTCTCAATACATCCGCAAGTCGT
>CLocus_11280TGCATATTTGGGAAAACTCTTATATGATTTGCAACAAGCTAGCAGACAAATGGAATGCTAAAGCCCTGTCT
>CLocus_12559TGCATAGTGATGCAAAATCTAGAACTTTACAATCATTCACACAAATTTGGCCGCAATGAAGTTGCTTG
>CLocus_12752TGCATAGGATATATGTCGGTGTGATATCAAGCTGCATGTACATGTGCAATCTTTCGCTCTTGCCCAAA
>CLocus_13196TGCATTGTCTCTCATGGGCATGGGCTTGCAGCTTGGGCAAGATCTTGAATGGTGCCACACTGAAAT
>CLocus_13476TGCATCCTCAACATTTCTCTTGTAGCACCACATGCTTTGCATTCTCAAAAATTTGGGTAGAGGATGAAT
>CLocus_13845TGCATAAATTTGTGCTGTTGATGGTGTAGCTGCAATATTCTTGCTATATTGCTGTATGTTTTACAGA
>CLocus_14433TGCATAAAATCTCAAGGGATATGCAAAAGGTTGTGAACCTAGTTTGGGCGAGCTTAGAGGAAGTAAAAA
>CLocus_14584TGCATAGAACGTAATTTAGTTCTTCTTCTTGTATCCCACTGTGCAAGTTGGCATATTTGTCATGAAGGC
>CLocus_14959TGCATACATCCGCATAACTAAGGAGGGAACCTGCATTTCACTTCGTTCTGAGCAAACCAAGACCAAA
>CLocus_15257TGCATTTCTCTGTTTTCGAAATGTTTCTATGTGTAATGCTGTTGAACAAATCATTTGTACGGGCTTT
>CLocus_15323TGCATGGTCAATTTGATGCTGCAATTTGGAGTCTTCAGTAAATGGTAGAATCAGGCATAAGGCCCAACGA
>CLocus_15507TGCATCTTTTGAGTTGATTCTTGAAGTGATTTGACACTAGGATGAAGATTCTCCACTAATGAGATGCT
>CLocus_15958TGCATTCGCTGTGCAATTTACGAAGTTCAAGTGAGTAACCTCTCAGATTTCATGCAACTCTCAAAGTTACAG
>CLocus_16375TGCATTTTCTTTTAGCTTTTAGCACTTCCCTTGTCTGTTTGAACCTTTGAGTTCAAGTTGAAGTGAAACT
>CLocus_16698TGCATATCAATGCCATCAATCGCTGAGGAACTGAACGTAGCATACTTCTTCCAAGAACAGGATCCCA
>CLocus_17772TGCATATGGAGGTTGGACCAACCGATGAACCGCCCTTCAGGGGTGACCATAAATAGTTCTACGGCTC
>CLocus_18338TGCATGAATGGAGATTGCAAGAACCCTGGAGAAACGAAAGAGATGGCCAAAAAGTACGGACACCGCGTG
>CLocus_18726TGCATAAACCACAACTCCAAACGCTGATCATAAGAACAAATGTAGCACCTGAGTGAGGCCAGGCTTGCT
>CLocus_18982TGCATAAATCTGAGGTTCTGCAATTTGACAGCTGAAGAAGCCTTCTCATCTATTCTGGTGATTCCCA
>CLocus_19486TGCATAAATATGATCAATGATAGAAAAAGCTATCCTTCCATTACCTCACTAAAGAAAGTTAGGGCTA
>CLocus_19620TGCATTAGCCCTTTCTCGTAATTGCTTACCATGACCATAAGGTCACCAATGCTAAAGAGTATGCTTGA
>CLocus_20193TGCATAAAATTTACAGGCAGCTATTGCCAGAGTTGCGACATGTGGAATAAGATACACGAACTCAAG
>CLocus_20625TGCATAGTGCTTACATGCAAGGCCAAGATTGTGGGAGTCTGTAACGGAAGTGAAGCTACACAACCTG
>CLocus_21265TGCATACATGTTGCTTCAATATTCTGCGTACTGTTTATTTTCTTTTGTCTGTTGTCAGTATTTAGT
>CLocus_21306TGCATAAATTTGAACGTTTATCATACATATACCCGGTGCCAAATGCTTGAAGTTGAGGACGAACATTTTA
>CLocus_21436TGCATCCCAATTTGAAGTCAATGTGGCAGGCAGGTTGCTAGATTGTGCTTTCATGGCAATTAGGCCCAAAAT
>CLocus_21970TGCATAAATTTGATAGGAAGTTGATTTCTGCAATATTTTCAATCATCTCAACATCACAGCGCTTCTGCA
>CLocus_22990TGCATTCCTGCTCTTTCATTGGTTGCTATGCTTGTATCTCTGACTCATGCAGCGTATCTATGTGCATTA
>CLocus_23247TGCATAATCTAATTATCATCAATCCAATGGCATGTGATACATAAAAAAGTGAAGTTGGGTAATAGCTTTTC
>CLocus_23745TGCATGATGACCATTTCTCACAAACCCGACATCTAATTTACGAGGTTGCCTCATGTACATTTACAGA
>CLocus_23954TGCATTTTCCAAATTTCTTCTAGCAGAAAGAAATAAAGAGCTCAACGCCATAAAAAAGTCTTACAGAT
>CLocus_24702TGCATCTATGTAAGAACTAAAAAAGGTATAGTCCCTTAGGAAGGACAAGGTGATCCTGATGATTAGAAAT
>CLocus_24975TGCATATGGAGTGTGTGTGGGATGGGGTTTACGATTCTACTTATGATGGCGTGGGTTGAAAAATTACA
>CLocus_25121TGCATGTTAGATGAGTTTTCATCCCTTCTGAACCTCATAGACCGGAAAGTTACCCGTATGAAGATTATAT
>CLocus_25468TGCATATACATTTTTCACAAACATTTTGGAAAGGTGGAACCTCACTTGTAGCAAGCTCATAGCGAATAA
>CLocus_25893TGCATCATTTCTAACTAGTCAACCAATCCAACTTCTTTCAGCCGACATTGCCATGAACATTTCTCTCCA
>CLocus_26459TGCATCTTAAGGAACTCAAAACTGCATAGGTCCGTGGAATATAAATGGCTATTGCTCATATTTT
>CLocus_26677TGCATCCCAATAATCAGTAATCAAGAAATGATCAACCAACCAAAATATTGCAGGAAAAAGTGATCTGCATAG
>CLocus_26684TGCATAGTAGGGGTTGAGTACTCATTTGCATAAGGGCTAGCGAGTCAATCAAGCGATTACAGAACTCTTA
>CLocus_26742TGCATTGGACGGAAGGAAAAAGGAACCAAAAGAGTGGCAGGTCCGATACCTTCCCTGGTCAGAAAAAGGA
>CLocus_26756TGCATGTTGGTATCTGGTAGAGAACCGGCCAGGAATGAAATGGCATTCCCGTATCAGTGATGTCGTGTTT
>CLocus_27029TGCATATGCTATTGCGGAGATTGCGCGCAGACCTATCTTATTGGCGTTGGGAAATTCAGATGTTTACAGA
>CLocus_27169TGCATAGGACGAAAGACAGCTACAAAAAGGTTTCATACGGACACATTGCAGAAACGTTGCATAAATTACAGA
>CLocus_27546TGCATCAACCCGAAATATAGCATGACATTACCTTACCTTGACCTGACTATTTCTTGTAGTCCACATCA
>CLocus_27893TGCATAAAACCTTTGACCTTGTCTGTAGAGGCCCTGATAAGACCATGTGCCGAACCATTTGGGACCAT
>CLocus_28292TGCATGAAATTTAGGTGTTGAAGAACTATAAAACAATGTAGTGTAATGTTCCGATGGTAAAGATCGAGTGGG
>CLocus_28531TGCATATTTACACGTAACCATGCTTGTCTTCTAGATCAAAAGTTAGATACATATGCTTCTATTACAGA
>CLocus_29077TGCATTTTTCATGTAGGATTGATGGTGTGTTGCGAGATTATGAAAATTTACTTCAACATATTTCTCGGT
>CLocus_29204TGCATGTTGACGCTGTAAGAGGCGTTTTTCTCTACTTACAGCAAGGTACTTTGTCAACATATTTGCGCC
>CLocus_29264TGCATTTCCAAAATGTATCTTTGCTATGAAAGTTTGTAGGACAGTTTGTGGTAGACATGATGTTGCTT
>CLocus_29408TGCATATTGCTGTTGATGTAAATGGAGCGTGAATGACACGCCAGGTAGCAATCAATCAGGTGGCAGATGA
>CLocus_29626TGCATGAAGTCAAGGCTCTGGATCTTTTGAATACTCCTTGTACGTATCTTTTACTATACTTTTACAG
>CLocus_30901TGCATAGAGTGGACCTTGGAGCTGAGTATAAATTAGCTTCTGCATCAAAAGGAGCCAAGGCTCTGGC
>CLocus_31250TGCATTTATTTGACACAATAGGCATCTAGCCTTGATGGTCCAGATCACAGAAATTTGACATGAACAATCCC
>CLocus_31335T GCATGGATTGTAACATTTGACAGAAAAAGTCACTGCGCTGAGCGTGGACTACTATTAGATTCACAACAAT
>CLocus_31656 TGCATTTGACATTTACCATTTGGAAGGCCCTTCAACAAATCTCCGTCAACTCTCGATACCTCAATTTGATC
>CLocus_31847 TGCATAAATGACGTTGGAAGATCAACCGATAAGTCTCACACCTCAGAAGAACCCACATCAGCATTTACAGA
>CLocus_31999 TGCATTGCTTGATGTATGTAAGGCACTACTTGCTAAGTGTGTTGTTACAGATAGTGACAGACAATGGGC
>CLocus_32595 TGCATCATATCTTTGGCTTTAGGAAATGAATCAGGCTATGGACCTAATCATGCTGCATTTGCTGGTATGT
>CLocus_34886 TGCATTTATGTAGCACAAGTAAATAGACCATAGACGTAACCTCAACAAATCCCATCCAAAATACCTGCG
>CLocus_34943 TGCATTGAATATTTTGTATGATTTTTCGAGAACCCCAAAATAGTTGGAATTTGGATCGAGTGAATGG
>CLocus_37049 TGCATAAGAATTATACAAATATTGACTACAGAAGTACTAGTAGGATTTTGTGCTTTTCGGACATTAGTGT

>CLocus_37361 TGCATTTTGCATTGAGAATTTGCTCAATCGATAACATTTGATGTTCAAATTCACCTGCCAGACCGTGAGA
 >CLocus_39922 TGCATTTCACTACCTTACACAGCAGCTTTGGGACGCGGAAACCCCTAGCTTCTTTTGCTCTTTTACACCCC
 >CLocus_41261 TGCATTCAGATTGGCTGCGGGGCAATGTACGAGAAATGGACCGGCTGAATTCGATTGAATTTGGACGGGAGA
 >CLocus_42437 TGCATTTCAACTGACAGTCTTACTCTTGCTATGATAGCATGTTAGACTCCCACTCTTTTCAGTAGTTA
 >CLocus_42496 TGCATGTCGTCGCCGAAATTTGAGATAATTTGTCATCGATTGGAACAAGAATTCAGACGATTTACAGA
 >CLocus_42783 TGCATGATCTTTTGTATGTATGTAAGGAAATTTTCTTGACCAAGTTGACTTTCTTGAAAAATGTGCA
 >CLocus_42938 TGCATCTACCTCTTGCTAAGAAAGGAGCCATCGATCCGCTTAGCATGCCAATCGTGCTCAGAAATG
 >CLocus_42943 TGCATTTGTGTTTAGTTTCTGTCACGCAGCAAACTTGCAACAGAATCATAGGACCGATGACTTATGCTG
 >CLocus_43567 TGCATAGGTACCCAGGTTCTTGCTCGACCTCTAATCCATACATTTTATATCACCACCTTCTCTCCCA
 >CLocus_43623 TGCATTTCAAGATTATGAAATGTAGGAGTGGTATCGGCAGCTTTAGATGGGACGGAGCTGAATCAGATAGT
 >CLocus_43803 TGCATTTGCTTGGTGTGCTTGCCGAGGGAGTTCTCAGTAAAGCGCTTTGCTTGTAGTGAAGAGATTAGT
 >CLocus_44207 TGCATTATATTGTGACATTTGCAAGTGCAGCTGGGGGACACCCATTGTGACGCTAGGGAAAGTCAAGATA
 >CLocus_44366 TGCATTGTAATAAGGTCCTCCATCATAAGCTGTCCAAGTCTCTTAGCAGCGTCAGTTGCATATATCTGAG
 >CLocus_45044 TGCATTTCTGGGTATTGGTTGGAAGACTATTGAACAACAGGGGTGTGCCAGATATGTTCCACATTTCCA
 >CLocus_45827 TGCATTTCTGCTGCTAGATGTCTGTCCTCAGAAAGTCTTAGTGACAAAGAAACACAGGTTACAGA
 >CLocus_46763 TGCATACCCAAAAAATTAGGGGGGAGACCAAGAGGTATCACAACCTTTTACACAATTTACGACGACATTT
 >CLocus_46938 TGCATTTTACAGGGTCGATGTATGGCATTGATTCAAGCTTCATCTGACGTACAAGAACAAGGAGCCGTG
 >CLocus_48562 TGCATGCTTGGATACACATCAAGGCATACACATCTCAATTCAGATGCTTCAAGTGGTCTCGACA
 >CLocus_52149 TGCATCTGATCTTTGAACCTGAAACCTCTTATTTGGAACAAGAAGGAAATTCACCCCTTGAGGGCCAT
 >CLocus_61580 TGCATTTGCTCTTCTCATTTGTCCTATCCCTGAACGAAATCTTGATGAATAGAGAGATACAAGATGCT
 >CLocus_67731 TGCATGGGTGAAAGGTTCTACCAATGATGGACTTCTTATCAAAATCATTTTCATATGTGTGATTCTTGGTA
 >CLocus_1033 TGCATGACACAACAGCCATAGGGCTTTTCCCATTTGCTTTTCTGCTTCTTCTTATCTTATTAGTAGTTG
 >CLocus_10453 TGCATGAAGACACAGAATTCACCAATTTCTTGGAAGAGGAAGCAGAGAAACAAGAATTCCTGCTACCTAG
 >CLocus_10594 TGCATTTCTATGTTTACCCACACCTCAAATTTGCAATTTCTGGTCATTGAGAATCATGCCAATTCGGATTA
 >CLocus_11619 TGCATTATTTTCAGTTTATATCTTCAAAAAAGATGAGAGAATTAGGATCGACAAGAGCAAGACACATTAC
 >CLocus_11751 TGCATAACTAGTGGAAATCAATTAGTTTTATGTTGGGATTGTGGAAAAGTTGTTGAGTTCTGCTGTGAAAT
 >CLocus_11764 TGCATTTCTTCTTTTCTTTCCCCCTCTCTTTTGGCACTCTCTCCCTTTCTCTCCCTTTCCCTGGGAACA
 >CLocus_11812 TGCATCTACACAAGACGTATATTTACAGGGACTATTGGAGCCCTAATAGGACTTAGAGTCCAATTGGGA
 >CLocus_11890 TGCATATTGGAACATAATCACCATAATGAGATGCACACCAGGAAGAGAGAAAAAGGGTAAGGAGGAG
 >CLocus_12602 TGCATTTGCTGATCAGAAATCAATGGCTGAAGGTAGCAACATTGAACCAATATCTTTAGTTCATAAGCT
 >CLocus_12733 TGCATTTCCATTTCTGTCAGATGGATTATGAAAAATCCAGAGAGCTACTCAACGTATTCTATGTCCTA
 >CLocus_12791 TGCATGCTTTCAATTTCTCAATGTCTATTTATCTTTTGTGCTGCTGGCTTTTACAGTTACTTTTT
 >CLocus_12955 TGCATAAATTAGACGATCCCGTACAGAACCGCCCAAGATTAGGTGGCTGAACAGTCATAATTTATTAC
 >CLocus_13028 TGCATTTGCATATTTTGTCTTTTATTGTGGGAGTCCAAGTGGCTTCTTGGTGGGCTTCAGATTTTTACAG
 >CLocus_13083 TGCATCTCTACTTAGCATTTGCTCTCAGCAATAACACCCAGACCAATGCACTGAAGGACCAAAAAAGTGGCTG
 >CLocus_13193 TGCATTTTACGCTCTCTTGGATTGGACAAAATTTCACTGTCCAATACATGCTCAGAAATAGATCAGACT
 >CLocus_13315 TGCATAGGTTTTATGTTTATCTCCTACGTAAAAGGAAGAAACAGCAACAGCAGAGCCCATGCAGACAAG
 >CLocus_13491 TGCATGCCATTGGGACAGCCGAAAGGGAGCTGGCAAAGACAGCCGAATTCAAAATTTCTTGAAATTTGCT
 >CLocus_13843 TGCATACCTCTTCCCAAGTTCAGGCAATGCTAACATAGGTCGACTAACCAATCTCTTCTTTAGCTCTTAC
 >CLocus_13931 TGCATTTCTGAGTCGATGTCAACAAGGAATCACTAGCATTTATCTCTCAAACCTAACAGCAGCTTTGAAAAG
 >CLocus_14078 TGCATAAATCTCATGAAATCGTACATTGTCAGCCCTACCAACTAATACGCTGTAGAAGATGAATACCA
 >CLocus_14336 TGCATCTTGATTTTCAACGGACTTTATGTTGCAAAATGGACATGGACATTCAAACGGATTTTAGCTTACA
 >CLocus_14338 TGCATGCTCAAAAATGATTGATTGTCTCCGTCATCCGAGAGAACTTACATTTTGTAGCCAGTTACAGA
 >CLocus_1448 TGCATTTCTTGTTTTCAACTTGGAACTGCGGTAGGATAAAGAGTAAGAGCTGACCTTGAGTGCACTATA
 >CLocus_14538 TGCATTTCTAGTCTGTTACAACGTATGCAACCTTCAATTACTCAACCCATAAAGACCAAAATGGCTTTA
 >CLocus_14657 TGCATTGTACGAATGCCAGTAAAGATGGAATACTTTTCAACAAAATCAGAGGGGAAGCAAGTGCCCAA
 >CLocus_15092 TGCATTCAGGATTCTCATGTTGAAGCAGAGCTATCCTAACACCAAGTTGAACAAGTGTGCTAGGATCAGG
 >CLocus_1584 TGCATTTCAACAGCAATTTGAACTTTATCAAAGTTAGGGGGAAAAATAGTAGTTTGTAGCTGAAAAATTTGA
 >CLocus_16062 TGCATTTGCTGTCATGAAATCGTACATTGTCAGCTTGAATCTCTTCAAAATTTATGGCAGTTCCACTACACGCAACG
 >CLocus_16133 TGCATACAAAGATAGTTATGATGATAGTTATGGGAAAGCAATTTGAACTCCCCCAAAACACCATCTCTGTA
 >CLocus_16158 TGCATATTGTCAATCATCATATGACCCCGAAATTTGGGGAGTTCCGGAAAGTCCGCAAAAGAAATTTACAG
 >CLocus_16394 TGCATTTACAGGCTCGTTTCATCGATCTTGTCTACTAACACCCAAATTTCAAGGCTAACCCCTTTAGTCTC
 >CLocus_16651 TGCATGATATTCAGAAACGGTTGGGGGTATGCAACCATGTCTCAAGGCTCCCTTAGGTTGTAGTTCAA
 >CLocus_16770 TGCATATTGAGGCAGCGTGTGATTGACTTTCCAGCTCAGCATACCCGATGTTTTCAGCTTAGTTAGATTCA
 >CLocus_16810 TGCATAGATTGTCCACTCTTGGGAAAAACACATCTAATTTGACCTTTCCACAATGTCTCTCTTTACAGA
 >CLocus_17088 TGCATAACATCTAAGGAGTGATTGTGACTGATTCATGATTTTGGTTGGGGAATTTGTAGGCTCATGCT
 >CLocus_17409 TGCATCAAACTGACGCGCATCCCTTCCCAACAAAATCTCAACTCCAAATGGCAGCAAAATAAACACGATA
 >CLocus_17524 TGCATCCATTGCTTCTCTACTAGAAGTCTGTTGAAACAAGAAGGCATGGCCACAGCAAAAGTGATAGTTG
 >CLocus_1836 TGCATGCTTGATTAGTTGGTTGATTGATCGGTGTTGGGGTTCTGTAACCTACGCTCTTGTGGATGATTATCAA
 >CLocus_18581 TGCATAAATTGGCATAACATGGAGGGAGAGAAGAGGGAGTGAGAGAAGACCCCTTATGTTGGCTAGTTTCATAG
 >CLocus_19001 TGCATAGTGAAGCCAAGATTGCAAAAGTGTGTTGTGATGTTTCTGAAAACCTGGACGTCATTGAGGGAAGT
 >CLocus_19463 TGCATGTTCTTGAAGAGCAGCACTTCTAAAAGGATGAACATGCAACAAAGCAACCAAAAAAATGGTG
 >CLocus_19749 TGCATTGTACATTGCAGCTTTGAATGTAATGGCACAAGATTTTCAAGCGAGGCTACGTTTTGAGTACAT
 >CLocus_20185 TGCATTTAGTTTAGCTGGATGTCATGTAGTGGGAGCATGGTATTGTTAGTTAGCGGAGTTACGAAGAGA
 >CLocus_20463 TGCATATATCTTTTTCGCACTACTTCTTTCAGGCTGTGAGTTTGCACAAACACGCAACAAGTGAT
 >CLocus_2057 TGCATTATTCATGAGTCCACCATTCATCCCGAGTCACATGAATCTGCTTACGACGGGTACACTGCAGC
 >CLocus_21178 TGCATAAGTAATTTATGTCATGTAGCATTTTGAAGACGACTGGTTGGCAGTTTGTAGGTTGGTTGTGAG
 >CLocus_21906 TGCATTTAGGACTTTAATCTACGTCATATGTGATTCTACTATGTTTGGCTAATACTAATGTGCA
 >CLocus_22086 TGCATAAAAACCGCTGTCTCTACCATGATCTGACCATGCACCTAGACAGAGCCTGCAGTTATGATCAT
 >CLocus_2239 TGCATATTGGCTACATTTTCAACCTGGCCAAAGGATCTCTGCGAGAAACCAATTATGAGATTTCTTGAAGTA
 >CLocus_22401 TGCATCTTTTCTTAGGCTATCACTGTGACGCCCCCGTCCCGTATCTTAAGTCAACTGGGAACGAGAAGGA
 >CLocus_22454 TGCATATCGGAAGGGATCTATCCCAAGGCTTGAGGCTAGCGATCGATCAGCTAAGCTTGATCCCA
 >CLocus_22722 TGCATATGTCAGATGGATGTCGCTTGGTAATCAATTGAATGGTGTGAGGCTCGTGGATGAAGGGCTCGGT
 >CLocus_22819 TGCATTTATAATGTGTAGGCACTAGAGAAATTTGTAACAATAATTGTGGCCAAGCTTGAGTATGAGAAA
 >CLocus_23071 TGCATATATGATGCTATGCAAAAGACCATTCAGATATCATCTAATCCGCACATGTGGACTGTTTTCTCTC
 >CLocus_23571 TGCATCTCATTGTGGGTCCCATTTGCAAGAAACCTGTGCTTTGACAGCAATTTGTTTCAGATGTCCGTATCA
 >CLocus_23639 TGCATACTAGCAGATACACAATTCATGGGGTTTACTTTGGTCTGTTTGGCATAGCCTGGTTCGATAAT
 >CLocus_24078 TGCATATGTAAGAATGCAGTCCAGACATGCAGTTTCTAACTTGAAGAACTAAGCTATTCTTTACAGAT
 >CLocus_24290 TGCATATGAATGAGGTCCTCAGACAATTGCAACTCATCAAGGATGCCTTTCTTCTGCAATTCATCATCC
 >CLocus_24567 TGCATTTGCCATAAAAAGTGTCTCGGAACATAAAACCATCTCTTCTGAGCAGTAGGATTGCAGTCAATCTGG
 >CLocus_24966 TGCATCTTCTTCCACAATAATCCAAAGGAAAGCTTCTAAGTCTTACTTGAATTTACGACGTGATCCCA
 >CLocus_25076 TGCATACCTTGACCTTAGGTGGTACATCCAGTTTCCAGATTTTCTACTATATGGTTTGTACACCTTTACT
 >CLocus_25197 TGCATCAATAACCTTTGACAGATCAATCTGGGAGGAAGAAGAATAACCTTCTCTGATCAAAACATTTGAAA
 >CLocus_25417 TGCATGCAATGCAAGACGTAATTTGCAAAAGCTGCAAAAGTACTTTCACTTCAACGCAACTCAATTTGGCTA
 >CLocus_2546 TGCATTTCCAGCAGATGACGGAACGAGTAGGAAGACATGGCAACATCATCGAAGCCTTGACGGTGGCAA
 >CLocus_25959 TGCATAAACAAGATACATCAATCTCCGATCAGTGCATCAGAAAGAGCAACAACCTTTTATTCTTAC
 >CLocus_26340 TGCATCTAACCTGCAGTAATATCCTAAGATTTTCCCATTTTGGTAAATGATCTTGGCAAATTCAGAT
 >CLocus_26837 TGCATTTCTGATCTGCAATGGGCCAAATTAAGTTAGATGGTCAAAATTTATGCTATGTGGCATCGAAAGATT
 >CLocus_27748 TGCATAAATCTGAGGCTATAACCTTGGTAAGCTCATAAGGACAAGCATTGCAATATATACTGGAAATGTT
 >CLocus_27868 TGCATTATCTGGGCGAAGCCAAACATCAGTGTGGCCATCCCAACTTCGAGTCAACCATCCAATCTGTT
 >CLocus_28020 TGCATAACTCAGGCTGATGCTGTTTCAATGCTGTTATCTGGGCTTGGAGTTCCATGGCTTGCTCTTAGT
 >CLocus_28141 TGCATGGCTTAGGCAGTATTTTCAAAGTTCATTGATGATCACAATGGTTGTTGGAGAGAAGGCTAGCAA
 >CLocus_28518 TGCATATGAAGACACACATCTTAGTCTTGAACCCATAAATCTGTTCTGCAATAGTCAAGACATGCTTTA
 >CLocus_28539 TGCATACGCAATGTGGATTCTTGTGGTAATTTGTGAATTTGGATTGGAAAAATTTCTTCTGTGCCATCTC
 >CLocus_29654 TGCATATGAATACACACATAATCAGAAAGGTAGCAAGCTTATTATCTATACAAGAGCTTAGTCTTA
 >CLocus_30111 TGCATACAAAATGCAGATGAAATCACAATAAGGGGCAAAATAGCAGAAATCATACAAGATCATTTTACAGA
 >CLocus_30193 TGCATAATTCAGCAAAAGACTCCACAAATTTGACATTCACAGCTTCTAGTGGACTCAGTGAGAGAGGCACT

>CLocus_30754 TGCATTTCACCAACCAAAAAACCACCATCAGCACTAACTTGCCAGACTTCATCTGCAACCATCTAAACA
 >CLocus_30811 TGCATGCACTTGGGTGGATGAATACAAGGAGCTTGGCGAGGATTGCACCATTTCAACGTGCCGAAATTTTC
 >CLocus_30895 TGCATCCTAGTTTGTGTAATGGCAATCATGGTCAATGCCGCCACCGTTGCCAATGATTTCGCCGACCATC
 >CLocus_30960 TGCATCTAGTTCGCTTTCTAAACTAGCTGTCTCTGAGAAAATGCTCCTGTCTAGCCTCAATTCACTCTG
 >CLocus_31190 TGCATAGGTGGAATTTTGAAAACTTTGAACTTTTGTACTTTTGGCCTTTTCCCTCTAGGTTTCCCTC
 >CLocus_31653 TGCATATCCAGTGATGGGAAGGAGTGGAGTTGTAATCTATCAAGTGTCAAGTAAATGTGGACTGTCTGGTG
 >CLocus_31977 TGCATACATTAGAAATTTTGGATACTTGTTCCTTCAAGGTAGAAAAGTGTCTTTCTGCAGCAGCTTTTGC
 >CLocus_32103 TGCATTGGTTGATAAAGTACAGAACACGGGATGGGCTGGGACCCGGTGAATCACACCATCACTGTCTACTGA
 >CLocus_32545 TGCATGTAAGGGTGTGTCTTGAATCCATGTATTGTGCTTGATCATCCATGTATTTCATCATTTACAGAT
 >CLocus_32800 TGCATGGATCTTTGAACTATATGAAGATGGATAACGCTTCTTGAAATGTTGACGACCAACACCAAGAA
 >CLocus_33000 TGCATTTGTGCATTGATAGGAGTGGTTGCAAAACAGGAACTGGGACCCGAGAATGTTGTGCGACTAGAAGA
 >CLocus_33778 TGCATTATCACCCCTGAGCACGCGCTAGCATGATTTTCTACCTTATTACTTCGGAATAGGATGACCCGCTA
 >CLocus_34530 TGCATGATCTTGTGATTGTATGCTTTTGGATGAATTTTCTGAAATGTTCACTTTTTTGGTAAATGTGCC
 >CLocus_34965 TGCATTCAACACACAATGTTTCCACATCTTTTCTCATGTGACGGCCAAATAGAAATGCTCACTCAAAACATC
 >CLocus_35380 TGCATTTGGCCATAACCCCTGCCATCCCTCACATTCAATCTAGATACCCCTTGGCCATTGCACTACAAATAA
 >CLocus_35828 TGCATGTACATGTGCAATATCTGACGTTTCAAGTTTATTATGTGATTTTTGTGACGTACACAGCAGAGAAAG
 >CLocus_36950 TGCATGCTTCACTTTCTTCACTGCCAAAGGAGGACCTAGCCCCAGACCATGTATTTTTGTGCTGCCGT
 >CLocus_38082 TGCATGGTGGAGAGACGGGAAGATGTGAGAGAGAGTTGGAAGTGAATTCACGAAATAAAAGGTTTGTGA
 >CLocus_38181 TGCATCAAGAAGTTGAGGAACAAAAGCATACCAATGGTTATGGTGAATGAGGTTATCATGGTACTAAAG
 >CLocus_39785 TGCATCAATCGCTAGTATGGACTTTGGCGTAGTGTGAGTGACCAAGACATCATAAATCAATGCGCTTCAAG
 >CLocus_41168 TGCATTTATCTTTGGCCAGTACTACTGACTTTTGTATGTGCATGACGAACACTGGCTGCGGTGGGACAGTAC
 >CLocus_41188 TGCATAAAATATTATCAAGCGTTTATGAGATGTTACGTTATCAGCATAATAGCTCCTACAAGACATCC
 >CLocus_41238 TGCATCTCTGTCTTCTGCTGCCAATATGTTTCTGTGTCAGCGTACGAATTGCTGGAGGAAACATTTTCTT
 >CLocus_42272 TGCATGCCAGAACAAAGATCGTTTTTGGCCAAATGGTGAATTTGTATACATAGCTCTGGCTAGTTTGA
 >CLocus_43192 TGCATAAAGATGTCAACGCAAAATGATGCCCTGCTTACAGGGACAGGGTGGGGGAGGTTGAGATGGTGA
 >CLocus_43766 TGCATAGGAGTATGCAATCTTATGTCTATTTTGAACAACATGGCCATCTATAACGTAGCTCAACAAAT
 >CLocus_4387 TGCATATTGTGCAACAAATTTATTTTCAACAAAGGATGTGCCGAAACTATGTTTCCATTCTTATCACA
 >CLocus_43925 TGCATAGGAATATTGTTGGTTTGTACAACAGACCTCGCAAGCAAGGGCCATTTTCCATTACTGAAACTT
 >CLocus_45156 TGCATCATGCTGTATCCATATGAGCTTTCATCTCCACACCCCTCAACTGTGGATCTGGACGAATATGAAT
 >CLocus_45720 TGCATGATTTCATGTATGAACATTAGGGCAATACATGGACATTTCTTTCGACGAAATCCTTTTTGCGCAACC
 >CLocus_45893 TGCATTTAGGCCATAGGAATCGGCCCTGGCCCTATAAGAACCTTCCAGACTGTTCTAGCAGCAGCTGGC
 >CLocus_4638 TGCATGTGTGAATGTGATTGTGTGCATGATAATGATGCAAGCATCTACTCAAAGGAGATTTTTTATGATTG
 >CLocus_48278 TGCATCCGAGGTGAAGAAGTTTTCATGGATTAGCTAGTTTCTATAGGAGGTTTTCCTCACTTTAGTAC
 >CLocus_49075 TGCATTGTAGAGCGCCCAATGTTAGAGTCTGTTCTATGGGATTTCTGGACTCTGCTTCCAAACTTCTTT
 >CLocus_50085 TGCATGTTTGAACATACTTGTGACGTTTCTCACTTATGTCAGGTTGAAACTTCTAAGAAACTTCAAGCAGA
 >CLocus_50822 TGCATCAATTGGGTGTCTAGAAATGGGATCACTAGTTTATAGATGCGAGCGGACTTGGATCACTTCAAT
 >CLocus_51188 TGCATTTGTAACTTTTGGTACCAATTTCTGGAGGCTTCGCGCAATCCATACAAGGATTTTTTACAGAT
 >CLocus_5195 TGCATAAGAAGTTGTTTGTGGGTCCGGATCGTCTGTGTTGTTGCTATTGCTATTTCTATTACAGAT
 >CLocus_52126 TGCATCTTGCACAGACTCAATGTCAACAACTCATCAAAATTTGGCATCTTGGGACAAACATCTCTATAGCA
 >CLocus_538 TGCATTCCATTCAATTGAGTTTGCATAGAATAGGGTTGTTCACTTCACTATTGGTCACTTCCATTTGAA
 >CLocus_5548 TGCATAGGAATAAATGGCAAGTAGAACACGATTTCTAAGTAAATGAAGCTACACTACCGATATGTTAC
 >CLocus_5620 TGCATCTTTGGCACCTTTGCTGCTTCTGTTGTTTTCGAAGAGTCTTAGGCTTTTTATCTTTTCGACCAT
 >CLocus_61178 TGCATGAATTTGTAGTGTGTTGGATTGTGAATTTCTGTTGAGGGGATTTTCTCTGAGTATGTTGTTGATCA
 >CLocus_6183 TGCATAATCCAGCTCGACCTGTGCTGCAAGGTTCCATCTAATGGGTCAAAATTTATGAGTTATTACAG
 >CLocus_61836 TGCATCTGCTCATTTCTATTGAGGATCCAATAGCATTATCAATTTTGGTGGTGTGAAGGCCCTGTGGT
 >CLocus_6412 TGCATATGGTCAATGTTCAAGAGAAAAATGCACCGGTAAAAGAATAAATTCAAATTCATTCTCTATTTC
 >CLocus_6528 TGCATACATTTTTCGGGTGGAACCTTATATAGCAGAATTGAGGCTCTGTCCAAAAAGAAAGGTCAAGGCC
 >CLocus_691 TGCATTTGAGTTGTATCAAATGCTAAGTAGTCATTCAACAATCGCCACATAAGAGTTCGAAAAATTTACAGA
 >CLocus_69100 TGCATCTAGTGGATTAGGTTCCAGAGGTGTGACTGAACTTATTGTTTACATGGTCCCACTATTACAGA
 >CLocus_7092 TGCATCTTATTTACTGCGACCTAAGAGGGAGATAACACAAGTAGGAGTGAAGAGGAAAGGACGGAGAAAT
 >CLocus_759 TGCATGTAATCAAACTCAAAATCTTTTATTCAGCAACCTATCCCTTCTCAGCATAGGTTTTTACAGAT
 >CLocus_7839 TGCATCTCAAGGAATGTCCCTTTGGCTCATTCCATCAATTAGCAGACGTTTGTATGTACAACTAGCTTTCT
 >CLocus_789 TGCATGCAAGGATGAAGAGGACACTACAAGGTCTGAATGTGATATCGGAAAGGCAAGTATCTGCCAACT
 >CLocus_7922 TGCATGAAATGATGATTCAAATCTAGAAAATGTCACACCATCATCTTATGGATCCAATGGAGTTTCAAC
 >CLocus_7967 TGCATCTTCACTCTGATAAAGAGTGAAGAATGATTGAACACTAGTTTACAAAAGCTTTGAATCATGTG
 >CLocus_820 TGCATAAGATCTTGAATGAGTTTTATGAACCTTATGGACGGTTCGAATAACCAAAATGATCATGTATTGAT
 >CLocus_82158 TGCATCAAGATCTGCTGACGCTGACGTTCAATTTCACTTCTACTGTTTACGTTTACGTTTCTGCTGCT
 >CLocus_857 TGCATAACAAAGGACAGCACTCTGCGCTCCACATTTGTAAGCTAAGACGCTGCTGCTGCTGCTGCTGCT
 >CLocus_92219 TGCATCGCTCACTTGTAAATGGCACAATTTGCTACTCTTACTTGTGGAATCATGTGAGGACTCTGCATA
 >CLocus_9326 TGCATATGGCAGTCAAGTACCAATGAGTGATCTACGTTCAATCAACACAAATGTAGGATGAGAGAT
 >CLocus_820 TGCATAAGATCTTGAATGAGTTTTATGAACCTTATGGACGGTTCGAATAACCAAAATGATCATGTATTGAT
 >CLocus_9196 TGCATAGAACGCAAGCAGCTTCTGTTTCTTCTGTTCAACCTTAGCATCTCAAGTAGTCAACCTTACAGAT
 >CLocus_9404 TGCATAGCAGTCAAAAACAGCCTCTCTGCTGCTCGAACCTCCAATAGCTATAAAAAACAGGGTAAGCA
 >CLocus_19861 TGCATCCATTTATTGTGCAAAAGATTTTGGACACGATGATGTGATGATTACATCATTTGCAAAAAGTTTA
 >CLocus_29372 TGCATCTCTGTTCTTCTGCTCTCTGCTCTCTGCTCTCTGTTGTTTCTGACAAAGCCTACATGGCCGAACCA
 >CLocus_35687 TGCATATATGCAATATGGAATTCGTGAGGTTGAAGAGCTCATATAGGAATTAGGAAGTTGTTGCTGGCTT
 >CLocus_45169 TGCATCAAGAATTACGTATAATGTTGTTGTCAGACCGGAGGGTCAAGATAAACAGGAGCCAATGAGAATCA
 >CLocus_45410 TGCATCAAAACATATACATCCAGCAGCTTCTCACAGCAAGCCTAGCAGCTTGTCTGAGGAGTCAAGCTG
 >CLocus_47572 TGCATTTCAATCCGACTGAAAGAGCCAACAGATTGATTGTGCAAAATGATGATGAACAGCACACAAGATGAC
 >CLocus_61998 TGCATCATGATGTTAGTGAAGTGAATAATGTTGGCTGCTTCTCATGTGTCACATTTCTTTTCTTCTATTT
 >CLocus_65834 TGCATCATTTCCATATCTGTTGAAATCAATCATGGTAAGCCCTAGTCTGTTGCTGATTCCATCCGCTA
 >CLocus_85475 TGCATCCAAATAATGTTCACTAGCTGCAATTCACGTTCTCCATAAAAGATCTCTTGTGATAGCCTTACCC
 >CLocus_88285 TGCATCACTGGCAGTAAAGTCTGCCAAAAGCCTTTTGGTAAGGTTCAAGGGGTTACTTGTGCTGTTGAG
 >CLocus_113237 TGCATCAAGTGTGTTACTATGTGGCCCAAACTTGTACTTTTACATGGATTTTGTCCCTATCACAG
 >CLocus_48645 TGCATCAAGACTAGTTAGGCCAATTTTGGTCAAGGAGGGCCATTTTGCAGCTTTTGGTGAAACTACAAT
 >CLocus_93457 TGCATCTGACATTTCTATAGCGTTAGCATATAGGCAATTAGCCATGCGCATTTTCTTCTTCCATCTTAG
 >CLocus_50447 TGCATCTTTCTATGAGGTTCTTGTAGCTAAAGCATTTTCTATTTGTTCTTTTATTGTGGAGTACACCAT
 >CLocus_14749 TGCATATTGTTGATAGCAAAAGGCAATCTCCGATCAGCAGTGTATCAACCTGAGATTTTCTTAGTTG
 >CLocus_23792 TGCATTACAGGGCAGATTCTTTCATGTTTTTGTGTAATCACATAACCTTCTCCAGCGTATTTTGTGTGA
 >CLocus_35624 TGCATCTCTCTCCACAACCAAGTCAATCAACCAACCAACCCAGAACCGTCTTAAATTCCTTTCTTTCT
 >CLocus_47727 TGCATGATGTTGTCACAAATAGGAAGAAAGCAAGCTGGTGAATTTGTTGAACAATGGTATCTAGCCACACG
 >CLocus_6040 TGCATTTAGGCCAAGAGAGGGCTTACTCTTACGATAGTACTTCTGCTTCTCAGAGCCGTCCATCGTAA
 >CLocus_9954 TGCATTTCCGTGCTGAGAACCACTAGTGTGCGTGAATGCGAGAAGCTTGAAGCGGTCCTTCTAGTTTA
 >CLocus_12952 TGCATTTGTTCTATGGCACATTACAGGAGGAAGTTTATATGGGGCAACTGCCTGGCTTTGTTGCTCTGAGA
 >CLocus_25124 TGCATAGCTTTTCAAGTTTCAAGTCAAGATTTTGGTAACCTTGCAGAGTATCCATATCTTCCCATGGT
 >CLocus_28439 TGCATACAAGTCTCATAGTGTCTTGTCTTGCAGCAACAGAAGATTCAACATACACATGGTTACAGAT
 >CLocus_32851 TGCATAACCAGTTGTAATAATGTTCTCTAGTGTACTATTGCCAAAAAGAACAGATAGTGGGCTGGAAG
 >CLocus_7761 TGCATAATATGCTCATAGTAACAATCAAGGACCTTCAAGGCACCATTTGGAGAGCCAGAGCATAAAGCT
 >CLocus_12563 TGCATAAAAATGACTTGGGAATAAGATTGATACAGACTTGGGTGGCTAATGTGCAATATTTTCAATACAGA
 >CLocus_14852 TGCATTGTACCATTTACATTGAAACGGCTATTAGAATGTCAAAAAAATGAAAAAATCTTCTATAAGAT
 >CLocus_35730 TGCATGGTGGGAAATGGGAACACCAAGCTGTCTATTGTTGGGAGGAGGGGAGAGAGTGAACAAACAT
 >CLocus_73469 TGCATCATCTCTTGAATGGGGAGTCAAGTGAAGTATTTGATGGGAGGCCATTTGGGAGTAATACAGAT
 >CLocus_3674 TGCATATAGAAGTTTATCTTTTATAGGAACAATAGCTTTCATGTAGAATGAATGCTAACATACCCAA
 >CLocus_13789 TGCATAGTGGCTTGTCTATGTAATGACGAGGTCAAGCCAAATGTCAAGTAATCAACTATCCCATGA
 >CLocus_22343 TGCATAAGTATAGGGATCTTACGACACATGAGCTAAAATGCAAAATTAATCTTGTGCTGACGAGATTA
 >CLocus_11066 TGCATCTGATTTTGTACTCAAAAGGTCAAGCTAAGGTGCGTGAATGTGAACCTCAATAAGGCAAAATGTC
 >CLocus_4882 TGCATGCAACCTGGGATTGATGATTTGTAATGATGTGATATCTACTGTTTGGCTTTGATTACA

>CLocus_10938 TGCATTTCATTGTGCAAAATAGTGGACTATTTATTTGTGAAAAAAGAAAAAACATGCAGGCATGGA
 >CLocus_21601 TGCATCCATTTCGGAGCAACTAAAGTTTTACCCCAAAGACATTATGGCCCTGTTCTTCACTGCTTACAGA
 >CLocus_23758 TGCATGTGCATCTTCACACTATTTGAAACCATGGAAAAAATGCTCGTCACACAAGATTTTCGGGTAGA
 >CLocus_24089 TGCATACGCAATCAGTTTCGGCTCTATGTGCTTACCAAAATAATTGCAATAGATCCGCTAGGAAAAGAAA
 >CLocus_37178 TGCATTCTTCATGTCCATTTCAGAGATTTCCCTAATTTACCAGTAGCTAATCAATATTACACATACCATA
 >CLocus_42308 TGCATGGGAGGATATCGAAGACTGTAACAGAGCGACATAAGGAAGTAAATGCCTGGGACATTCACTTTTT
 >CLocus_42414 TGCATGTATCAGTAAATACAGGCTCCAGGTGGAGACGCCGAGTAACATCTCAGAGACAAAGCCCAATT
 >CLocus_44668 TGCATGTAAGCCGAGTGTAGGAATGCCAGCAGGATGTGTCTGACCTGGACTTGGGGAGAAGCTTCCGT
 >CLocus_75386 TGCATGCACAAATTAGCTGAGATTGGATTTCACCCGCGTCTCTCTGCAAAAAACAAGGATCTTACAGA
 >CLocus_6935 TGCATTATACATGCTTCAAAACACCTTGTCTTGTCTTGGGAAAAATTCAGTACTCAATATTGGCTTACAGAT
 >CLocus_14438 TGCATAATCATATTTTTCAACAAGTTTTGGAAAGGTAGAACTCATCTTGTAGCAAGCTCACAGCAAAACA
 >CLocus_24646 TGCATCTCGATCCCATACAGGCGTGGCCAGTGGCGGAAC TAGGAGGGCGGGTGTGGGCACGCCCTAAGGT
 >CLocus_27086 TGCATGGCAATATAGCAAAAGCAGTATACCCCACTGCTGTTTTATCAGAGTTCCTGCCGGGAGAGTAG
 >CLocus_43912 TGCATATCGTGATGAAAGAGCCCTCAGTATTACAGCTGCTAAATAGGCCTAAAAATTAATTTTGGATTGT
 >CLocus_115971 TGCATAAGGTGTAATACATGGAGAAATTCAGGCAATGGTTGCTCAACAAATTCACAACTTTTTGGAACC
 >CLocus_2885 TGCATGAGCGAAGTATACAAGGTGATCGTATAGTCTACCGCAGGATGGGAACAAGGAAAAGACTTCTGC
 >CLocus_6072 TGCATTCAAGATTGGAGCAGCGCAATTTGGCTTGGACATTCTTGTGTGCAGGGCAGGAATGCAGTGCAT
 >CLocus_11029 TGCATTTTATGTGGTTCAATTCGATCTTGGAAAGCTTTTTCCAAATCATTAGTAGTTCCTTCAATACAGGCAAC
 >CLocus_16423 TGCATGTGCCCTGTGTCTCAGTGTCTATTTCTAAGGCAAGTAGACTTCTTTCGAATGCGGATAAGTTGA
 >CLocus_17390 TGCATGACTTATGAAATGATTGGCTAAGTTTGTGTCTGTGTAATATGTTTACAGGCAAGAAGATCTTCA
 >CLocus_18795 TGCATGCTAGTCCATAGTGGCGGAGACTCAGGCTAAATCATCCGAGGGCATCTGACGCAGATTAGG
 >CLocus_18908 TGCATAGATTATCAATTTGCGTAGGGATCCGTTTGTGGCAATGAATGTTCTTGTGAGTAATCAACACCA
 >CLocus_20596 TGCATGAGAAATTGACCCATTCTGTTTACAGGCAATTTAGTTTCCAGTCTATGGCATATCTTGGCTTGAT
 >CLocus_22046 TGCATGAGGAGCAATTCGCCCAAAGCCGAGGAAAAAATCCCACTGTTGATTATGCTGCCGCGTGGATAT
 >CLocus_35255 TGCATGACAAATTACAAGCCTAATTTCTTCCCTGTGCTAATATAGTATCTCTAAGGATGACAATTG
 >CLocus_37143 TGCATATGCACGAGGTGGGCAAGAGTCAAAGTTGCCACAACTACTGAAGGAGATGGCTGCACCTTACAGAT
 >CLocus_44171 TGCATAAAATCAAGGTTTAGACATGACGTCTTCTTCAACGAAGTATGAGGTGGAGAAATTTGATGGTTCGA
 >CLocus_46214 TGCATGTCTCGTGGCGTGGGGTCTTTGACGCAATGAATCGGAAAGTTCTTCTGGGTAATCCGGCTGGCC
 >CLocus_2546 TGCATTTCCAGCAGATGACGGAAACGAGTAGGAAGACATGGCAACATCATCGAAGCCTTGACGGTGGCAA
 >CLocus_8722 TGCATGCTAATCTTCTTCTTTGGTTTGGAAAAAACATGGTAAATATCTTCTTGTGAAAGTGAAT
 >CLocus_13484 TGCATAAACTCATGAAAGAGTTCAAATCATTGTACAAAATTTGTGCAATAAATAAACAACACATGGCG
 >CLocus_16223 TGCATGCTCATATAGCGATGATGAATTCAGCTGGAATGAAAAACCGGAATATTGCTCTTCTGTCATTGTT
 >CLocus_32349 TGCATTGCCGAATTTGCAATTCTCAGTCTCAGTCTGATGCTCACAACGCGTATTATGGACGTAGCTTTGGCC
 >CLocus_165 TGCATGATGGGGCCACGTCTAAAATCCTACTCTATATGAATGGTCTTGGGCTAAGAAAAAGCAATTTTTG
 >CLocus_2281 TGCATTATAATGCAAAGGTTGGAGAGAAACAGACCATAAATTTGAGGTGATTGTCATAGGCAAAATACAG
 >CLocus_8324 TGCATACCTTATTGAAATCCATCACTTACACATTTCTTGGCAATCACAGAATTCAGTAATCAAGGAATT
 >CLocus_8703 TGCATTTCTTGGGATGAAGACATGAATACGAAATAAGAGTTTAGTGCATAACAAAAGATGTCATTACAG
 >CLocus_11306 TGCATACGGAATTTTGGCCGAAATTTGGAACAGCTCGGAAGTCTTTTCCGCCCTCTCGATCGACCTTCCCT
 >CLocus_13484 TGCATAAACTCATGAAAGAGTTCAAATCATTGTACAAAATTTGTGCAATAAATAAACAACACATGGCG
 >CLocus_13742 TGCATCTGGCATGGACATAAGTTTCTCTATTGTCAAAGACATTCTTCCGCAACATCTCAATCAAGACAC
 >CLocus_16291 TGCATGCTGCTCTGGTTGGATCTCTTACGGGAAGTTCCGACTTGGCCCGGGCGTGGCTTCCGTCATATA
 >CLocus_20470 TGCATATATGAACCATGAGCAGTTGAGATTCTTGCACCTTGGCCCTTACTCTATGAGGTTCACTAGGCTTG
 >CLocus_23311 TGCATCTTCTAGTCAACAAGACGAACAAAGTAGAAATAGAACAAACCAAGAACAGAACCGAGCAATCAAG
 >CLocus_26986 TGCATAAATTAGGAACAATGTACGGGAAGAAAAAAGTGGTCAAAGTTGGGCTACATAATCATTTTAGTCCAC
 >CLocus_29699 TGCATATAAAGGGGTTTCAAACACGGGCTGCCCTTTGCTAATGTAAGTGGTAGTCCGACGTGCAAAGCC
 >CLocus_31949 TGCATGCAAAAAATTTGAATGCAGGTGATATTGGCCCTTGGAAATCTTGGCCACCAAGAAATCAGGATTA
 >CLocus_39876 TGCATTCTTCTTCTTCTTGGGACTCATTTTCCAACAGATGCCACACAGGTTCAATTTTCAACTTCAGCCAT
 >CLocus_43511 TGCATCTTCCATTGCACTCCTCTACACATGTGAATATCATAACTTTGTCTCCAATTAGCTACCGTGAT
 >CLocus_47265 TGCATGATATGTTTGTCTTGGAGAGAAGTCATCTTCTTGGATCCCTAATGGTTGCGATATTCTTTACAG
 >CLocus_49852 TGCATTCCCTTGTGGTGGAGCACACTCATTTACCCTAACTCATAGTCTAGAATTCAACACTACATAAG
 >CLocus_51784 TGCATTACTTGTGAATACACATGATGTCTTCAAGTATTGAGGTGCCAGTGTGGAGATGCCTAACTTTTT
 >CLocus_62012 TGCATCTAAACCTAGAGCATGTAGAGCTAGGAGGTGTGCAGGAACGTTTCCCTTGCCTCCTAATAAAGGAA
 >CLocus_598 TGCATTATATCAACGACCAAGATTACAATGTACATAAACCTGAAGCAAATTAGCGAACGATTATGCTCT
 >CLocus_1582 TGCATATCTCTTTCTCAGAGACACTTTTGGCATTCGAAGTGTGTGCTGCCATGCTTGGTTTATCCAT
 >CLocus_3226 TGCATTTTGTAGTAAATGTTGCAAGATTCCATCAACAATGGTATTGAAACCAGAATTCGGAAATCATCC
 >CLocus_5874 TGCATTACAAGAAGTAACACGAGGAAAGAAATGAAATCTCACGTTGGAGTAAGCAGCAGATCTGAATTTCT
 >CLocus_8237 TGCATCTTACGCCAAGCAGCTGTCTTTGTTTGGCATCGTAATCACATGCTTGTCTTCTACTCATTT
 >CLocus_9174 TGCATACGTGATTGTGCTACCTTCTAATCGCATTGCCCAATATTGCATCAGAATAACACAAATCAAAT
 >CLocus_11410 TGCATGAGAAAGAGATGATATCAGTCCCTTCGACTCGGTTATAAGACTAGCTACAATAAGCCCGAAAGG
 >CLocus_11563 TGCATGCCTGCCTGTACGCACAAGATATTGACCTATGTTCTGCTGTCAGCTGTCTACCTGTTCTGGGTTA
 >CLocus_15984 TGCATGGGTACATGAATTTGATTCTCCATTTCTCCCACTGACAAGGAGGGGCAACTAATGACAAACGGCA
 >CLocus_28906 TGCATAAAACCCGAATAAGAAAGCCGCTGCTAGCCTTGTAATAAGGTGCTCGCCGACTAACATCTGGAGAT
 >CLocus_40667 TGCATAAAGGGCATGACACAGCGCACGAATATGTCTACCTTCTTAGGCACAAATCAAGCTCCTTTTCT
 >CLocus_46501 TGCATGTATCACCATTCTTGACCAAGTCTTATCTATCTCATCTTCTCTGAGCCATTGAAACTGCCATTT
 >CLocus_48425 TGCATTATGGGAACCTCATAGATACCGTGAAACATCAAGTCTGCATAATAAGGCACGTAGATGGCTGCC
 >CLocus_73710 TGCATCCATCCAAATCAACATTGGGCATTTGGATGCAAACGGAGGTGTAATGGACAGTTGACAACCTTT
 >CLocus_81476 TGCATGAAAGAACATCCATCACACTAAAGCATAAAGTGACACAATATTGAGCTCACCTTCTTACAGAT
 >CLocus_96965 TGCATCCACACTCTATCGAAAAATTTTACCTTATATATCTCAGGTGCAACATATACACCTGCAATTTGCA