

Universidade Federal do Amazonas
Programa de Pós-Graduação em Biotecnologia

ILDENEI ALMEIDA DOS SANTOS

**ESTUDO GENÔMICO DE POPULAÇÕES DE PIQUIÁ (*Caryocar villosum* (Aubl.) Pers.,
Caryocaraceae): elucidando o processo de domesticação de populações da espécie**

Manaus/AM

2026

Ildenei Almeida dos Santos

**ESTUDO GENÔMICO DE POPULAÇÕES DE PIQUIÁ (*Caryocar villosum* (Aubl.)
Pers., Caryocaraceae): elucidando o processo de domesticação de populações da espécie**

Dissertação apresentada ao Programa de Pós-graduação
em Biotecnologia, para obtenção do título de mestre em
Biotecnologia, da Universidade Federal do Amazonas.
Área de concentração: Biotecnologia Agroflorestal.

ORIENTADORA:

DRA. DORIANE BATISTA PICANÇO

COORIENTADORA:

DRA. RUBANA PALHARES ALVES

MANAUS/AM

2026

Ficha Catalográfica

Elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

-
- S237e Santos, Ildenei Almeida dos
Estudo genômico de populações de piquiá (*Caryocar villosum* (Aubl.) Pers., Caryocaraceae): elucidando o processo de domesticação de populações da espécie / Ildenei Almeida dos Santos. - 2026.
47 f. : il., color. ; 31 cm.
- Orientador(a): Doriane Batista Picanço.
Coorientador(a): Rubana Palhares Alves.
Dissertação (mestrado) - Universidade Federal do Amazonas, Programa de Pós-Graduação em Biotecnologia, Manaus, 2026.
1. Genômica. 2. SNPs. 3. Diversidade genética. 4. Estrutura genética. I. Picanço, Doriane Batista. II. Alves, Rubana Palhares. III. Universidade Federal do Amazonas. Programa de Pós-Graduação em Biotecnologia. IV. Título
-

Ildenei Almeida dos Santos

**ESTUDO GENÔMICO DE POPULAÇÕES DE PIQUIÁ (*Caryocar villosum* (Aubl.)
Pers., Caryocaraceae): elucidando o processo de domesticação de populações da espécie**

Dissertação apresentada ao Programa de Pós-graduação
em Biotecnologia, para obtenção do título de mestre em
Biotecnologia, da Universidade Federal do Amazonas.
Área de concentração: Biotecnologia Agroflorestal.

Aprovado em vinte e quatro de abril de dois mil e vinte e seis.

BANCA EXAMINADORA

Prof^ª. Dra. Doriane Batista Picanço, presidente
Universidade Federal do Amazonas

Prof^ª. Dra. Ana Flávia Francisconi, membro
Universidade de São Paulo

Prof^ª. Dra. Izeni Pires Farias, membro
Universidade Federal do Amazonas

DEDICO,

Aos meus pais Ivo e Luzineide pelo incentivo e apoio para alcançar meus objetivos. A minha esposa Livia pelo encorajamento. Aos meus irmãos, irmãs e amigos (as).

AGRADECIMENTOS

A CAPES pela concessão de bolsa de estudo durante todo o mestrado, apoio fundamental para conclusão deste trabalho.

A Dra. Doriane Batista Picanço, minha orientadora. Por cada palavra e “puxão de orelha”. Foi uma honra ter sido seu aluno. Pela oportunidade de ganhar experiência em laboratório, executar as pesquisas laboratoriais no LEA (Laboratório de Evolução Aplicada) me deu a oportunidade de conhecer pessoas excepcionais que contribuíram para que pudesse concluir minha dissertação.

A Dra. Rubana Palhares Alves, minha coorientadora. Por cada revisão e conversa, o sentimento de gratidão por tudo que fez por mim, obrigado.

Ao Mestre Érico Polo Técnico do departamento de Genética (DGEN), expert em Genética computacional. Sua ajuda foi primordial para conclusão do meu mestrado, sou expressamente grato por todo o apoio e conhecimento que pude obter ao seu lado.

A Dra. Izeni Farias, por ceder equipamento e recursos humanos do LEGAL, para construção da biblioteca genômica.

A Dra. Sandra, por me ensinar as metodologias de se construir a biblioteca genômica para este estudo.

A Universidade Federal do Amazonas por meio do programa de pós-graduação em Biotecnologia pela oportunidade de realizar o curso de mestrado.

Ao departamento de genética, pela estrutura fornecida para iniciar e finalizar minha dissertação.

Aos colegas do LEA, que tive a honra compartilhar bancada. Por toda ajuda e companheirismo.

Aos meus pais Ivo e Luzineide, que sempre me incentivaram emocionalmente e financeiramente em cada etapa da minha vida profissional.

A minha esposa Livia, que desde o início me deu apoio em cada etapa. Sou extremamente grato por ter você em minha vida.

A Deus, por me abençoar e cuidar da minha vida.

RESUMO

ESTUDO GENÔMICO DE POPULAÇÕES DE PIQUIÁ (*Caryocar villosum* (Aubl.)

Pers., Caryocaraceae): elucidando o processo de domesticação de populações da espécie

O piquiá (*Caryocar villosum*) é uma árvore amazônica com diferentes usos, tais como alimentação, uso medicinal e madeireiro. A espécie apresenta populações incipientemente domesticadas, resultantes de um processo coevolutivo que provavelmente se iniciou no final do Pleistoceno e início do Holoceno. Atualmente, povos e comunidades tradicionais selecionam e cultivam piquiazeiros, sendo esperadas diferenças genéticas entre árvores de sistemas de cultivo recentes, como quintais, e árvores de florestas maduras. Assim, o objetivo deste estudo foi analisar a diversidade e a estrutura genética de piquiazeiros em quatro áreas, comparando quintais e florestas maduras, a fim de compreender o processo de domesticação da espécie. As áreas estudadas foram: Floresta Nacional (FLONA) Caxiuanã, FLONA Tapajós, FLONA Tefé e Reserva de Desenvolvimento Sustentável (RDS) Amanã. Foram utilizadas 207 amostras de tecido foliar de piquiazeiros, e foi realizada genotipagem por sequenciamento envolvendo dupla digestão de DNA (ddGBS). Para as análises, foram considerados 6.900 polimorfismos de nucleotídeo único (SNPs). Para caracterizar a diversidade genética, foram calculados parâmetros como riqueza alélica (RA) e número de alelos privados rarefeitos (APR). A estrutura genética foi avaliada com base em duas abordagens: sNMF e fineRADstructure. A taxa de imigração foi estimada utilizando o programa MIGRATE. Foi encontrada maior variação genômica entre as áreas (8,92%) do que dentro das áreas — ao comparar quintais e florestas (2,14%) —, o que sugere a existência de quatro populações genéticas. Considerando cada área, a maior variação entre quintais e florestas foi observada na FLONA Caxiuanã (6,30%) e a menor na RDS Amanã (-0,4), indicando que, nesta última, não há diferenciação genética entre árvores cultivadas e de floresta. Os marcadores SNPs mostraram alta similaridade entre os sistemas. Quanto à diversidade genética, a FLONA Tapajós apresentou os maiores valores (RA = 8.117; APR = 505), seguida da FLONA Caxiuanã (RA = 8.066; APR = 438), FLONA Tefé (RA = 8.007; APR = 352) e, por fim, RDS Amanã (RA = 7.984; APR = 255). Considerando todas as áreas, as florestas apresentaram maior diversidade (RA = 12.333; APR = 442) do que os quintais (RA = 12.310; APR = 435). No entanto, quando analisadas separadamente, apenas na FLONA Tapajós os parâmetros foram maiores na floresta, o que pode sugerir seleção nos quintais, resultando em redução da variabilidade genética. Nas análises de estrutura genética, o sNMF indicou que quatro é o número mais provável de grupos genéticos. Houve estruturação entre pelo menos três áreas, com a FLONA Tefé e a RDS Amanã apresentando alta similaridade, podendo ser consideradas extensões uma da outra. Já o fineRADstructure indicou a existência de dois grupos: um formado pela FLONA Caxiuanã e FLONA Tapajós, e outro pela FLONA Tefé e RDS Amanã. A FLONA Tapajós apresentou a maior taxa média de imigrantes (0,98). Os resultados indicam que, nas diferentes áreas, houve interações distintas entre humanos e o piquiá. Nesse contexto, a FLONA Tapajós se destaca por apresentar maior diversidade genética e maior taxa de imigração, além de menor diversidade genética nos quintais, o que sugere que a área pode ter sido um centro de origem da domesticação da espécie, ainda em processo incipiente.

Palavras-chave: Genômica; SNPs; Diversidade genética; Estrutura genética.

ABSTRACT

GENOMIC STUDY OF POPULATIONS OF PIQUIÁ (*Caryocar villosum*): elucidating the species populations of domestication process

The piquiá (*Caryocar villosum*) is na Amazonian tree with diverse applications, including food, medicinal, and timber uses. The species exhibits incipiently domesticated populations, a co-evolutionary process that likely began between the late Pleistocene and early Holocene. Currently, traditional peoples and communities select and cultivate piquiá trees, 8ncipie to expected genetic 8ncipiente between trees in 8ncipie cultivation systems, such as home gardens, and those in mature forests. Thus, the objective of this study was to analyze the genetic diversity and structure of piquiá trees across four 8ncip, comparing home gardens and mature forests to understand their domestication process. The study 8ncip included the Caxiuanã National Forest (FLONA), Tapajós FLONA, Tefé FLONA, and the Amanã Sustainable Development Reserve (RDS). A total of 207 leaf tissue samples were used, and genotyping-by-sequencing was performed using double-digest DNA (ddGBS). For the analyses, 6,900 single nucleotide polymorphisms (SNPs) were considered. To characterize genetic diversity, parameters such as allelic richness (AR) and the number of rarefied private alleles (RPA) were calculated. Genetic structure was evaluated based on two approaches: sNMF and fineRADstructure. The migration rate was calculated using the MIGRATE software. The results showed 8ncipi genomic variation between 8ncip (8.92%) than within 8ncip 8nci comparing home gardens and forests (2.14%), suggesting the existence of four genetic populations. Within each population, the 8ncipien variation between home gardens and forests was found in the Caxiuanã FLONA (6.30%), while the lowest was in the Amanã RDS (-0.4%), indicating no genetic differentiation between cultivated and forest trees in the RDS. Regarding genetic diversity, the Tapajós FLONA showed the highest diversity (AR: 8.117; RPA: 505), followed by the Caxiuanã FLONA (AR: 8.066; RPA: 438), Tefé FLONA (AR: 8.007; RPA: 352), and finally the Amanã RDS (AR: 7.984; RPA: 255). Considering all 8ncip, forests exhibited 8ncipi diversity (AR: 12.333; RPA: 442) than home gardens (AR: 12.310; RPA: 435). However, 8nci analyzed separately, parameters were 8ncipi in the forest 8nci in the Tapajós FLONA, suggesting that selection in home gardens resulted in a reduction of genetic variability. In the genetic structure analyses, sNMF indicated that four is the likely number of genetic clusters. Structure was observed between at least three 8ncip, with the Tefé FLONA and the Amanã RDS being very similar, potentially representing extensions 8ncipi another. The fineRADstructure revealed the existence of two main groups: one formed by the Caxiuanã FLONA and Tapajós FLONA, and the other by the Tefé FLONA and Amanã RDS. The Tapajós FLONA presented the highest average immigration rate (0.98). These results indicate that distinct human-piquiá interactions occurred across the study 8ncip. In this 8ncipien, the Tapajós FLONA stands out by presenting the highest genetic diversity and immigration rate, alongside lower genetic diversity in home gardens, suggesting that this 8nci may have been the center of origin for the domestication of the species, with na ongoing 8ncipiente domestication process.

Keywords: Genomics; SNPs; Genetic Diversity; Genetic Structure.

LISTA DE FIGURAS

Figura 1 - Localização das quatro unidades de conservação estudadas: FLONA Tapajós, FLONA Caxiuanã, FLONA Tefé e RDS Amanã, Amazônia brasileira.....21

Figura 2 - Matriz de 207 indivíduos e 6.900 SNPs e testados valores de K de 1 a 10 populações ancestrais. A barra retangular em vermelho, indica o N° de K=4 como ideal como melhor resultado do modelo.....31

Figura 3 - Valores de entropia cruzada e matriz de ancestralidade com o K=4 sendo a melhor distribuição, as cores definem a presença em porcentagem na formação alélica de cada população..... 32

Figura 4. Matriz de coancestralidade entre os indivíduos das quatro unidades de conservação, azul escuro (FLONA Tefé), azul claro (RDS Amanã), vermelho (FLONA Tapajós) e rosa (FLONA Caxiuanã). As cores verde musgo, refere-se ao sistema de plantio quintal e verde escuro ao sistema de plantio “floresta” 33

LISTA DE TABELAS

Tabela 1. Quantidade de acesso amostral por unidade de conservação e ambiente de ocorrência.....	22
Tabela 2. Estimativas de diversidade genômica para 6.900 marcadores SNPs avaliados em 207 indivíduos de <i>Caryocar villosum</i> em quatro unidades de conservação. Número de indivíduos (N), heterozigosidade observada (Ho), heterozigosidade esperada (He), coeficiente de endogamia (FIS), significativo com intervalo de confiança de 95% de Ho (Ho_IC95%), significativo com intervalo de confiança de 95% de He (He_IC95%), significativo com intervalo de confiança de 95% de FIS (FIS_IC95%), número de alelos (NA), riqueza alélica (RA), alelos privados (Ap), alelos privados rarefeitos (APR). Os números destacados em negrito indicam diferenças estatisticamente significativas.....	26
Tabela 3. Estimativas de diversidade genética genômica para 6.900 marcadores SNPs avaliados em 207 indivíduos de <i>Caryocar villosum</i> em quatro unidades de conservação, considerando os sistemas de quintal e floresta de cada unidade de conservação. Número de indivíduos (N), heterozigosidade observada (Ho), heterozigosidade esperada (He), coeficiente de endogamia (FIS), significativo com intervalo de confiança de 95% de Ho (Ho_IC95%), significativo com intervalo de confiança de 95% de He (He_IC95%), significativo com intervalo de confiança de 95% de FIS (FIS_IC95%), número de alelos (NA), riqueza alélica (RA), alelos privados (Ap), alelos privados rarefeitos (APR). Os números destacados em negrito indicam diferenças estatisticamente significativas.....	27
Tabela 4. Estimativas de diversidade genética genômica para 6.900 marcadores SNPs avaliados em 207 indivíduos de <i>Caryocar villosum</i> em quatro unidades de conservação, considerando os sistemas de quintal e floresta. Número de indivíduos (N), heterozigosidade observada (Ho), heterozigosidade esperada (He), coeficiente de endogamia (FIS), significativo com intervalo de confiança de 95% de Ho (Ho_IC95%), significativo com intervalo de confiança de 95% de He (He_IC95%), significativo com intervalo de confiança de 95% de FIS (FIS_IC95%), número de alelos (NA), riqueza alélica (RA), alelos privados (Ap), alelos privados rarefeitos (APR). Os números destacados em negrito indicam diferenças estatisticamente significativas.....	28
Tabela 5. Análise de variância molecular (AMOVA) baseada a partir da média de 6.900 loci SNPs encontrados em 207 amostras de <i>Caryocar villosum</i> em quatro unidades de conservação da Amazônia Brasileira, e, dentro dos sistemas quintal e floresta. Valores em negrito indicam as estatísticas correspondentes a diferenciação entre grupos. Os números destacados em negrito indicam diferenças estatisticamente significativas.....	29
Tabela 6. Análise de variância molecular (AMOVA) em <i>Caryocar villosum</i> par a par, comparando as variações de duas unidades de conservação por vez. Valores em negrito indicam as estatísticas correspondentes a diferenciação entre grupos. Os números destacados em negrito indicam diferenças estatisticamente significativas par a par.....	30
Tabela 7. Análise de variância molecular (AMOVA) em <i>Caryocar villosum</i> dentro das unidades de conservação. Valores em negrito indicam as estatísticas correspondentes a diferenciação entre os sistemas. Os números destacados em negrito indicam diferenças estatisticamente significativas	30
Tabela 8. Estimativas de fluxo gênico entre as unidades de conservação com o MIGRATE...	34

Tabela 9. Estimativas de fluxo gênico entre as unidades de conservação com o MIGRATE, centro de distribuição.....	35
---	----

SUMÁRIO

1. INTRODUÇÃO	13
2. OBJETIVOS	15
2.1 Objetivo geral	15
2.2 Objetivos específicos	15
3. REFERENCIAL TEÓRICO	16
3.1 Domesticação de populações de plantas	16
3.2 Marcadores moleculares	18
3.3 Piquiá: distribuição, importância e interações com humanos	20
4. METODOLOGIA	22
4.1 Amostragem	22
4.2 Extração de DNA	23
4.3 Construção da Biblioteca Genômica e sequenciamento	24
4.4 Processamento de sequência bruta	24
4.5 Diversidade Genética e AMOVA	24
4.6 Estrutura Genética	25
4.7 Migração e Fluxo Gênico	26
5. RESULTADOS	27
5.1 Diversidade genética	27
5.2 Estrutura Populacional	30
5.3 Migração e Fluxo Gênico	35
6. DISCUSSÃO	37
6.1 Diversidade Genética	37
6.2 Comparação Quintal vs. Floresta e o Processo de Domesticação	38
6.3 Estrutura Genética e variações populacionais	39
6.4 Fluxo Gênico e Padrões de Migração	40
7. CONCLUSÃO	42
8. REFERÊNCIAS BIBLIOGRÁFICAS	43

1. INTRODUÇÃO

A domesticação de populações de plantas é um processo co-evolutivo, no qual humanos selecionam e manejam os indivíduos, propagando fenótipos de seus interesses, podendo resultar em mudanças nos genótipos e fenótipos das populações manejadas em comparação com as populações silvestres (Clement, 1999a). Enquanto processo, as populações de plantas podem apresentar diferentes graus de domesticação, sendo classificadas de acordo com a dependência do manejo humano e alterações esperadas nos genótipos e fenótipos: incipientemente domesticadas, semi-domesticadas e domesticadas (Harlan 1992:64; Clement, 1999a).

Em populações incipientemente domesticadas espera-se divergências fenotípicas e genotípicas pequenas, especialmente quando há fluxo gênico entre estas populações e populações silvestres (Clement, 1999a; Miller; Gross, 2011; Casas *et al.*, 2016). Na Amazônia, existem evidências de que o processo de domesticação de populações de plantas ocorre desde o final do Pleistoceno e início do Holoceno, sendo o bioma um importante centro de origem de domesticação (Clement, 1999a, 1999b; Clement *et al.*, 2010). Estima-se que pelo menos 85 espécies de árvores e palmeiras tenham populações domesticadas, dentre elas, 51 apresentam populações incipientemente domesticadas (Levis *et al.*, 2017). As espécies com populações incipientemente domesticadas são as mais abundantes na paisagem (Coelho *et al.*, 2021), e um padrão de distribuição agregado é muitas vezes um indicador da domesticação (Levis *et al.*, 2017).

Uma das espécies amazônicas com populações incipientemente domesticadas é o piquiá (*Caryocar villosum* (Aubl.) Pers., Caryocaraceae) (Clement, 1999a, 1999b; Alves *et al.*, 2016). O piquiá é exclusivamente neotropical, distribuindo-se por toda Amazônia (Prance; Silva, 1973).

No baixo rio Tapajós a população humana local seleciona frutos não amargos para cultivo e existem piquiazais (agregações de piquazeiros) ocorrendo nas proximidades de sítios arqueológicos de terra preta de índio, antigos assentamentos humanos (Alves *et al.*, 2016). Além do recente processo de seleção, a distribuição agregada da espécie, próxima aos sítios arqueológicos, e evidências genéticas sugerem um processo histórico de domesticação de piquiá na região (Alves *et al.*, 2016; Francisconi *et al.*, 2021). Esta região também foi identificada como um centro de diversidade genética da espécie (Francisconi *et al.*, 2021). Alves (2023), através de estudos ecológicos e genéticos, comparando a Amazônia Oriental e Ocidental encontrou maior ocorrência, abundância e diversidade genética na Amazônia Oriental, sugerindo a região da FLONA Tapajós como centro de origem da domesticação de populações de *C. villosum*.

Apesar da importância do piquiá, do constante uso por populações tradicionais e do potencial comercial, a espécie carece de estudos. Estudos populacionais da espécie podem contribuir com a elucidação do seu processo de domesticação, numa perspectiva histórica. Neste sentido, é interessante caracterizar geneticamente diferentes formas de ocorrência como, quintais (áreas cultivadas atuais correspondendo à domesticação recente) e agregações em florestas maduras (podendo corresponder à domesticação passada) (Alves, 2023). Além disso, a compreensão de como a variação genética da espécie está distribuída pode contribuir para a tomada de decisões para conservar a diversidade genética e auxiliar em planos de manejo (Loveless; Hamrick, 1984).

Estudos genéticos de domesticação de populações de plantas na Amazônia geralmente utilizam marcadores moleculares, principalmente marcadores microssatélites ou de repetição de sequência simples (SSR) (Pereira *et al.*, 2018). Para piquiá existem estudos utilizando marcadores microssatélites relacionados à caracterização da distribuição da espécie na Amazônia brasileira, verificando se o Rio Negro seria uma barreira geográfica natural para a dispersão (Collevatti *et al.*, 2009); e comparando árvores de quintais e florestas para investigar o processo de domesticação em duas unidades de conservação (Resex Rio Ouro Preto e FLONA Tapajós) (Francisconi *et al.*, 2021). No entanto, problemas com alelos nulos são bastante comuns na genotipagem com microssatélites, podendo resultar em estimativas não verdadeiras, envolvendo o aumento da homozigosidade observada e alterações das estimativas da diversidade genética (Van Oosterhout *et al.*, 2006). Abordagens baseadas em sequenciamento de nova geração e identificação de polimorfismo de nucleotídeo único (em inglês: single nucleotide polymorphisms – SNPs) são mais informativos e a redução dos custos para o desenvolvimento destas técnicas os têm tornados mais acessíveis (Seeb *et al.*, 2011).

Para *C. villosum*, só existe um estudo utilizando marcador SNPs, o qual demonstrou maior diversidade genética na região mais à leste da Amazônia brasileira, levantando a hipótese de que essa região tenha sido um centro de origem de domesticação de populações da espécie (Alves, 2023). No entanto, no referido estudo, o baixo número amostral, bem como a baixa qualidade dos dados genômicos, não permitiram realizar análises populacionais mais refinadas e nem comparar árvores de quintais e florestas para aferir os índices de diversidade nos dois sistemas de plantio. Assim, nessa proposta utilizamos marcadores SNPs para estudar a estrutura e diversidade genética de *C. villosum* em populações situadas no interior ou no entorno de quatro unidades de conservação da Amazônia brasileira, comparando geneticamente as árvores cultivadas em quintais e presentes em florestas maduras.

2. OBJETIVOS

2.1 Objetivo geral

Caracterizar a diversidade e estrutura genética de populações de piquiá (*C. villosum*) de quatro unidades de conservação da Amazônia brasileira (Floresta Nacional - FLONA Tapajós, FLONA Caxiuanã, FLONA Tefé e Reserva de Desenvolvimento Sustentável - RDS Amanã), comparando árvores em quintais e florestas maduras, a partir de dados genômicos.

2.2 Objetivos específicos

1. Caracterizar a diversidade genética de piquiazeiros de quintais e florestas maduras.
2. Analisar a diversidade e a estrutura genética das populações nas quatro unidades de conservação, inferindo possíveis centros de origem de domesticação da espécie.
3. Buscar estrutura genética e estimar o fluxo gênico entre árvores das quatro unidades de conservação.
4. Averiguar evidências genéticas que apoiam a Floresta Nacional Tapajós como principal centro de domesticação da espécie.

Hipóteses

1. Há estruturação genética e não há fluxo gênico entre as populações das quatro unidades de conservação, devido à distância geográfica.
2. Em uma mesma unidade de conservação há fluxo gênico entre os sistemas de quintais e florestas maduras, devido ao manejo humano.
3. Nos quintais existe menor diversidade genética do que nas florestas maduras, devido a seleção humana.
4. Existem evidências genéticas que apoiam a Floresta Nacional Tapajós como centro de domesticação da espécie.

3. REFERENCIAL TEÓRICO

3.1 Domesticação de plantas na Amazônia

Os humanos podem ter chegado às terras baixas da América do Sul por volta de 20.000 anos antes do presente (AP) (Lahayee *et al.*, 2013), data contestada por Roosevelt (2013) que estima em 13.000 AP. As comunidades humanas pré-cerâmicas ocuparam a área de Porto Velho e Rondônia em 12.000 AP (Miller, 1992). O conhecimento sobre as plantas fez com que os humanos realizassem a seleção de fenótipos, determinando quais plantas seriam promovidas, manejadas e cultivadas, tornando-as mais úteis e adaptadas para a sua própria intervenção (Harlan, 1992).

A seleção e propagação das plantas é um processo denominado de domesticação e é descrito como co-evolutivo, pois são esperadas mudanças tanto nos fenótipos e genótipos das plantas quanto dos humanos (Harlan, 1992; Clement, 1999a). A seleção humana pode ser intencional, promovendo mudanças na população, embora ocorra com mais frequência de forma não intencional, na medida em que não há o objetivo de mudar as características populacionais (Darwin, 1859). A seleção de características desejáveis da planta (por exemplo: frutos adocicados e maiores) pode corresponder a uma seleção direcional que, geralmente, resulta em maior variabilidade fenotípica da população domesticada, pois inclui tanto os tipos comuns na população silvestres, quanto os tipos novos (Clement, 1999b).

As mudanças das características fenotípicas, diferenciando as populações domesticadas das silvestres, são conhecidas como síndromes de domesticação (Clement, 1999a). Alguns exemplos de síndromes de domesticação incluem: o aumento do tamanho das sementes e diminuição da dispersão, o que facilita o controle de cruzamentos; a dormência de sementes, de modo que a germinação ocorra na época de produção desejada; o crescimento compacto, a fim de diminuir a competição; e o aumento da autopolinização e da reprodução vegetativa, para manter os fenótipos similares (Gepts, 2004).

Em populações domesticadas, além das mudanças fenotípicas, também ocorrem mudanças genotípicas. A propagação seletiva das linhagens de interesse pode resultar no progressivo estreitamento da base genética das populações domesticadas (Tanksley; Mccouch, 1997) e, dessa forma, a redução da diversidade genética é um processo característico das populações domesticadas (Gepts, 2004).

A domesticação de plantas, enquanto processo, pode apresentar diferentes graus. De acordo com Clement (1999b), populações domesticadas podem ser classificadas em: incipientemente domesticadas, que apresentam fenótipos pouco diferentes dos encontrados na natureza; semi-domesticadas, as quais passaram por seleção humana mais intensa; e totalmente

domesticadas, quando a pressão de seleção gera perda crítica da adaptação ecológica, havendo grande dependência do manejo humano e dos ambientes alterados. Em populações domesticadas, nos seus diferentes graus, espera-se menor variabilidade genética devido aos gargalos genéticos, decorrentes do processo de seleção (Franciscone *et al.*, 2021).

A redução da variabilidade genética pode resultar também do efeito da deriva genética, devido à coleta de um limitado número de plantas (Weber *et al.*, 2001; Hollingsworth *et al.*, 2005; Lengkeek *et al.*, 2005). Estima-se que nos últimos mil anos, 80% da variação genética de cultivos domesticados foi perdida em relação a seus parentes silvestres, como aconteceu com as culturas da soja, trigo, café, feijão e pêssego (Tanksley; Mccouch, 1997; Tuxill, 1999; Lópezgartner *et al.*, 2009; Bitocchi *et al.*, 2013; Cao *et al.*, 2014).

Agregações florestais podem ser resultado de manejo e domesticação humana passados, assim como em ambientes manejados no presente, como roças e quintais, podem ter o processo de domesticação recente e ainda em curso (Alves *et al.*, 2016). Lins *et al.* (2015) estudaram ocupações modernas no Rio Urubu, um tributário do Rio Amazonas a 200 km de Manaus, e encontraram 205 espécies cultivadas em 40 quintais, com uma média de 8,2 (± 6.6) espécies nativas da Amazônia por quintal, sendo as espécies mais comuns: cupuaçu (*Theobroma grandiflorum*), presente em 65% dos quintais, mandioca (*Manihot esculenta*), em 45%, e pimenta amazônica (*Capsicum chinense*) em 43%. Os diferentes graus de domesticação das populações de plantas, encontrados no estudo foram explicados pelo manejo histórico, abandono, reocupação das áreas por povos indígenas (Lins *et al.*, 2015).

A uva Amazônica ou mapati (*Pourouma cecropiifolia*) é outra espécie com populações domesticadas. Um estudo ecológico realizado por Pedrosa *et al.* (2018), em comunidades e vilas do Amazonas, mostrou que a domesticação do mapati resultou no aumento do fruto, da sua massa e da massa da semente. No entanto, o processo de domesticação, resultou em uma diminuição da quantidade de frutos presentes em cada cacho, tendo, surpreendentemente, uma maior variabilidade genética em plantas sob seleção humana.

Moreira *et al.* (2015), estudou plantas cultivadas e não cultivadas de urucum (*Bixa orellana*) analisando suas morfologias, em seis principais rios da bacia Amazônica. Os autores observaram uma maior produção de pigmentos nas plantas cultivadas e a abertura dos frutos após a maturação (Moreira *et al.*, 2015).

Muhlen *et al.* (2019), estudou populações cultivadas de mandioca amarga e doce no Brasil, com objetivo de analisar a diversidade genética e estrutura populacional da espécie, utilizando dez marcadores microssatélites SSR. Os SSR nucleares revelaram alta diversidade populacional. O número médio de alelos foi de 9,7, enquanto a média de heterozigosidades

observadas e esperadas foram respectivamente de 0,576 (desvio padrão $\pm 0,174$) e 0,655 (desvio padrão $\pm 0,198$), a análise de variância molecular indicou divergência intermediária entre mandioca amarga e mandioca doce ($FST = 0,081$), com maior variação dentro dos grupos.

As regiões das variedades doces da bacia Amazônica estão intimamente relacionadas entre si, havendo também um risco de perda da ancestralidade do alto Rio Negro à medida que a distância geográfica daquela região aumenta. A ancestralidade geral da mandioca brava brasileira é predominante em toda a Amazônia central e oriental. Neste contexto, o uso de marcadores moleculares tem se tornado uma importante ferramenta para estudos de genética de populações, sendo utilizados para quantificar diversidade genética, estrutura genética e fluxo gênico entre e dentro de populações de espécies vegetais.

3.2 Marcadores moleculares

Marcadores moleculares, são fenótipos moleculares provenientes de quaisquer segmentos expressos ou não de DNA que são herdados geneticamente e se prestam à diferenciação de um ou mais indivíduos (Ferreira; Grattapaglia, 1998; Milach, 1998). Dentre as vantagens da utilização de marcadores moleculares em relação aos marcadores morfológicos, estão o fato de serem mais numerosos e apresentarem maior quantidade de polimorfismos. Além disso, comportam-se como caracteres de herança simples e previsíveis, podendo não ser afetados por variações do meio ambiente. Desta forma, os marcadores moleculares permitem quantificar a diversidade genética, seguir o movimento de indivíduos, estimar a endogamia, caracterizar as novas variedades e espécies e avaliar os padrões históricos de dispersão da espécie ou de determinadas populações (Freeland, 2005).

Diferentes técnicas podem resultar na origem de marcadores moleculares, em que cada técnica utiliza uma estratégia particular para detectar os polimorfismos. Diferenças entre as técnicas é dada por características como a abundância no genoma, níveis de polimorfismos detectados, especificidade de *locus*, reprodutibilidade, requerimentos técnicos e custos (Mondini *et al.*, 2009).

Os avanços nas abordagens genômicas proporcionam novas oportunidades para avaliar a variação ao nível do genoma completo e para compreender os mecanismos de adaptação em diferentes espécies de cultivos (Bevan; Uauy, 2013). As ferramentas genômicas também são valiosas para a caracterização de populações selvagens de espécies agrícolas e para apoiar a sua utilização em programas de melhoramento (Brozynska *et al.*, 2016).

As abordagens genômicas foram facilitadas por novos métodos de sequenciamento de próxima geração (em inglês *next-generation sequencing*; NGS). Tais métodos estão permitindo a descoberta de muitos polimorfismos de nucleotídeo único (em inglês *single nucleotide polymorphism*; SNPs) com custo reduzido (Allendorf *et al.*, 2010; Kess *et al.*, 2015). Várias abordagens genômicas baseadas em NGS dependem do sequenciamento de regiões de DNA associadas ao local de restrição (RADseq) (Baird *et al.*, 2008), com isso os SNPs vêm sendo amplamente utilizados em estudos de genômica populacional.

Os SNPs são variações na sequência de DNA que afetam somente uma base nitrogenada (adenina, timina, citosina ou guanina) na sequência do genoma entre indivíduos de uma espécie ou entre pares de cromossomos de um indivíduo. Um SNP é uma substituição da linha germinativa de um único nucleotídeo em uma posição específica no genoma (Li *et al.*, 2009). Além dos marcadores SNPs serem abundantes, suas bases moleculares permitem que haja uma distribuição homogênea pelo genoma (Caetano, 2009). Com isso, avanços tecnológicos recentes trouxeram metodologias de alto desempenho e acurácia, baixo custo para prospecção, caracterização e genotipagem de marcadores SNPs (Baird *et al.*, 2008).

O polimorfismo revelado em *locus* RAD-seq originam-se nas sequências adjacentes aos sítios de restrição, onde são identificados SNPs. Alternativamente, a ocorrência de mutações nos sítios de restrição ao longo do genoma dos indivíduos, causam a presença ou ausência de determinada sequência de RAD-seq nos indivíduos de uma população (Baird *et al.*, 2008), podendo então, ser também analisados como marcadores dominantes. O sequenciamento de *locus* RAD-seq se dá usualmente por meio da plataforma Illumina (Bentley *et al.*, 2008). O sequenciamento massivo e paralelo de amostras multiplex permite o rápido descobrimento de milhares a milhões de *locus* RAD-seq e por conseguinte a genotipagem em larga escala de populações com muitas centenas ou milhares de marcadores SNPs (Alves-Pereira, 2019).

Os SNPs são os tipos de variantes genéticas mais frequentes e dispersos ao longo do genoma das espécies, estimando-se que correspondam à 90% da variação genética observada em humanos (Collins *et al.*, 1998). Em uma análise de múltiplos táxons, incluindo plantas, é provável que variantes do tipo SNPs ocorram a cada 200-500 pares de bases em regiões não-codificadoras, e a cada 500-1000 pares de bases em regiões codificadoras (Brumfield *et al.*, 2003). Este advento das tecnologias de sequenciamento de nova geração, aliado à abundância e ocorrência ao longo de todo o genoma tornaram os SNPs os marcadores de maior apelo para a genômica de populações (Willing *et al.*, 2012).

3.3 Piquiá: distribuição, importância e interações com humanos

A espécie *C. villosum* - conhecida popularmente como piquiazeiro ou piquiá - pertence à família Caryocaraceae, sendo exclusivamente neotropical (Prance; Silva, 1973). A planta atualmente está classificada dentro da ordem Malpighiales, composta por 25 espécies agrupadas em dois gêneros: *Caryocar* e *Anthodiscus*. São poliploides, com $2n = 46$ cromossomos (Ehrendorfer *et al.*, 1984). O piquiá é amplamente distribuído pela Amazônia (Prance; Silva, 1973). Ao Norte ocorre desde o escudo da Guiana até Caiene (Guiana Francesa); ao leste, até o estado do Pará e Roraima (Brasil), e à oeste, até o Rio Negro. Na região sul da Amazônia, a distribuição é limitada pelo escudo brasileiro, que começa em São Luís (Maranhão, Brasil), ao leste entre Pará e Amazonas, até o oeste do norte de Rondônia (Prance; Silva, 1973).

O pericarpo do piquiá apresenta bastante óleo. No início de 1.900, a espécie foi sugerida como uma cultura oleaginosa alternativa, mas ela fracassou comercialmente nas plantações da Malásia, devido ao mercado altamente competitivo com gorduras e óleos mais baratos de fontes mais abundantes (Lane, 1957). Apesar do fracasso no passado, o potencial comercial da espécie ainda é discutido podendo representar uma fonte alternativa de renda para povos tradicionais da Amazônia (Clement, 1999c; Shanley; Medina, 2005).

Além do valor comercial, os povos tradicionais amazônicos usam a espécie para diferentes finalidades, tais como alimentação, uso medicinal, madeireiro, atrativo de caça, dentre outros (Shanley; Medina, 2005; Alves *et al.*, 2016). Yamaguchi e Souza (2021) estudaram as propriedades fitoquímicas dos frutos de piquiá coletados em comunidades e feiras do município de Coari, no Amazonas, detectando a presença de antraquinonas e leucoantocianidina, compostos com potenciais benefícios à saúde humana, além de destacarem o uso da espécie por povos tradicionais da região para fins alimentícios e medicinais.

Na FLONA Tapajós no Pará, povos tradicionais das comunidades Maguari e Jamaraquá citaram como principais usos do piquiá: alimentício e medicinal (Alves *et al.*, 2016). O óleo da polpa do fruto foi indicado para tratamento de queimaduras na pele, micoses, dores musculares, dentre outros (Alves *et al.*, 2016). O óleo também foi mencionado como cosmético, sendo usado para a fabricação de sabões, perfumes, cremes e repelentes de insetos (Alves *et al.*, 2016). Estes povos, preferem consumir e selecionar frutos menos amargos para cultivo (Alves *et al.*, 2016). Plantações intencionais são feitas em quintais e roças, sendo também encontradas agregações em floresta madura, próximas aos sítios arqueológicos; essas agregações possivelmente foram resultantes de manejos realizados pelos Povos Indígenas no passado, indicando um processo de domesticação histórico na região (Alves *et al.*, 2016).

Segundo Clement (1999a) as populações do Tapajós são incipientemente domesticadas, ou seja, elas foram modificadas pela seleção humana, mas o fenótipo médio, para as características que são objeto de seleção, ainda está dentro da variação encontrada na natureza. Nesta mesma região, Francisconi *et al.* (2021), em um estudo utilizando sete *loci* microssatélites, comparou árvores florestais e de quintais, encontrando um pouco menos de diversidade genética nos quintais, interpretando como sendo uma consequência do processo de domesticação recente da espécie. Além disso, ao comparar populações da FLONA Tapajós com outra região no oeste da Amazônia (Reserva Extrativista - RESEX Rio Ouro Preto) foi encontrada maior diversidade genética na FLONA e quatro grupos genéticos (Francisconi *et al.*, 2021). Os autores levantam a hipótese de que a região pode ser um centro de domesticação da espécie (Francisconi *et al.*, 2021).

Um outro estudo genético, com 11 populações de piquiá ao longo da Amazônia brasileira, envolvendo SNPs, encontrou um grupo monofilético, formando as populações situadas mais à oeste da Amazônia, possivelmente resultante de um ancestral do leste da Amazônia (Alves, 2023). A autora sugere que populações de piquiá provavelmente foram domesticadas no leste da Amazônia, com posterior dispersão por humanos para o oeste (Alves, 2023).

Já um outro estudo genético, com dez microssatélites, demonstrou que o Rio Negro não representou barreira geográfica para a dispersão da espécie (Collevatti *et al.*, 2009). A falta de diferenciação genética entre as árvores de lados opostos da bacia pode estar relacionada a fatores ecológicos, como a polinização e dispersão de sementes a longa distância, visto, por exemplo, que a polinização é feita principalmente por morcegos do gênero *Phyllostomus* (Gribel; Hay, 1993; Martins; Gribel, 2007), mas também pode estar relacionada à dispersão humana, onde há preferências por frutos não amargos para consumo, fazendo com que levem sementes com essas características e as cultivem (Alves *et al.*, 2016).

O conhecimento científico sobre uma espécie é primordial para tomada de decisões estratégicas para a sua conservação. Segundo o estudo de Alves *et al.* (2016), nas comunidades de Maguari e Jamaraquá na FLONA Tapajós, o piquiá está entre as dez espécies mais utilizadas pelos seus moradores, não sendo incluída no manejo madeireiro da FLONA. O piquá pode apresentar tanto benefícios sociais quanto econômicos, havendo também importância biotecnológica a ser melhor explorada (Yamaguchi; Souza, 2021).

4. MATERIAIS E MÉTODOS

4.1 Amostragem

Para objeto de estudo foram utilizadas 217 amostras de folhas de árvores de piquiá presentes dois tipos de ambientes: i) quintais - áreas de cultivo com manejo mais intenso, nesta categoria também foram incluídas roças, pomares e capoeiras mais novas; e ii) florestas maduras - áreas com árvores maiores, com manejo menos intensivo. As amostras foram coletadas no interior e/ou entorno de quatro unidades de conservação (UCs) da Amazônia brasileira: FLONA Tapajós, FLONA Caxiuanã, FLONA Tefé e RDS Amanã (Figura 1; Tabela 1). Após a coleta, as amostras foram acondicionadas em filtro de papel, secas com sílica em gel e guardadas em freezer -20° , no Laboratório de Evolução Aplicada (LEA), do Instituto de Ciências Biológicas (ICB) da Universidade Federal do Amazonas (UFAM).

Figura 1. Localização das quatro unidades de conservação estudadas: FLONA Tapajós, FLONA Caxiuanã, FLONA Tefé e RDS Amanã, Amazônia brasileira.

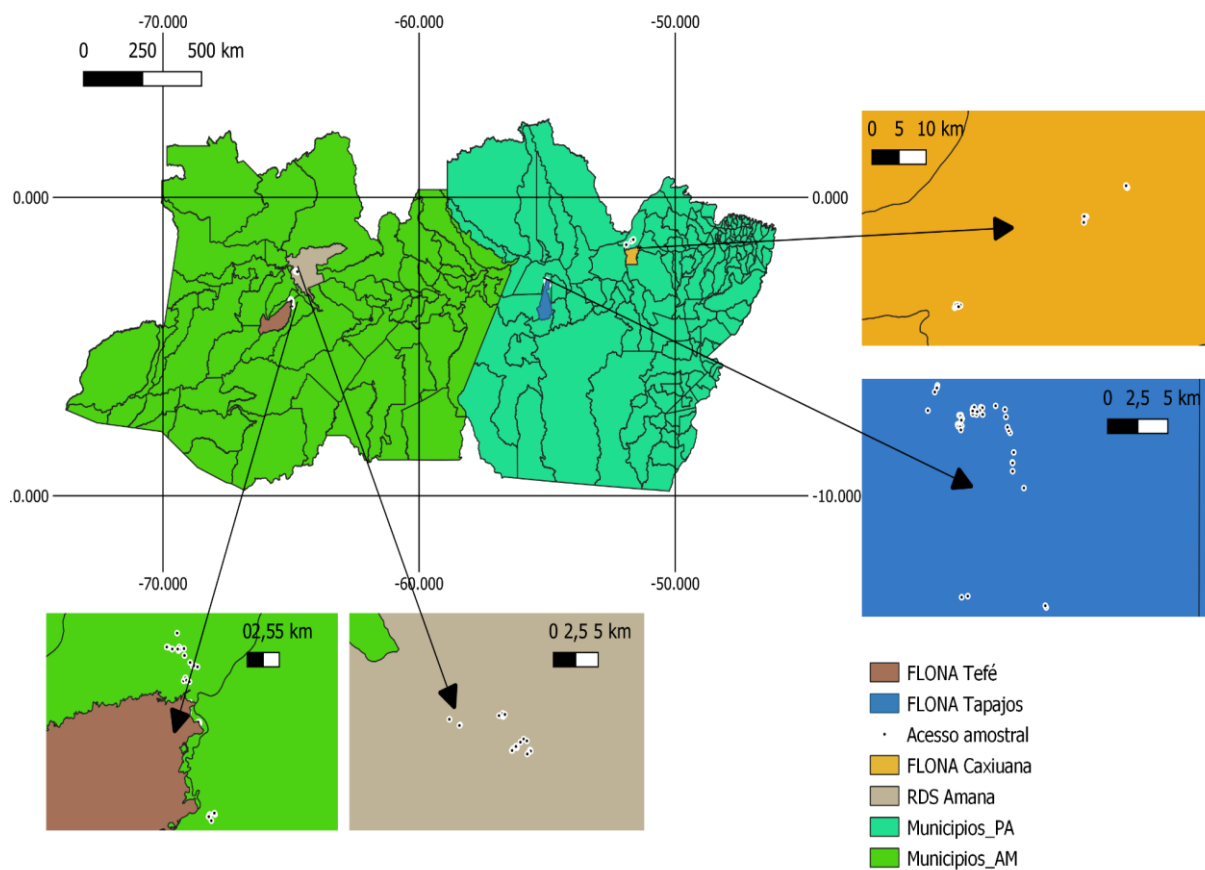


Tabela 1. Quantidade de acesso amostral por unidade de conservação e ambiente de ocorrência.

Unidade de conservação	FLONA Tapajós		FLONA Caxiuanã		FLONA Tefé		RDS Amanã	
Local de ocorrência	Florestas	Quintais	Florestas	Quintais	Florestas	Quintais	Florestas	Quintais
Quantidade amostral	50	15	19	33	22	36	11	21
Total de Amostras	65		52		58		32	

4.2 Extração de DNA

A extração de DNA foi realizada no LEA – ICB/UFAM. O DNA foi extraído das folhas por meio do protocolo utilizando CTAB 2 % (Brometo de Cetiltrimetilamônio) com adaptações (Doyle; Doyle, 1987). Após a maceração de 50mg de tecido, utilizando precelisys, foram adicionados 1000 µl de CTAB 2 %, previamente aquecidos a 65 °C, 14µl de 2-mercaptoetanol e 2µl de proteinase K (20mg/µl). As amostras foram então colocadas por 60 minutos no banho-maria a 60°C, sendo homogeneizadas com movimentos suaves a cada dez minutos.

Após esse período, foram adicionados 600µl de CIA (clorofórmio-álcool isoamílico) e centrifugado por dez minutos à 12.000 rpm. Após a centrifugação, toda a fase aquosa foi transferida para um novo tubo de 2 ml, adicionado 80µl de CTAB 10%, 600µl de CIA e centrifugado por dez minutos à 14.000 rpm. Após a centrifugação, a fase aquosa foi transferida para um novo tubo de 1,5ml. Foi adicionado junto a fase aquosa 500µl de isopropanol gelado e 34µl de acetato de sódio 3 molar, e incubado por uma hora no freezer a -20°C.

Na sequência, foi centrifugada por dez minutos à 12.000 rpm em 4°C. Após a centrifugação, foi descartado o sobrenadante, ficando somente os pellets nos tubos. Foram realizadas duas lavagens com 1.000µl de álcool 70%. Após as lavagens, os pellets foram secos em temperatura ambiente por aproximadamente duas horas. Após a secagem, foi adicionado 50µl de tampão TE com RNase e realizada incubação em banho-maria, por 30 minutos à 37°C. O DNA extraído foi armazenado no freezer -20°C. A quantificação das amostras foi estimada pela intensidade de fluorescência emitida pelo corante Gel Red sob UV em géis de agarose a 0,8 % (eletroforese), comparada aos de padrões com pesos moleculares e concentrações

específicas e conhecidas de DNA do fago λ (20, 50, 100 e 200 ng). A eletroforese foi realizada em uma cuba contendo tampão TBE 1 % a uma voltagem de 2 V/cm.

4.3 Construção da biblioteca genômica e sequenciamento

A redução da complexidade genômica foi realizada seguindo o método de genotipagem por sequenciamento (GBS), envolvendo dupla digestão de DNA (ddGBS) (Poland *et al.*, 2012). Os procedimentos foram realizados no Laboratório de Evolução e Genética Animal (LEGAL) - ICB/UFAM. Para cada amostra, foram utilizados aproximadamente 5 μ g de DNA genômico. Para purificação das amostras foram utilizados *beads* magnéticos. Sequências de adaptadores e barcodes foram ligados aos fragmentos do DNA genômico total digeridos com as enzimas de restrição FSP B1 e EcoR1. Os fragmentos, associados aos sítios de restrição desta enzima, foram então submetidos ao sequenciamento na plataforma Illumina, pela empresa Norte Americana GENOME. As sequências geradas foram separadas por indivíduos e tiveram suas sequências barcodes removidas pela empresa que realizou o sequenciamento, resultando em fragmentos RAD de cerca de 150 pb.

4.4 Processamento de sequência bruta

As sequências obtidas do sequenciamento das 217 amostras, foram filtradas e agrupadas por similaridade com iPyrad (Eaton; Overcast, 2020). Foram eliminadas sequências contendo mais de 5 bases com escore de qualidade menor que 30 e que apresentavam cobertura inferior a 6 cópias, bem como aquelas menores que 35 pb. As sequências de adaptadores e códigos de barras já haviam sido retiradas pela empresa responsável pelo sequenciamento. Foi realizado agrupamento (intra e interindividual) entre sequências com no mínimo 85% de identidade. Para evitar possíveis parálogos, foram descartados os loci que apresentaram mais de 50% de sítios heterozigotos e aqueles que continham mais de 20% de sítios variáveis ou mais de 8 *indels*. Destas, 6 foram descartadas por terem cobertura virtualmente nula (0 a 7), ficando, portanto, 216 amostras. Foram eliminados quatro indivíduos com maior quantidade de dados perdidos (*missing data*), que aumentava a matriz resultante em mais de 50%.

4.5 Diversidade Genética e AMOVA

A matriz original de SNPs não ligados continha 28.469 SNPs e 216 amostras de *C. villosum*. As filtrações foram realizadas no software R (R Core Team, 2021), incluindo a remoção de loci com frequência alélica mínima (MAF) inferior a 3, loci tri e tetra-alélicos, e SNPs com baixa cobertura entre os indivíduos, de modo que a amostra com menor cobertura mantivesse pelo menos 50% dos loci. Além disso, foram excluídos indivíduos com elevada proporção de dados ausentes (*missing data*), que elevavam a matriz resultante em mais de 50%.

Após esses procedimentos, a matriz final foi composta por 207 amostras e 6.900 SNPs, garantindo que todas as amostras apresentassem, no mínimo, 50% dos loci.

Foram estimados parâmetros de diversidade genética: entre as quatro UCs entre sistemas quintais e florestas de todas as UCs conjuntamente; e entre quintais e florestas separadamente, dentro de cada UC. Heterozigosidades esperada e observada e Fis foram calculados para cada locus usando-se a função `basic.stats` do pacote `hierfstat` (Goudet; Jombart, 2021). Os índices gerais foram obtidos a partir das médias entre loci, e intervalos de confiança a 95% por meio de 1.000 bootstraps de loci. O número de alelos foi computado para cada UC, somando-se o número de diferentes alelos presentes em cada locus. A riqueza alélica foi calculada usando-se a função `allelic.richness` do pacote `hierfstat`. Alelos privados em cada população foram contados com a função `private_alleles` do pacote `poppr` (Kamvar *et al.*, 2015), e rarefeitos a partir da população com menor número amostral usando um script customizado.

Para a realização da análise de variância molecular (AMOVA), foi utilizado o pacote `poppr` (Kamvar *et al.*, 2014), que automatiza a análise de AMOVA do pacote `pegas` (Paradis, 2010). Foram testados grupos fora do contexto hierárquico, apenas UCs e apenas sistemas, no geral, assim como comparações par-a-par entre as UCs e comparações dos sistemas dentro de cada UC, separadamente.

4.6 Estrutura genética

A estrutura populacional foi estimada a partir de duas abordagens, uma que utiliza todos os SNPs de um determinado locus e outra que utiliza um SNP aleatório de cada um dos loci polimórficos. A estrutura populacional (K) e as proporções de ancestralidade de cada indivíduo foram calculadas a partir de SNPs, por meio de uma fatoração não-negativa esparsa da matriz de genótipos com seleção guiada pela entropia cruzada (Frichot *et al.*, 2014). Foram testados valores de K de 1 a 10 populações ancestrais. Valores de entropia cruzada foram calculados para cada número de K (10 repetições), em cada um dos seguintes valores de alfa: 1, 10, 100, 1.000 e 10.000. Alfa é o peso de regularização que controla a esparsidade das proporções de ancestralidade, e serve como uma variável de ajuste para datasets pequenos (*i.e.*, com matrizes menores que ~10 mil SNPs). Estes valores foram escolhidos conforme as recomendações de Frichot *et al.* (2014). Os valores finais de entropia cruzada foram estipulados como sendo a menor média obtida para cada valor de K dentre os diferentes valores de alfa testados. A partir do resultado de entropia cruzada, 100 matrizes de admixture foram estimadas com o número ótimo de K, e aquela com o menor número de entropia cruzada, dentre as 100, escolhida como resultado para a matriz final.

Os loci RAD obtidos com iPyRAD foram convertidos para o formato finerad com o script fornecido com o fineRADstructure para este fim, o qual agrega os haplótipos por locus e remove loci inaptos (*e.g.*, invariáveis ou com amostragem insuficiente; colunas com deleções/excesso de bases ambíguas). Após a conversão 27.491 loci foram mantidos. Para manter consistência com as análises anteriores, foram utilizados os mesmos 207 indivíduos. Em seguida, foi estimada a matriz de coancestralidade com RADpainter, que quantifica, para cada indivíduo, a fração de haplótipos "copiados" de cada outro indivíduo. A inferência de estrutura populacional foi realizada no fineSTRUCTURE, este método foi desenvolvido para detectar estrutura populacional fina, fazendo uso de SNPs ligados em um mesmo locus gênico, gerando uma matriz de coancestralidade comparando, para cada locus, os SNPs variáveis entre todos os pares de indivíduos e, depois, usa essa matriz para fazer uma clusterização dos indivíduos por meio de uma MCMC. Com MCMC (1 milhão de burn-in; 1 milhão de iterações amostradas e mil de "thinning"), a partir da matriz de coancestralidade. A convergência/estabilização foi avaliada por inspeção visual do tracer da Posterior, extraído do log do fineSTRUCTURE e plotado como série temporal; confirmei ausência de tendências sistemáticas após o burn-in antes de interpretar a partição/árvore final.

4.7 Migração e fluxo gênico

Foram utilizados SNPs não ligados e apenas SNPs bialélicos foram retidos. Os loci foram requeridos a ter ambos os alelos representados em pelo menos duas populações, reduzindo, assim, potenciais vieses devidos a alelos privados ou variantes específicas de população. Além disso, os loci foram mantidos somente quando havia genótipos disponíveis para pelo menos 64 cromossomos (32 indivíduos diplóides), correspondendo ao menor tamanho amostral entre as populações (RDS Amanã). O conjunto de dados final foi composto por 1135 SNPs, formatado em HapMap, um dos formatos de dados reconhecidos pelo Migrate-n v5.0.4 (Beerli, 2006; Beerli; Palczewski, 2010). As análises foram conduzidas em um framework Bayesiano para estimar a taxa de mutação escalada pelo tamanho populacional (θ) e a taxa de migração escalada pela mutação (M), seguindo as recomendações de Beerli (2009). O programa foi executado com 1 cadeia longa (100 mil amostras, registradas a cada 100), 4 cadeias com diferentes temperaturas (Metropolis coupling) e um burn-in de 100 mil amostras por réplica.

5. RESULTADOS

5.1 Diversidade genética

As quatro UCs apresentaram heterozigosidade observada (H_o) inferior à heterozigosidade esperada (H_e), o que se reflete em coeficientes de endogamia (FIS) positivos em todas as UCs. Em alguma delas, os valores de FIS foram superiores a 0,42, indicando o cruzamento entre indivíduos aparentados.

De modo geral, as UCs apresentaram parâmetros de diversidade genética semelhantes. No entanto, a FLONA Tapajós apresentou os maiores valores de número de alelos (12.395), riqueza alélica (8.117) e alelos privados rarefeitos (505), enquanto a RDS Amanã apresentou os menores valores desses parâmetros (Tabela 2).

Tabela 2. Estimativas de diversidade genômica para 6.900 marcadores SNPs avaliados em 207 indivíduos de *Caryocar villosum* em quatro unidades de conservação. Número de indivíduos (N), heterozigosidade observada (H_o), heterozigosidade esperada (H_e), coeficiente de endogamia (FIS), significativo com intervalo de confiança de 95% de H_o ($H_o_IC95\%$), significativo com intervalo de confiança de 95% de H_e ($H_e_IC95\%$), significativo com intervalo de confiança de 95% de FIS ($FIS_IC95\%$), número de alelos (NA), riqueza alélica (RA), alelos privados (Ap), alelos privados rarefeitos (APR). Os números destacados em negrito indicam diferenças estatisticamente significativas.

Parâmetros de diversidade	RDS Amanã	FLONA Caxiuanã	FLONA Tapajós	FLONA Tefé
N	32	52	65	58
H_o	0,0836	0,0873	0,0871	0,0801
H_e	0,161	0,171	0,178	0,162
FIS	0,379	0,409	0,424	0,421
$H_o_IC95\%$	0,080; 0,086	0,084; 0,090	0,084; 0,089	0,077; 0,082
$H_e_IC95\%$	0,156; 0,165	0,167; 0,175	0,174; 0,182	0,158; 0,166
$FIS_IC95\%$	0,366; 0,392	0,399; 0,420	0,413; 0,435	0,410; 0,433
NA	11.059	11.658	12.395	11.767
RA	7.984	8.066	8.117	8.007
AP	111	281	306	417
APR	255	438	505	352

As estimativas dos parâmetros de diversidade genética dentro de cada sistema de plantio nas quatro unidades de conservação indicaram que a heterozigosidade observada (H_o) foi inferior à heterozigosidade esperada (H_e) em todos os sistemas. Consequentemente, os coeficientes de endogamia (FIS) foram positivos, com valores superiores a 0,41 em alguns sistemas.

Observou-se ainda que, na RDS Amanã e na FLONA Caxiuanã, os coeficientes de endogamia foram mais elevados nos sistemas de quintal (0,397 e 0,397, respectivamente) (Tabela 3).

De modo geral, verificou-se similaridade na diversidade genômica entre os sistemas de quintal e floresta. No entanto, o sistema de floresta da FLONA Tapajós apresentou os maiores valores de número de alelos, riqueza alélica, alelos privados e alelos privados rarefeitos (Tabela 3).

Tabela 3. Estimativas de diversidade genética genômica para 6.900 marcadores SNPs avaliados em 207 indivíduos de *Caryocar villosum* em quatro unidades de conservação, considerando os sistemas de quintal e floresta de cada unidade de conservação. Número de indivíduos (N), heterozigosidade observada (Ho), heterozigosidade esperada (He), coeficiente de endogamia (FIS), significativo com intervalo de confiança de 95% de Ho (Ho_IC95%), significativo com intervalo de confiança de 95% de He (He_IC95%), significativo com intervalo de confiança de 95% de FIS (FIS_IC95%), número de alelos (NA), riqueza alélica (RA), alelos privados (Ap), alelos privados rarefeitos (APR). Os números destacados em negrito indicam diferenças estatisticamente significativas.

UCS/Sistema	RDS Amanã Flo	RDS Amanã Qui	FLONA Caxiuanã Flo	FLONA Caxiuanã Qui
Ho	0,084	0,083	0,089	0,086
He	0,160	0,159	0,167	0,169
Fis	0,342	0,367	0,356	0,397
Ho_IC95%	0,081; 0,088	0,080; 0,086	0,086; 0,092	0,083; 0,089
He_IC95%	0,155; 0,166	0,155; 0,164	0,162; 0,172	0,164; 0,173
FIS_IC95%	0,326; 0,358	0,353; 0,380	0,341; 0,371	0,384; 0,409
Número de Alelos	9.934	10.569	10.415	11.272
Riqueza Alélica	7.870	7.951	7.973	8.038
Alelos Privados	3	37	22	115
APR	110	100	163	200
UCS/Sistema	FLONA Tapajós Flo	FLONA Tapajós Qui	FLONA Tefé Flo	FLONA Tefé Qui
Ho	0,087	0,085	0,077	0,081
He	0,178	0,171	0,162	0,161
Fis	0,415	0,381	0,412	0,399
Ho_IC95%	0,084; 0,090	0,082; 0,088	0,074; 0,080	0,078; 0,084
He_IC95%	0,174; 0,182	0,167; 0,176	0,158; 0,167	0,157; 0,165
FIS_IC95%	0,40; 0,42	0,36; 0,39	0,39; 0,42	0,38; 0,41
Número de Alelos	12.136	10.696	10.716	11.273
Riqueza Alélica	8.115	8.017	7.960	7.989
Alelos Privados	181	21	23	107
APR	217	196	131	167

Foram estimados parâmetros de diversidade genética nos sistemas de quintal e floresta em cada unidade de conservação (UC). Os resultados indicam que a FLONA Tapajós apresentou os maiores valores de riqueza alélica (8.115) e de alelos privados rarefeitos (217) no sistema de floresta entre as UCs. De modo geral, observou-se pequena diferença entre os sistemas de quintal e floresta, destacando-se, contudo, um maior número de alelos privados rarefeitos (196) no sistema de quintal da FLONA Tapajós (Tabela 4).

Quando considerados apenas os sistemas de plantio (floresta e quintal), os dados de SNPs indicaram elevada similaridade entre os sistemas. Ainda assim, alguns parâmetros, como a riqueza alélica (12.333) e o número de alelos privados rarefeitos (442), foram ligeiramente maiores no sistema de floresta, enquanto o sistema de quintal apresentou valores de 12.310 e 435, respectivamente.

A análise dos intervalos de confiança da heterozigosidade esperada, heterozigosidade observada e do coeficiente de endogamia revelou sobreposição entre os valores estimados, reforçando a similaridade genética entre os sistemas (Tabela 4).

Tabela 4. Estimativas de diversidade genética genômica para 6.900 marcadores SNPs avaliados em 207 indivíduos de *Caryocar villosum* em quatro unidades de conservação, considerando os sistemas de quintal e floresta. Número de indivíduos (N), heterozigosidade observada (Ho), heterozigosidade esperada (He), coeficiente de endogamia (FIS), significativo com intervalo de confiança de 95% de Ho (Ho_IC95%), significativo com intervalo de confiança de 95% de He (He_IC95%), significativo com intervalo de confiança de 95% de FIS (FIS_IC95%), número de alelos (NA), riqueza alélica (RA), alelos privados (Ap), alelos privados rarefeitos (APR). Os números destacados em negrito indicam diferenças estatisticamente significativas.

UCs/Sistema	Floresta	Quintal
Ho	0,086	0,084
He	0,181	0,179
Fis	0,441	0,439
Ho_IC95%	0,083; 0,088	0,082; 0,087
He_IC95%	0,178; 0,185	0,176; 0,183
FIS_IC95%	0,432; 0,451	0,430; 0,448
Número de Alelos	13.363	13.439
Riqueza Alélica	12.333	12.310
Alelos Privados	361	437
APR	442	435

5.2 Estrutura populacional

A análise de variância molecular (AMOVA), realizada com 207 indivíduos das quatro unidades de conservação (RDS Amanã, FLONA Tefé, FLONA Caxiuanã e FLONA Tapajós), revelou maior variação genética entre os sistemas de plantio (quintal e floresta), correspondendo a aproximadamente 11%, com valor de p altamente significativo ($p < 0,001$). Também foi observada uma divergência genética de 8,9% entre as unidades de conservação.

Os valores de F_{ST} foram baixos (inferiores a 0,12), porém estatisticamente significativos, indicando que a maior parte da diversidade genética está distribuída dentro das subpopulações, ou seja, entre sistemas de plantio e indivíduos (Tabela 5).

Tabela 5. Análise de variância molecular (AMOVA) baseada a partir da média de 6.900 loci SNPs encontrados em 207 amostras de *Caryocar villosum* em quatro unidades de conservação da Amazônia Brasileira, e, dentro dos sistemas quintal e floresta. Valores em negrito indicam as estatísticas correspondentes a diferenciação entre grupos. Os números destacados em negrito indicam diferenças estatisticamente significativas.

Fonte de variação	SSD (Soma dos quadrados)	MSD (Média quadrática)	Porcentagem de variação
Entre UCs	14.121	4.707	8,9%
Entre sistemas	4.275	1.068	2,1%
Entre indivíduos	138.157	694	88,9%
Total	156.553	759	

A AMOVA par a par indicou que a maior porcentagem de variação genética ocorreu entre a RDS Amanã e a FLONA Caxiuanã (15,10%), seguida pela comparação entre a FLONA Caxiuanã e a FLONA Tefé (14,90%). Por outro lado, as menores porcentagens de variação foram observadas entre a RDS Amanã e a FLONA Tefé (3,10%) e entre a FLONA Caxiuanã e a FLONA Tapajós (5,60%) (Tabela 6).

Tabela 6. Análise de variância molecular (AMOVA) em *Caryocar villosum* par a par, comparando as variações de duas unidades de conservação por vez. Valores em negrito indicam as estatísticas correspondentes a diferenciação entre grupos. Os números destacados em negrito indicam diferenças estatisticamente significativas par a par.

Fonte de variação	Soma dos quadrados	Média quadrática	Porcentagem de variação	Estatística F
AMA/CAX	5.584	5.584	15,1%	0,151
AMA/TAP	4.302	4.302	10,5%	0,105
TAP/TEF	5.775	5.775	10,4%	0,104
AMA/TEF	1.537	1.537	3,1%	0,031
CAX/TAP	3.226	3.226	5,6%	0,056
CAX/TEF	7.369	7.369	14,9%	0,149

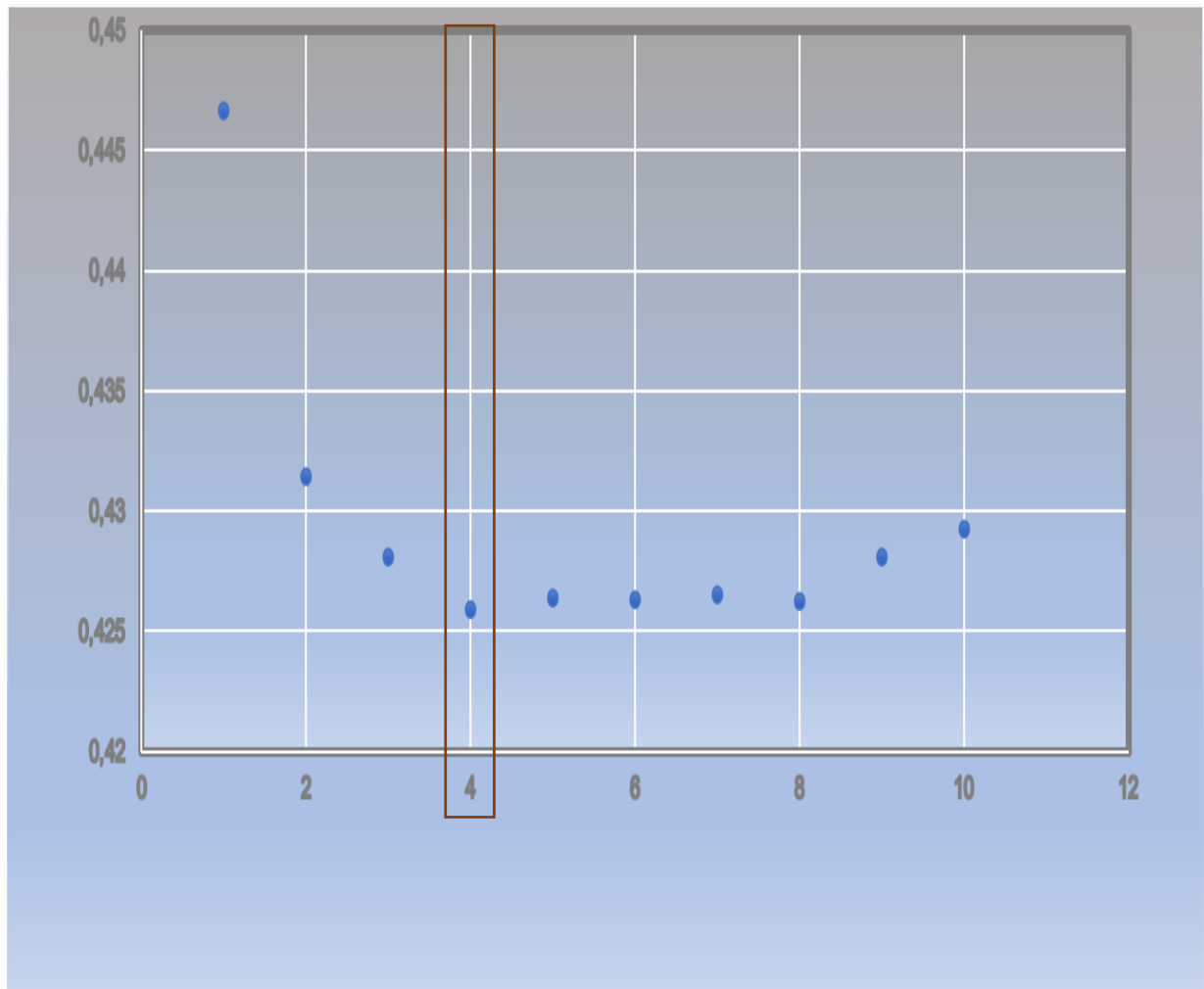
Quando realizada a AMOVA utilizando todos os indivíduos dos sistemas de quintal e floresta de cada unidade de conservação, observou-se que a maior variação ocorreu nos sistemas da FLONA Caxiuanã (6,30%), enquanto a menor ocorreu nos sistemas da RDS Amanã (Tabela 7).

Para a análise da estrutura populacional, foi estimada a quantidade ideal de K a partir da entropia cruzada. Foram avaliados valores de K de 0 a 10, sendo que o menor valor de entropia cruzada indicou K = 4 como o mais adequado, conforme mostra a Figura 2.

Tabela 7. Análise de variância molecular (AMOVA) em *Caryocar villosum* dentro das unidades de conservação. Valores em negrito indicam as estatísticas correspondentes a diferenciação entre os sistemas. Os números destacados em negrito indicam diferenças estatisticamente significativas.

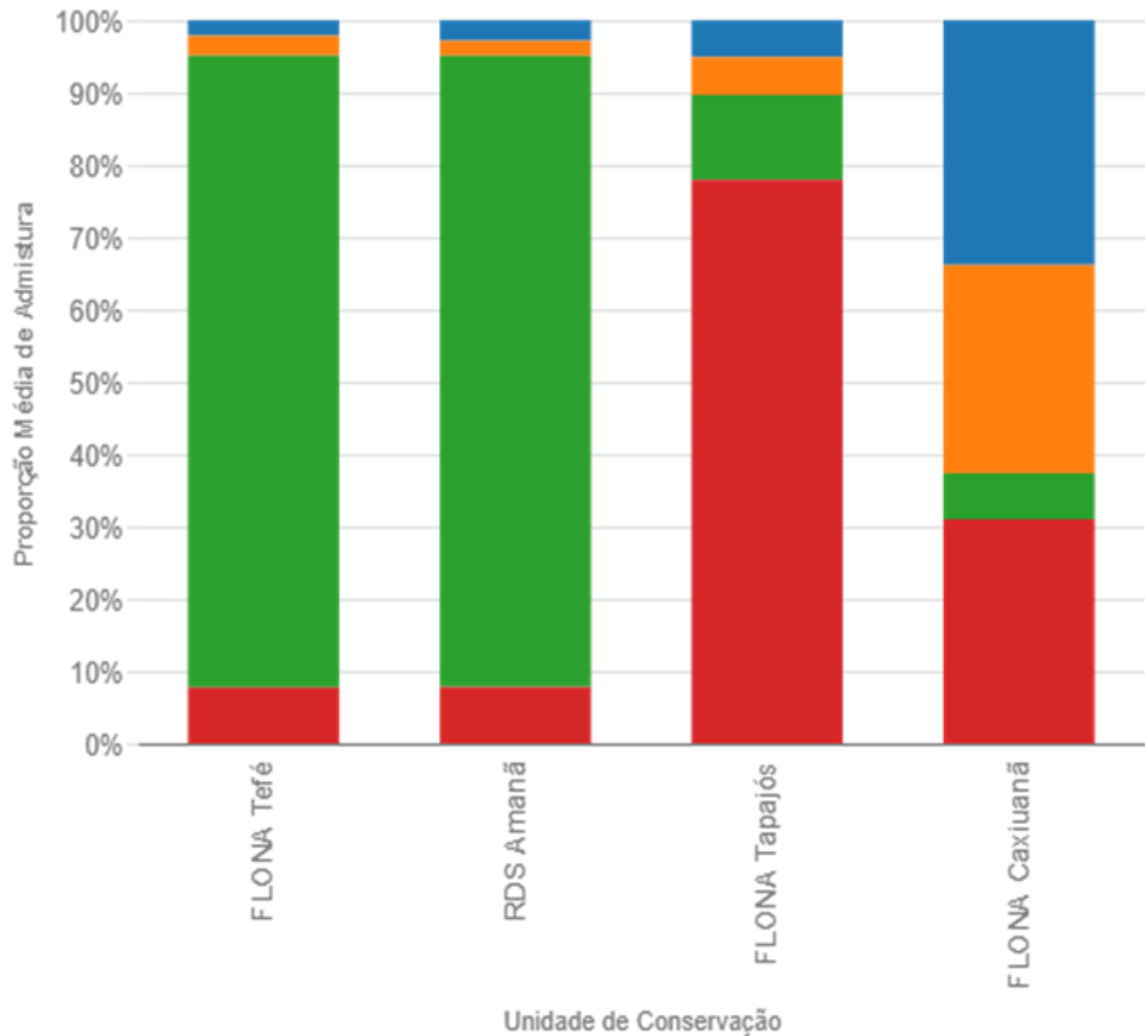
Dentro Sistema	Soma dos quadrados	Média quadrática	Porcentagem de variação	Estatística F
RDS Amanã	614	614	-0,4%	-0,004
FLONA Tefé	939	939	1,4%	0,014
FLONA Caxiuanã	1.779	1.779	6,3%	0,063
FLONA Tapajós	943	943	1,1%	0,011

Figura 2. Matriz de 207 indivíduos e 6.900 SNPs e testados valores de K de 1 a 10 populações ancestrais. A barra retangular em vermelho, indica o N° de K=4 como ideal como melhor resultado do modelo.



No sNMF, realizado nas quatro unidades de conservação com a presença de indivíduos de quintal e floresta, o número ideal de grupos foi $K = 4$, havendo estruturação em pelo menos três unidades de conservação. Duas unidades de conservação (RDS Amanã e FLONA Tefé) apresentaram-se como uma extensão uma da outra. Dentro dos quatro grupos genéticos, três são distintos, havendo estruturação (Figura 3).

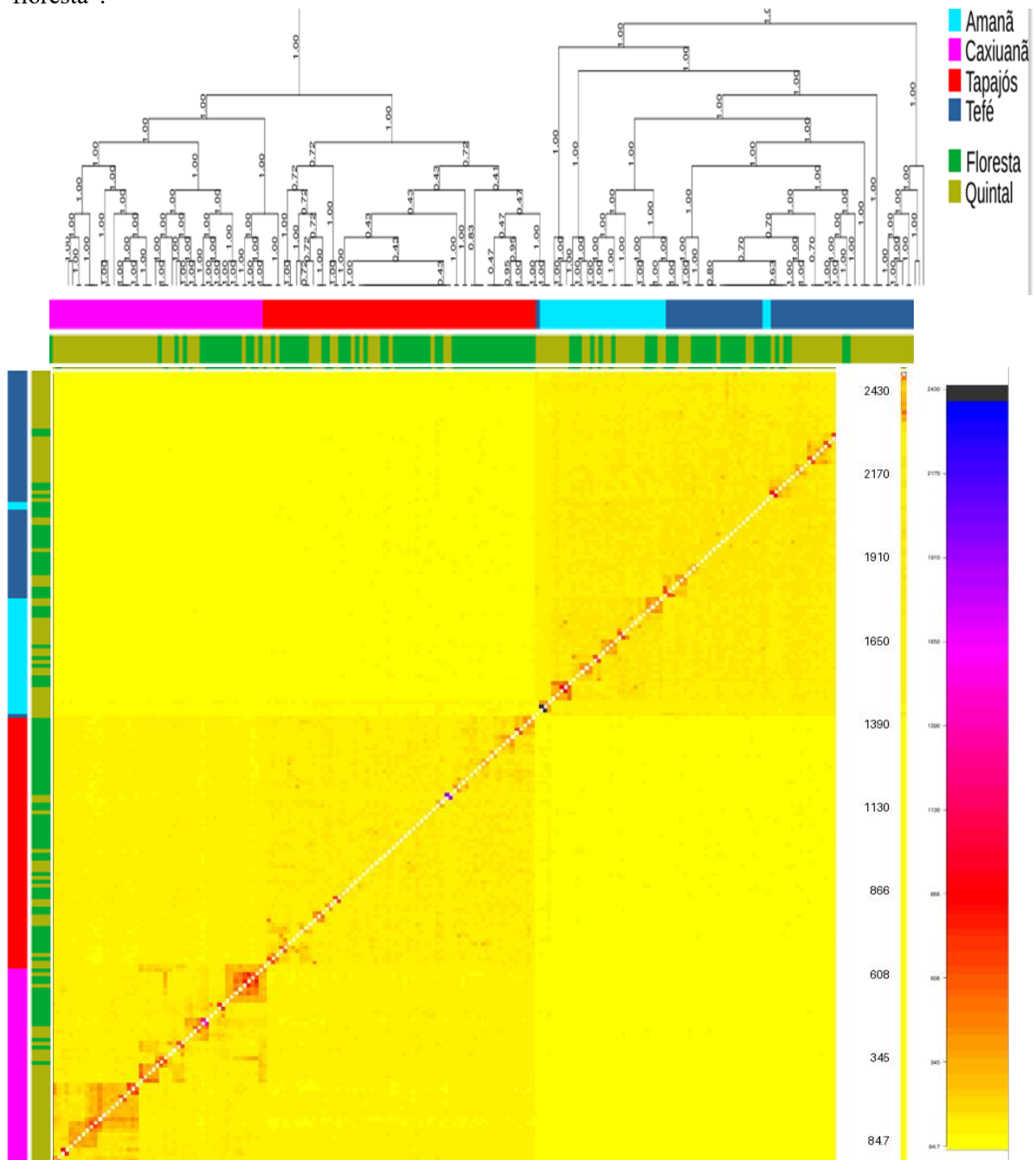
Figura 3. Valores de entropia cruzada e matriz de ancestralidade com o K=4 sendo a melhor distribuição, as cores definem a presença em porcentagem na formação alélica de cada população.



Foi realizada uma matriz de semelhança entre indivíduos em análise utilizando o software fineRADstructure. A matriz é simétrica, e a diagonal principal (de cima à esquerda para baixo à direita) representa a semelhança de um indivíduo consigo mesmo (Figura 4).

Os indivíduos da FLONA Caxiuanã e da FLONA Tapajós formam blocos distintos em relação aos indivíduos da RDS Amanã e da FLONA Tefé na matriz, agrupando-se também no dendrograma. Os grupos Floresta (verde escuro) e Quintal (verde musgo) também se agrupam de forma coesa, apresentando maior relação entre si (Figura 4).

Figura 4. Matriz de coancestralidade entre os indivíduos das quatro unidades de conservação, azul escuro (FLONA Tefé), azul claro (RDS Amanã), vermelho (FLONA Tapajós) e rosa (FLONA Caxiuanã). As cores verde musgo, refere-se ao sistema de plantio quintal e verde escuro ao sistema de plantio “floresta”.



5.3 Migração e Fluxo Gênico

A FLONA Tapajós é o destino mais relevante em termos de impacto na taxa de substituição populacional por migração, especialmente no fluxo RDS Amanã–FLONA Tapajós (1,009). A estimativa do parâmetro central (média/mediana/moda) é altamente consistente para todas as rotas, sugerindo distribuições de probabilidade geralmente simétricas.

O fluxo RDS Amanã–FLONA Tefé é o que apresenta maior incerteza (maior dispersão na estimativa do parâmetro). Já o fluxo FLONA Tefé–RDS Amanã apresenta a menor incerteza (0,940). O parâmetro de dispersão do modelo (theta destino) é praticamente constante em todos os fluxos migratórios (Tabela 8).

Tabela 8. Estimativas de fluxo gênico entre as unidades de conservação com o MIGRATE.

Origem	Destino	média	theta destino	migrantes/geração
Ama	Cax	84,375	0,046	0,972
Ama	Tap	86,264	0,047	1,009
Ama	Tef	83,849	0,046	0,967
Cax	Ama	87,783	0,044	0,955
Cax	Tap	83,171	0,047	0,973
Cax	Tef	84,131	0,046	0,971
Tap	Ama	86,600	0,044	0,942
Tap	Cax	84,896	0,046	0,978
Tap	Tef	84,182	0,046	0,971
Tef	Ama	86,416	0,044	0,940
Tef	Cax	83,095	0,046	0,957
Tef	Tap	81,953	0,047	0,959

A FLONA Tapajós apresenta os valores mais altos para o centro da distribuição (quantis, moda, mediana e média) em comparação com as demais unidades, com valores em torno de 0,047. Já a RDS Amanã apresenta os valores mais baixos, entre 0,043 e 0,044 (Tabela 9). No entanto os valores são estatisticamente bem similares, indicando que não há diferença entre as unidades de conservação.

A FLONA Caxiuanã e a FLONA Tefé apresentam valores de centro de distribuição muito próximos, na faixa de 0,046 (Tabela 9).

A RDS Amanã é a única população com razão menor que 1 (0,963), o que indica que o número médio de emigrantes por geração é maior do que o de imigrantes. A FLONA Caxiuanã, a FLONA Tapajós e a FLONA Tefé apresentam razões próximas ou superiores a 1, indicando que o número médio de imigrantes é igual ou ligeiramente superior ao de emigrantes por geração.

A FLONA Tefé apresenta a maior razão (1,019), indicando a maior diferença relativa entre imigrantes e emigrantes a favor dos imigrantes. A FLONA Tapajós apresenta a maior taxa média de imigrantes por geração (0,980), enquanto a RDS Amanã apresenta a maior taxa média de emigrantes por geração (0,983) (Tabela 9).

Tabela 9. Estimativas de fluxo gênico entre as unidades de conservação com o MIGRATE, centro de distribuição.

População	Quantil 2.5%	Quantil 97.5%	moda	mediana	média
Ama	0,043	0,044	0,044	0,0436	0,0435
Cax	0,046	0,046	0,046	0,0464	0,0461
Tap	0,047	0,047	0,047	0,0468	0,0468
Tef	0,046	0,046	0,046	0,0462	0,0462
MÉDIA MIGRANTES / GERAÇÃO					
Pop	Imigrantes		Emigrantes		Razão
Ama	0,946		0,983		0,963
Cax	0,969		0,966		1,002
Tap	0,980		0,964		1,017
Tef	0,970		0,952		1,019

6. DISCUSSÃO

6.1 Diversidade Genética

A utilização de uma matriz robusta de 6.900 SNPs proporcionou uma resolução genômica inédita para o piquiá, superando limitações comuns de marcadores SSR, como alelos nulos e homozigosidade superestimada. Essa abordagem permitiu identificar que, apesar da seleção humana, a estrutura genética da espécie permanece preservada em escala regional, com divergências moderadas entre as Unidades de Conservação.

O estudo revelou que os valores nos parâmetros de diversidade genética entre sistemas de quintal (manejados) e floresta madura são bastante similares. Diferentemente de outras culturas estudadas por Nascimento *et al.*, (2024), que encontraram assinaturas de seleção intensas em populações de açaí, não foi sustentada a hipótese de um gargalo genético acentuado nos quintais, indicando uma domesticação incipiente e recente, na qual a divergência genotípica ainda é pequena.

A alta densidade de marcadores SNPs utilizada neste estudo superou as limitações de homozigosidade superestimada comuns em genotipagens com microssatélites, que frequentemente enfrentam problemas com alelos nulos. Enquanto Francisconi *et al.* (2021) utilizaram marcadores microssatélites para investigar populações de piquiá, as estimativas podem ter sido influenciadas pela presença de alelos nulos.

O estudo demonstrou que, em todas as quatro UCs e em todos os sistemas de plantio (quintal e floresta), a heterozigosidade observada foi menor que a heterozigosidade esperada. Isso resultou em um coeficiente de endogamia (FIS) positivo em todas as UCs, com valores significativos, alguns deles acima de 0,42. Um FIS positivo e alto (acima de 0,4) indica um excesso de homozigotos em relação ao esperado pelo equilíbrio de Hardy-Weinberg. Isso sugere a ocorrência de cruzamento entre indivíduos geneticamente próximos ou mistura entre subpopulações (Ridley, 2006, p. 159). Colli-silva *et al.* (2023), observaram que há redução dos valores da diversidade genética em populações domesticadas de cupuaçu em comparação com populações selvagens da espécie.

Embora a endogamia seja esperada em populações domesticadas devido a gargalos genéticos e à seleção humana (Francisconi *et al.*, 2021; Gepts, 2004), também foi observado que populações de floresta apresentam altos coeficientes de endogamia. O cruzamento preferencial (acasalamento não aleatório) ou a estrutura populacional fina, mesmo em florestas maduras, podem levar a um histórico de cruzamento entre parentes próximos (Loveless; Hamrick, 1984).

Os parâmetros de diversidade genética são similares entre as populações de quintal (manejadas) e as de floresta madura, reforçando o estudo de Francisconi *et al.* (2021), que indica que a seleção humana ainda não causou um gargalo genético acentuado ou uma diferenciação profunda entre os grupos. No entanto, de acordo com Kurunes *et al.* (2021), a baixa diversidade genética entre as populações pode estar associada a outros fatores, como o efeito Wahlund ou a mistura entre subpopulações.

Foi encontrada pouca diferença entre as diversidades genômicas das populações, com alguns parâmetros bastante similares. Resultados semelhantes foram encontrados por Alves (2023), que estudou populações de *Caryocar villosum* pertencentes à Amazônia Ocidental e Oriental brasileira, e por Franciscone *et al.* (2021), que observaram valores nos parâmetros de diversidade bastante similares entre duas populações de piquiá.

6.2 Comparação Quintal vs. Floresta e o Processo de Domesticação

Diferentemente do observado em culturas de ciclo curto, nas quais a domesticação gera gargalos genéticos severos, o piquiá apresenta índices de endogamia elevados ($FIS > 0,40$) tanto em quintais quanto em florestas maduras. Nascimento *et al.* (2024) encontraram padrões semelhantes em estudos com o açaí (*Euterpe oleracea*), nos quais a biologia reprodutiva e o cruzamento entre indivíduos aparentados (endogamia natural), bem como a estrutura populacional fina em escala local, exercem papel fundamental na configuração genética, independentemente da intensidade do manejo atual.

A FLONA Tapajós foi a unidade de conservação com maior riqueza alélica no sistema de floresta, demonstrando maior profundidade temporal na diversidade natural e maior número de alelos privados rarefeitos, refletindo elevada diversidade genética entre os sistemas (Tabela 2), o que a posiciona como um possível centro de diversidade genética e reforça a hipótese de ser o centro de origem da domesticação da espécie. Este achado está alinhado com o estudo de Francisconi *et al.* (2021), que também encontraram maior diversidade genética na FLONA Tapajós, e com o estudo de Alves (2023), que estimou essa região como centro de origem.

A hipótese levantada neste estudo, de que haveria menor diversidade genética nos quintais do que nas florestas maduras, não foi sustentada. As estimativas de diversidade entre os sistemas de quintal e floresta foram muito similares, com sobreposição dos intervalos de confiança (IC95%), e, embora o sistema de floresta tenha apresentado maior riqueza alélica e maior número de alelos privados rarefeitos, o sistema de quintal apresentou maior número de alelos e de alelos privados. Essa distribuição mista de alelos sugere um fluxo gênico contínuo entre os sistemas ou um processo de domesticação recente e incipiente (Clement, 1999a), no

qual a divergência genotípica ainda é pequena. O padrão observado de diversidade genética elevada e coeficientes de endogamia semelhantes em ambos os sistemas (quintal e floresta) é consistente com a descrição do piquiá como uma espécie incipientemente domesticada (Alves *et al.*, 2016; Francisconi *et al.*, 2021). Em populações incipientemente domesticadas, espera-se uma divergência genotípica pequena, especialmente quando o fluxo gênico entre populações manejadas e silvestres é mantido.

A endogamia encontrada nos quintais (com exceção de algumas áreas) foi mais elevada na RDS Amanã e na FLONA Caxiuanã, onde o coeficiente de endogamia foi maior nos sistemas de quintal (Tabela 3). Esse padrão pode ser explicado pelo manejo humano, com a coleta de um número limitado de indivíduos e a propagação seletiva de linhagens de interesse, resultando em um estreitamento da base genética (Weber *et al.*, 2001; Hollingsworth *et al.*, 2005).

6.3 Estrutura genética e variações populacionais

As análises de estrutura populacional (sNMF e fineRADstructure) e a AMOVA fornecem percepções sobre a estruturação genética e variância entre as unidades de conservação, bem como sobre o potencial centro de origem da domesticação. A AMOVA mostrou que a maior parte da diversidade genética está contida dentro das subpopulações/indivíduos (11%), embora os valores sejam estatisticamente significativos. A variação entre as UCs foi de apenas 9%, enquanto a variação entre os sistemas (quintal e floresta) foi de 11%. Valores similares foram encontrados por Franciscone *et al.* (2021), com a utilização de marcadores microssatélites, onde as fontes de variação estavam principalmente entre os indivíduos (97,63%) e apenas 2,36% estavam entre os sistemas de plantio. Já Muhlen *et al.* (2019), estudou populações cultivadas e selvagens de mandioca, e encontrou maior parte da variação está dentro dos grupos com F_{ST} de 0,081.

O sNMF (com $K = 4$) revelou pouca ou nenhuma estruturação entre a RDS Amanã e a FLONA Tefé, que são quase uma extensão uma da outra, quando olhamos para o panorama geral das populações, notou-se que existe três estruturas bem definidas (figura 3). Diferente do que encontrados em populações domesticadas de mandioca, onde se encontrou estruturas bem definidas quando comparadas com populações selvagens (Pereira-Alves *et al.*, 2018). A matriz de semelhança (fineRADstructure) mostrou que a FLONA Caxiuanã e a FLONA Tapajós formam blocos distintos e se agrupam separadamente da RDS Amanã e da FLONA Tefé. Isso sugere uma estrutura genética em escala regional, na qual a distância geográfica exerce alguma influência, mas não é um fator absoluto. Os resultados encontrados no

FineRADstructure, encontrando dois grupos distintos, corroboram com o estudo realizado por Alves (2023), onde se analisou populações de piquiá considerando a Amazônia Ocidental e Oriental brasileira, e encontrou-se duas estruturas distintas, uma formada pela Amazônia Ocidental e outra pela Amazônia Oriental.

A FLONA Tapajós foi a unidade de conservação com maior número de alelos, riqueza alélica e alelos privados rarefeitos (Tabela 2), o que a posiciona como um possível centro de diversidade genética e reforça a hipótese de ser o centro de origem da domesticação da espécie. Este achado está alinhado com o estudo de Francisconi *et al.* (2021), que também encontraram maior diversidade genética na FLONA Tapajós, e com o estudo de Alves (2023), que estimou essa região como centro de origem.

Estudos de Moreira *et al.* (2015) e Pedrosa *et al.* (2018) mostram que a domesticação em espécies de uvas alterou drasticamente o fenótipo (tamanho de fruto e pigmentação). No piquiá, os SNPs revelam que a estrutura genética ainda está mais associada à geografia regional, sugerindo que a seleção humana focada em frutos não amargos ainda não isolou completamente essas populações.

6.4 Fluxo Gênico e Padrões de Migração

A hipótese de que haveria estruturação genética e não teria fluxo gênico entre as populações das quatro unidades de conservação, não foi sustentada. Houve pouca estruturação nas populações como demonstrado nas análises do SNMf e fineRADstructure. Os resultados do MIGRATE mostram que há fluxo gênico entre as unidades de conservação. A identificação da FLONA Tapajós como o principal destino de imigrantes e detentora do maior número de alelos privados rarefeitos (505) corrobora a hipótese de ser um centro de diversidade e de origem da domesticação. Esse achado está em consonância com as evidências arqueológicas de agregações próximas às Terras Pretas de Índio na região, encontradas por Alves *et al.* (2016) e Franciscone *et al.* (2021). Assim como a mandioca (Alves-Pereira *et al.*, 2020) teve sua dispersão mapeada por SNPs a partir do sudoeste amazônico, o piquiá parece ter na calha do Tapajós o seu núcleo de manejo histórico, de onde a diversidade genética foi exportada e mantida ao longo do tempo.

As estimativas de fluxo gênico (MIGRATE) entre as UCs fornecem a direção e a intensidade da migração. O fluxo entre a RDS Amanã e a FLONA Tapajós (Ama–Tap) é o mais relevante em termos de impacto na taxa de substituição populacional por migração, com uma taxa de 1,009 migrantes por geração. A razão entre imigrantes e emigrantes reforça que a RDS Amanã é a única população com valor inferior a 1 (0,963), indicando maior número de

emigrantes do que de imigrantes. Além disso, a RDS Amanã apresenta a maior taxa média de emigrantes por geração (0,983). Por sua vez, a FLONA Tapajós e a FLONA Tefé apresentam razões superiores a 1 (1,017 e 1,019, respectivamente), indicando que o número médio de imigrantes é ligeiramente superior ao de emigrantes.

A alta taxa de imigrantes na FLONA Tapajós, combinada com sua elevada diversidade genética, sugere que essa unidade de conservação está recebendo e retendo material genético de outras populações, o que pode estar associado à sua importância como centro de domesticação. O fluxo gênico, de modo geral, com razão próxima a 1 (exceto na RDS Amanã), sugere um movimento de genes relativamente equilibrado entre as UCs, o que pode ser facilitado por polinizadores de longa distância, como morcegos do gênero *Phyllostomus* (Martins; Gribel, 2007), e pela dispersão humana de sementes (Alves *et al.*, 2016).

7. CONCLUSÃO

A utilização de marcadores moleculares do tipo SNP representa um salto metodológico em relação aos tradicionais microssatélites (SSR), permitindo uma cobertura homogênea de todo o genoma e maior precisão na detecção de variantes genéticas.

Portanto, com base nos resultados obtidos e nos objetivos apresentados neste estudo, é possível concluir que a diversidade genética em piquiazeiros de quintais e de floresta compartilha parâmetros de diversidade bastante semelhantes, reforçando a incipiência da domesticação da espécie. Existe pouca estrutura genética entre as quatro unidades de conservação, tornando as populações estruturalmente semelhantes, não havendo padrões bem definidos que indiquem o principal centro de domesticação da espécie.

Caryocar villosum na Amazônia apresenta um processo de domesticação incipiente em curso. Embora o manejo humano não tenha gerado uma forte distinção na diversidade genética entre quintais e florestas (devido, provavelmente, ao fluxo gênico contínuo e à natureza incipiente do processo), há evidências claras de endogamia em todas as populações e um forte indício de que a FLONA Tapajós (com a maior diversidade alélica e maior recebimento de imigrantes) seja um centro de origem do processo de domesticação.

8. REFERÊNCIAS BIBLIOGRÁFICAS

- Alves, R.P. Conhecimento Ecológico Tradicional e Domesticação de Árvores na Amazônia Brasileira. Tese de doutorado em Ecologia, 117p. Instituto Nacional de Pesquisas da Amazônia, Amazonas, 2023.
- Alves R.P., Levis C., Clement C.R. Use and Management of Piquiá Suggest in situ Domestication along the Lower Tapajós River, Brazilian Amazonia. **Economic Botany**, New York, v.70, n.2, p. 199–201, 2016.
- Baird, N.A, Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A. Rapid SNP discovery and genetic mapping using sequenced RAD markers. **PLoS ONE**, San Francisco, v. 3, 2008.
- Beerli, P. (2006). Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, 22(3), 341–345.
- Beerli, P. (2009). How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In G. Bertorelle et al. (Eds.), *Population Genetics for Animal Conservation* (pp. 42–79). Cambridge University Press.
- Beerli, P., & Palczewski, M. (2010). Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics*, 185(1), 313–326.
- Bentley, D.R. et al. Accurate whole human genome sequencing using reversible terminator chemistry. **Nature**, London, v. 456, p. 53-59, 2008.
- Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodriguez, M., Biagetti, E., Santilocchi, R., Spagnoletti Z. P., Gioia, T., Logozzo, G., Attene, G., Nanni, L., Papa, R. Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. **New Phytologist**, v. 197, p. 300–313, 2013.
- Brumfield, R.T., Beerli, B., Nickerson, D.A., Edwards, S.V. The utility of single nucleotide polymorphisms in inferences of population history. **Trends in Ecology and Evolution**, Cambridge, v. 18, p. 246-256, 2003.
- Cao, K., Zheng, Z., Wang, L., Liu, X., Zhu, G., Fang, W., Cheng, S., Zeng, P., Chen, C., Wang, X., Xie, M., Zhong, X., Wang, X., Zhao, P., Bian, C., Zhu, Y., Zhang, J., Ma, G., Chen, C., Li, Y., Hao, F., Li, Y., Huang, G., Li, Y., Li, H., Guo, J., Xu, X., Wang, J. Comparative population genomics reveals the domestication history of the peach, *Prunus persica*, and human influences on perennial fruit crops. **Genome Biology**, v. 15, p. 1–15, 2014.
- Casas, C., Diez, F., Gopinath, G. and Gourinchas. Dominant currency paradigm. **Working paper**, 2016.
- Clement, C.R. 1492 and the loss of Amazonian crop genetic resources. I. The relation between domestication and human population decline. **Economic Botany**, New York, v.53, n.2, p. 189-196, 1999a.
- Clement, C.R. 1492 and the loss of Amazonian crop genetic resources. II. Biogeography of culture in contact. *Botânica Econômica*. **Economic Botany**, New York, v.53, n.2, p. 203-216, 1999b.
- Clement, C.R. Piquiá (*Caryocar vilosum*, *Cariocaraceae*). Páginas 100-109 em J. W. Clay, P. T. B. Sampaio e C. R. Clement, eds, *Biodiversidade Amazônica: Exemplos e estratégias de utilização*. Sebrae-AM, Programa de Desenvolvimento Empresarial e Tecnológico, Manaus, p. 100-199c.
- Clement, C.R., Cristo-Araújo, M., Coppens D'eeckenbrugge, G., Pereira, A.A., Picanço, R.D. Origin and domestication of native Amazonian crops. **Diversity**, Basel, v. 2, p. 72-106, 2010.

- CLEMENT, C. R. et al. Origin and Domestication of Native Amazonian Crops. **Diversity**, S. 1, v. 13, n. 6, p. 223, 2021.
- Coelho, S.D., Levis, C., Baccaro, F.B., Figueiredo, F.O.G., Antunes, A.P., ter Seege, H., Peña-Claros, M., Clement, C.R., Schietti, J. Eighty-four per cent of all Amazonian arboreal plant individuals are useful to humans. **Plos One**, v. 16, p. 1-17, 2021.
- Collevatti, R.G., Leoi, L.C.T., Leite, S.A., Gribel, R. Contrasting patterns of genetic structure in Caryocar (Caryocaraceae) congeners from flooded and upland Amazonian forests. **Biological Journal of the Linnean Society**, v.98, p.278-290, 2009.
- Collins, F.S., Brooks, L.D., Chakravarti, A. A DNA polymorphism discovery resource for research on human genetic variation. **Genome Research**, New York, v. 8, p. 1229-1231, 1998.
- Darwin, C. On the origin of species by Charles Darwin: A fac simile of the first edition with an introduction by Ernst Mayr. Harvard University Press, Cambridge, Massachusetts, USA, 1859.
- Ehrendorfer, F.; Morawetz, W.; Dawe, J. The neotropical angiosperm families Brunelliaceae and Caryocaraceae: First karyosystematical data and affinities. **Plant Systematics and Evolution**, v. 145, p. 183–191, 1984.
- EMBRAPA RONDÔNIA. Diversidade Genética Via Marcadores Molecular SNP Em Cafeeiros Amazônicos. **Boletim de Pesquisa e Desenvolvimento**, Porto Velho, n. 182, 2023.
- Ferreira, M.E., Grattapaglia, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3.ed. Brasília: Embrapa Cenargen. p. 22, 1998.
- Francisconi, A.F., Alves, R.P., Clement, C.R., Igor, G.D; Carvalho, A.S., Veasey, E.A. Genetic structure and diversity identify incipient domestication of Piquia ' Caryocar villosum (Aubl.) pers. along the lower Tapajo 's River, Brazilian Amazonia. **Genetic Resources Crop Evolution**, v. 68, p. 1487-1501, 2021.
- Freeland, J.R. **Molecular Ecology**. Hoboken: Wiley-Blackwell, 2005. p. 400.
- Gepts, P. Crop domestication as a long-term selection experiment. **Plant Breeding Reviews**, v. 24, p. 1 – 44, 2004.
- Goldstein, D.B.; Pollock, D.D. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity*, Oxford, v. 88, p. 335- 342, 1997.
- Gribel R, Hay JD. Ecologia da polinização de Cariocar brasiliense (Caryocaraceae) na vegetação de cerrado do Brasil Central. **Revista de Ecologia Tropical**, v.9, p.199–211. 1993
- Harlan, J.R. Crops and Man. American Society of Agronomy & Crop Science Society of America, Madison, WI, USA, 1992.
- Hollingsworth, P. M., Dawson, I. K., Goodall-Copestake, W. P., Richardson, J. E., Weber, J. C., Montes, C. S., Pennington, R. T. Do farmers reduce genetic diversity when they domesticate tropical trees? A case study from Amazonia. **Molecular Ecology**, v. 14, p. 497–501, 2005.
- Lane, E. V. Piqui-a -- Potential Source of Vegetable Oil for an Oil-starving IVorld. **Economic Botany**, v. 11, p. 187–207, 1957.
- Lengkeek, A. G., Jaenicke, H., Dawson, I. K. Genetic bottlenecks in agroforestry systems: Results of tree nursery surveys in East Africa. **Agroforestry Systems**, v. 63, p. 149–155, 2005.
- Levis, C., Costa, F.R.C., Bongers, F., Peña-Claros, M., Clement, C.R., Junqueira, A.B., Neves, E.G., Tamanaha, E.K., Figueiredo, F.O.G., Salomão, R.P., Castilho, C.V., Magnusson, W.E., Phillips, O.L.,

Guevara, J.E., Sabatier, D., Molino, J.F., Cárdenas López, D., Mendoza, A.M., Pitman, N.C.A., Duque, A., Núñez Vargas, P., Zartman, C.E., Vasquez, R., Andrade, A., Camargo, J.L., Feldpausch, T.R., Laurance, S.G.W., Laurance, W.F., Killeen, T.J., Mendonça Nascimento, H.E., Montero, J.C., Mostacedo, B., Amaral, I.L., Guimarães Vieira, I.C., Brien, R., Castellanos, H., Terborgh, J., de Jesus Veiga Carim, M., da Silva Guimarães, J.R., de Souza Coelho, L., de Almeida Matos, F.D., Wittmann, F., Mogollón, H.F., Damasco, G., Dávila, N., García-Villacorta, R., Coronado, E.N.H., Emilio, T., de Andrade Lima Filho, D., Schiatti, J., Souza, P., Targhetta, N., Comiskey, J.A., Marimon, B.S., Marimon Jr., B.-H., Neill, D., Alonso, A., Arroyo, L., Carvalho, F.A., de Souza, F.C., Dallmeier, F., Pansonato, M.P., Duivenvoorden, J.F., A. Fine, P.V., Stevenson, P.R., Araujo-Murakami, A., Aymard C., G.A., Baraloto, C., do Amaral, D.D., Engel, J., Henkel, T.W., Maas, P., Petronelli, P., Cardenas Revilla, J.D., Stropp, J., Daly, D., Gribel, R., Ríos Paredes, M., Silveira, M., Thomas-Caesar, R., Baker, T.R., da Silva, N.F., Ferreira, L.V., Peres, C.A., Silman, M.R., Cerón, C., Valverde, F.C., Di Fiore, A., Jimenez, E.M., PeñuelaMora, M.C., Toledo, M., Barbosa, E.M., deMatos Bonates, L.C., Arboleda, N.C., de Sousa Farias, E., Fuentes, A., Guillaumet, J.-L., Møller Jørgensen, P., Malhi, Y., de Andrade Miranda, I.P., Phillips, J.F., Prieto, A., Rudas, A., Ruschel, A.R., Silva, N., von Hildebrand, P., Vos, V.A., Zent, E.L., Zent, S., Cintra, B.B.L., Nascimento, M.T., Oliveira, A.A., Ramirez-Angulo, H., Ramos, J.F., Rivas, G., Schöngart, J., Sierra, R., 67 Tirado, M., van der Heijden, G., Torre, E.V., Wang, O., Young, K.R., Baider, C., Cano, A., Farfan-Rios, W., Ferreira, C., Hoffman, B., Mendoza, C., Mesones, I., Torres Lezama, A., Medina, M.N.U., van Andel, T.R., Villarroel, D., Zagt, R., Alexiades, M.N., Balslev, H., Garcia-Cabrera, K., Gonzales, T., Hernandez, L., Huamantupa Chuquimaco, I., Manzatto, A.G., Milliken, W., Cuenca, W.P., Pansini, S., Pauletto, D., Arevalo, F.R., Costa Reis, N.F., Sampaio, A.F., Urrego Giraldo, L.E., Valderrama Sandoval, E.H., Valenzuela Gamarra, L., Vela, C.I.A., ter Steege, H. Persistent effects of pre-Columbian plant domestication on Amazonian forest composition. **Science**, v.355, p. 4-7, 2017.

LI, H.; HANDSAKER, B.; WYSOKER, A. et al. The sequence alignment/map (SAM) format and SAMtools. **Bioinformatics**, Oxford, v.25, p. 2078–2079, 2009.

Lins, J., Lima, H. P., Baccaro, F. B., Kinupp, V. F. Pre-Columbian Floristic Legacies in Modern Homegardens of Central Amazonia. **PloS one**, v. 10, p. 1–10, 2015.

López-Gartner, G., Cortina, H., McCouch, S. R., Moncada, M. D. P. Analysis of genetic structure in a sample of coffee (*Coffea arabica* L.) using fluorescent SSR markers. **Tree Genetics & Genomes**, v. 5, p. 435–446, 2009.

Loveless, M. D., and J. L. Hamrick . Ecological determinants of genetic structure in plant populations. **Annual Review of Ecology and Systematics**, v.15, p. 65 – 95, 1984.

Korunes, Katharine L.; Goldberg, Amy (2021). «[Human genetic admixture](#)». Public Library of Science (PLOS). *PLOS Genetics*. **17** (3): e1009374. [ISSN 1553-7404](#). [PMC 7951803](#)  [. doi:10.1371/journal.pgen.1009374](#)

Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: population inference from RADseq data. *Molecular biology and evolution*, 35(5), 1284-1290.

Martins, P. S. Biodiversity and agriculture: patterns of domestication of Brazilian native plant species. **Anais da Academia Brasileira de Ciências (Brazil)**, v. 66, p. 219–224, 1994.

Martins RL, Gribel R. Polinização de Caryocar vill soma (Aubl.) Pers. (Caryocaraceae) uma árvore emergente da Amazônia Central. **Revista Brasileira de Botânica**, v. 30, p.35–43, 2007

Mignone, F., Gissi, C., Liuni, S., Pesole, G. Untranslated regions of mRNAs. **Genome Biology**, London, v. 3, doi:10.1186/gb-2002-3-3-reviews0004, 2002.

Milach, S.C.K. **Marcadores moleculares em plantas**. Porto Alegre: UFRGS. P. 140, 1998.

- Miller A.J, Gross B.L. From forest to field: perennial fruit crop domestication. **American Journal of Botany**, v.98, n.9, p. 1399-1405, 2011.
- Moreira, P.A., Lins, J., Dequigiovanni, G., Veasey, E. A., Clement, C. R. The Domestication of Annatto (*Bixa orellana*) from *Bixa urucurana* in Amazonia. **Economic Botany**, New York, p. 1-9, 2015.
- MONDINI, L.; NOORANI, A.; PAGNOTTA, M.A. Assessing plant genetic diversity by molecular tools. **Diversity**, Basel, v.1, p. 19-35, 2009.
- Muhlen, G.S., Pereira, A.A., Carvalho, C.R.L., Junqueira, A.B., Clement, C.R., Valle, T.L. Genetic diversity and population structure show different patterns of diffusion for bitter and sweet manioc in Brazil. **Genetic Resources and Crop Evolution**, v. 66, p. 1173-1790, 2019.
- NASCIMENTO, A. S. et al. Genomic consequences of açai extraction in the Amazon: insights into selective pressures, genomic diversity, and population structure. **Frontiers in Plant Science**, [S. l.], v. 15, 2024.
- Pedrosa, H.C., Clement, C.R., Schietti, J. The Domestication of the Amazon Tree Grape (*Pourouma cecropiifolia*) Under an Ecological Lens. **Frontiers in Plant Science**, v. 9, p. 1-14, 2018.
- Pereira, A.A. Diversidade genética, genômica e filogeografia de mandioca (*Manihot esculenta* Crantz): implicações para a dispersão do cultivo ao longo dos principais eixos fluviais da bacia amazônica brasileira. Tese de Doutorado em Ciências. P. 139. Universidade de São Paulo Escola Superior de Agricultura, São Paulo, 2015.
- Pereira, A.A.; Clement, C.R.; Rodrigues, D.P.; Veasey, E.A.; Dequigiovanni, G.; Ramos, S.L.F.; Pinheiro, J.B.; Zucchi, M.I. Patterns of nuclear and chloroplast genetic diversity and cassava structure along the main Brazilian Amazon rivers. **Annals of Botany**, v.121, p 625-639, 2018.
- POWELL, W.; MACHRAY, G.C.; PROVAN, J. Polymorphism revealed by simple sequence repeats. **Trends in Plant Science**, London, v. 1, p. 215-222, 1996.
- Prance, G.T., and Silva, M.F.d. A monograph of Caryocaraceae. **Flora Neotropica**, v. 12, p. 1-75, 1973.
- Ridley, Mark *Evolução*/ Mark Ridley ; tradução Henrique Ferreira, Luciane Passaglia, Rivo Fischer. – 3. ed. – Dados eletrônicos. – Porto Alegre : Artmed, 2007. Editado também como livro impresso em 2006. ISBN 978-85-363-0863-0 1. Biologia – Evolução. I. Título. CDU 575.8 Catalogação na publicação: Juliana Lagôas Coelho – CRB 10/1798. P. 156.
- Rindos, D. The origins of agriculture: An evolutionary perspective. **Academic Press**, San Diego, 1984.
- Seeb, J.E., Carvalho, G., Hauser, L., Naish, K., Roberts, S., Seeb, W. Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. **Molecular Ecology Resources**, Oxford, v. 11, p.1-8, 2011.
- Shanley, P. and M. Galvão, *Frutíferas e plantas úteis na vida Amazônica*. Belém, CIFOR, Imazon, 2005.
- TANKSLEY, S. D.; MCCOUCH, S. R. Seed banks and molecular maps: Unlocking genetic potential from the wild. **Science**, v. 277, p. 1063–1066, 1997.
- Tuxill, J. Nature's Cornucopia: Our stake in plant diversity. **Worldwatch Paper**, Washington DC: Worldwatch Institute, p. 148, 1999.

- VAN OOSTERHOUT, C.; WEETMAN, D.; HUTCHINSON, W. F. Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. **Molecular Ecology Notes**, v. 6, p. 255-256, 2006.
- Weber, J. C., Montes, C. S., Vidaurre, H., Dawson, I. K., Simons, A. J. Participatory domestication of agroforestry trees: An example from the Peruvian Amazon. **Development in Practice**, v. 11, p. 425–433, 2001.
- Willing, E.M., Dreyer, C., Van Oosterhout, C. Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many snps markers. **PLoS ONE**, San Francisco, v. 7, e42649, DOI:10.1371/journal.pone.0042649, 2012.
- YAMAGUCHI, K. K. L.; SOUZA, E. S. Uso de piquia (*Caryocar villosum*) como fonte de produtos biotecnológicos. **Revista Ensino Saúde e Biotecnologia da Amazônia**, v. 3, p. 18-27, 2021.
- LAHAYE, Christelle; HERNANDEZ, Marion; BOËDA, Eric; FELICE, Gisele D.; GUIDON, Niede; HOELTZ, Sirlei; LOURDEAU, Antoine; PAGLI, Marina; PESSIS, Anne-Marie; RASSE, Michel; VIANA, Sibeli. Ocupação humana na América do Sul até 20.000 a.C.: sítio Toca da Tira Peia, Piauí, Brasil. **Revista de Ciência Arqueológica**, Amsterdã, v. 6, pág. 2840-2847, 2013.
- MILLER, Eurico T. Adaptação agrícola pré-histórica no alto rio Madeira. In: MEGGERS, Betty J. (Ed.). **Pré-história sul-americana: novas perspectivas**. Washington: Taraxacum, 1992. pág. 219-232.
- ROOSEVELT, Anna C. A Amazônia e o Antropoceno: 13.000 anos de influência humana em uma floresta tropical. **Antropoceno**, Amsterdã, v. 69-87, 2013.
- GOUDET, J.; JOMBART, T. *hierfstat: Estimation and Tests of Hierarchical F-Statistics*. R package version 0.5-10, 2021. Disponível em: <https://CRAN.R-project.org/package=hierfstat>. Acesso em: 18 jun. 2025.
- JOMBART, T.; AHMED, I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, [S. l.], 2011. DOI: 10.1093/bioinformatics/btr521.
- KAMVAR, Z. N.; BROOKS, J. C.; GRÜNWALD, N. J. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, [S. l.], v. 6, p. 208, 2015. DOI: 10.3389/fgene.2015.00208.
- R CORE TEAM. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2021. Disponível em: <https://www.R-project.org/>. Acesso em: 18 jun. 2025.
- WICKHAM, H. et al. *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4, 2023. Disponível em: <https://CRAN.R-project.org/package=dplyr>. Acesso em: 18 jun. 2025.